Chronic Health Conditions and Economic Outcomes

YoonKyung Chung*
Research Fellow
Korea Energy Economics Institute
yoonie@keei.re.kr

October 2013

Abstract

With nearly half of the adult population afflicted with a chronic disease, the first onset of a chronic condition is a major life event. This paper examines the relationship between the onset of a chronic disease and economic outcomes among the working-age male population. Fixed effects models are estimated to quantify the temporal patterns of earnings, after-tax family income and food expenditures, controlling for unobserved heterogeneity across individuals. Also, a multiplicative fixed effects model is introduced to include zero earnings in the analysis. Using the Panel Study of Income Dynamics (PSID), I find that initial chronic illness onset has a persistent effect on earnings. In the short run, earnings losses are mostly driven by changes at the intensive margin, however, in the long-run, adjustments at the extensive margin also play a role. Family income responses are smaller than those of earnings. Food expenditures respond in a muted fashion, recovering over time. Finally, there exists significant heterogeneity in the health effects across education groups over time.

Keywords: Chronic Conditions, Labor Market Outcomes

JEL: I10, J17

*I am deeply indebted to A. Colin Cameron and Ann Stevens for guidance and support. I would like to thank Ankur Patel, Hilary Hoynes, Doug Miller, Meixin Guo and my colleagues at UC Davis for helpful comments and encouragement.

1 Introduction

ON SEPTEMBER 19TH-20TH 2011, a second-ever U.N. summit on health - a decade after the meeting on AIDS - was held in New York. The summit was dedicated to combating the growing incidence of chronic disease, especially in the developing world. In the U.S., chronic ailments are responsible for 70% of deaths and almost half of the adult population suffer from at least one chronic condition (Centers for Disease Control and Prevention (2009)). Heart disease, cancer and stroke are the three leading causes of death, and arthritis is the most common cause of disability. One in every 10 Americans suffers from major limitations in activity due to disabling chronic conditions.

These chronic conditions are costly. In 2007, approximately 75% of national healthcare expenditures were spent on overall chronic disease care (CDC (2009)). Even so, the biggest drain on the economy due to chronic disease is not the direct medical costs but the indirect costs of foregone economic output. According to a report by the Milken Institute, the total estimated "lost productivity" due to seven selected chronic conditions in 2003 was $1,046 billion; 4.5 times greater than the treatment costs of $277 billion (DeVol et al. (2007)).[1] This indirect cost of chronic disease is the focus of this paper.

This paper estimates the long-term effects of a negative health event on individual earnings, household income and consumption (food) expenditures. The negative health event of interest here is the onset of a chronic disease, such as heart disease, lung conditions, stroke, cancer, diabetes, hypertension and mental disorder. A chronic disease can be characterized by its persistent effect on health (a major determinant of one's labor productivity), the costliness of its treatment and its high prevalence rate among the adult population.

Most previous work has focused on the role of socioeconomic status (SES)[2] as a determinant of an individual's general health status (for example Goldman (2001) and Adams, Hurd, McFad-

[1]The seven chronic conditions are cancer, heart disease, hypertension, mental disorders, diabetes, pulmonary conditions, and stroke. The cost of "lost productivity" is estimated using the overall earnings loss due to lost workdays and presenteeisms for individuals with chronic disease.

[2]Such as income, wealth and education.

den, Merrill and Ribeiro (2003)). Less attention has been paid to the impact of specific individual health status events, including the onset of a chronic illness, on SES. This line of inquiry can be challenging as it requires detailed condition information, including type and onset date. Studies that do analyze the effect of chronic conditions on SES for the U.S. population have focused on the near elderly population using the Health Retirement Study (HRS) (see Wu (2003) and Smith (2003)) or the AHEAD (Smith (2003)). Unfortunately, the average age of onset among chronic disease is much lower than the common retirement age (columns 3-6 of table 1). In addition, more than half of the working-age male population are afflicted by at least one chronic illness (columns 1-3, table 1). Taken together, it is clear that examining the evolution of labor market outcomes for a working-age sample is helpful in understanding the economic impact of chronic diseases.

This paper quantifies the temporal effects of chronic disease incidence on the economic well-being of the working-age population. To do this, I use the panel structure of the Panel Study of Income Dynamics (PSID) to estimate a fixed effects model that allows the effect of the treatment event to vary over time. Here the treatment event is the first onset of a chronic illness. This estimation strategy was introduced by Jacobson, LaLonde and Sullivan (1993) and is used extensively in the job loss and disability literature (see Stevens (1997), Charles (2003), Stephens (2001), and Meyer and Mok (2008)). Unlike the specifications of Wu (2003) and Smith (2003), this model allows the measured chronic health event to have an impact prior to the event, and the inclusion of time fixed effects makes it possible to pool together the first chronic health events across time by re-aligning calendar time as event time. While zero earners are often dropped in the earlier disability studies using log-linear conditional mean models, I employ a multiplicative fixed effects model for earnings to include those with zero earnings in the analysis, incorporating those who may be hardest hit by chronic disease.

I begin by examining the direct impact of chronic disease on individual[3] earnings through changes in labor supply. This effect could be amplified or dampened at the household-level depending on the magnitude, persistence and the predictability of the shock, as well as household

[3]Individual is defined as household head in the PSID.

characteristics. Thus I also study after-tax family income, food expenditure responses and potential insurance mechanisms to examine whether and how households can insure against these idiosyncratic health events over time.

I find that the first chronic health event has a statistically significant impact on earnings, family income and food expenditures. Individual earnings are reduced by 12% at the time of the event, and fall further to a 18% decline in the long-run. Earnings loss around the time of the disease onset is mostly driven by reductions in working hours (intensive margin) rather than by labor market withdrawal (extensive margin). In the long-run, adjustments on the extensive margin also play a role; six or more years after the onset, the probability of employment is decreased by almost 6 percentage points. Considering that by age 50 approximately one half of the male population has a chronic condition, the estimated loss in aggregate earnings may be large.

I find clear, negative trends in labor market outcomes prior to the onset of a chronic condition, even in the presence of individual fixed effects. These pre-onset trends are concentrated among those with at least one year of education exceeding a high school degree and are unaffected by changes in sample age groups, family structure, and severity of illness. Individuals with a high school degree or less experience a persistent drop in earnings at the onset of a chronic condition, while after-tax family income partially recovers.

I do not find any reliable sources of insurance that buffer family income against idiosyncratic earnings shocks due to disease incidence, aside from social security disability benefits. Food expenditure responses are small with a 4% fall at the time of the event and they seem to recover in the long run.

The biggest challenge in identifying the effects of chronic health events is the issue of whether the onset of chronic illness can be considered an unpredictable random event. Individuals have private information that is unavailable to researchers, such as information on family medical history and their health behaviors. Therefore, these health events may not be random and may be a consequence of unobservable characteristics. If such confounding factors also directly affect the outcome variable, their omission in the regression equation would cause a bias in the estimates (omitted variable bias). Also, the direction of causation may be reversed, as emphasized in the epidemiology literature: Poor prior labor market outcomes may increase the likelihood of chronic health events in the future. For example, involuntary job loss may increase the individual's stress level and the likelihood of engaging in risky health behaviors, leading to an escalated risk of chronic disease incidence, such as heart attack or stroke (See Gallo, Teng, Falba, Kasl, Krumholz and Bradley (2006)). This would imply that a chronic health event is not a shock but an outcome predictable by prior poor outcomes (SES). In this paper, the longitudinal dimension of the data allows for individual-specific fixed effects that absorb permanent person-specific characteristics, such as genetic factors and time-invariant risky health behaviors. The individual fixed effect incorporates past outcome information, as long as there is a sufficient time lag.

Measurement error (or the mismeasurement of the timing and type of chronic illness), based on self-reports of past diagnosis, represents another challenge to identification. For example, conditional on having a chronic disease, those of low SES may be less likely to be diagnosed, due to limited access to health care (under reporting, see Chatterji, Joo and Lahiri (2010)).[4] At the same time, those individuals of lower SES with poorer labor market prospects have greater incentives to seek disability benefits (see Autor and Duggan (2003)), which may lead to a higher propensity to report a "marginal" chronic illness, as well as increasing the likelihood of false reporting (Baker, Stabile and Deri (2004)). While under-reporting would lead to a downward bias in the estimates, "over reporting" would lead to an upward bias; hence, the direction of the bias cannot be determined.[5] The potential impact of reporting errors on estimates are not directly addressed in this paper.

This paper is divided into seven sections. Section 2 reviews the relevant literature. Section 3 describes the data and section 4 presents the main specification with details of the estimation method used in the analysis. It also introduces an alternative specification: the multiplicative fixed-effects

[4]On the other hand, Cutler and Lleras-Muney (2006) report that those with more education are less likely to self-report a past diagnosis.

[5]These are examples of non-classical measurement error, where measurement errors are correlated with other (potential) control variables in the regression or/and the variable measured with error is an imperfect proxy. A more complete analysis of non-classical measurement error in chronic health conditions is being pursued in a companion paper.

regression model for earnings. Section 5 reports the main results and compares these results with those from the disability literature. Section 6 extends the analysis to severe chronic conditions. Lastly, section 7 concludes.

2 Literature Review

Many individual health studies have focused on explaining the well-documented positive association between socioeconomic status (SES) and health (see Marmot (1999), Cutler, Lleras-Muney and Vogl (2008)).[6] It is only recently that economists have begun to explore the pathways from health to SES, typically for the older population (see Smith (2003) and Adams et al. (2003)). This paper is part of that growing literature, examining the effects of health deterioration on labor market outcomes and economic well-being.

The ideal way to identify the effects of health would be to randomly assign a negative health event to individuals and then measure the difference in outcomes between the treated and non-treated. Since this is infeasible, researchers have tried to find other sources of identifying variation, particularly health measures or health events that are plausibly exogenous.

Smith (1999) was the first to recognize the onset of new chronic condition as a conditional "health innovation." Smith (2003, 2004, 2007) attempts to isolate the unanticipated part of chronic illness onset by controlling for predictors of a new health event, such as behavioral risk factors, baseline health and economic conditions. Using the first five waves of the HRS, the author finds that a major health shock between the first two waves reduces the probability of work by 15 percentage points,[7] while household income loss amounts to $4,033 over this period. Wu (2003), using the same dataset for married couples, finds a strong impact on household wealth through lowered earnings. The author points out that these changes in earnings are due to retirement rather

[6] Those with lower SES tend to have worse health outcomes, such as increased mortality (Adler et al.(1994)). Here the measures of SES include but are not limited to wealth, income, education, occupation and race.

[7] This is a similar finding to Datta Gupta, Kleinjans and Larsen (2011).

than reduced work hours. Neither of their estimation strategies allow for pre-onset effects. The result is that if there are important health effects prior to the onset of a related, new chronic condition, their estimates may be biased downward. They consider not only the first onset but any new onset as a health shock. In other words, the comparison or control group includes not only those who are always healthy but those who have already experienced a chronic health shock. It is likely that differences between those who always stay healthy and those who already have a history of chronic conditions go beyond time-invariant unobservables or baseline health measures. Lastly, as mentioned above, both papers focus on older workers.[8]

A recent paper has extended the previous literature by examining the effects of health shocks on earnings by education and age for the entire working-age Swedish population (Lundborg, Nilsson and Vikstrom (2011)). Using unanticipated hospital admissions in register data as health shocks, the authors stratify the sample by three age groups and estimate a fixed-effects model. They allow the effects to differ across two education groups by interacting education and an event time indicator. They find a negative impact that is almost twice as large for those within the low education group across all ages at the year of the health shock.[9] Their identification strategy relies on the condition that the relative earnings trends between high and low educated individuals are similar across treatment status prior to the shock.

Instead of controlling for pre-onset trends, the methodology below imposes no restriction on pre-event effects,[10]. At the same time, this strategy makes use of all initial chronic health events across time. This approach is often used in the disability literature to examine the effects of disability on economic outcomes.[11] To a certain extent, this paper can be viewed as an extension of the disability literature, since chronic illness is one main cause of disability. Since the same estimation

[8]Other popular candidates for exogenous health events are accidents and injuries. Dano (2005) uses road injuries and Crichton, Stillman and Hyslop (2005) accidental injuries as exogenous changes in individual health. Riphahn (1998) defines health shocks as a sharp and sudden drop in an individual's health satisfaction.

[9]Labor earnings are reduced by 5-6 percent for those with high education and 9-12 percent for those with low education. When the health shocks are classified into the ten most common diseases, they find that while the estimated effects do vary significantly across types of health shocks, the differences in the effects between high and low education group are similar across all types.

[10]This estimation strategy follows Jacobson, LaLonde and Sullivan (1993).

[11]Examples include Charles (2003), Meyer and Mok (2008) and Stephens (2001).

methods are used, the effect of chronic illness onset on the outcome variables can be compared to that of disability. It should be noted that there is a key difference between chronic illness and disability: Disability is generally defined in terms of work limitations, while chronic illness is a persistent health condition diagnosed independently of work capacity.

Stephens (2001), Meyer and Mok (2008) and Charles (2003) use the PSID to study the effect of disability. Stephens (2001) and Meyer and Mok (2008) find earnings decline by 10% to 15% around the time of the disability onset, which in turn drops by a further 22% in the long run for the overall disabled group.[12] Meyer and Mok (2008) also categorize disability into four groups based on duration and self-reported severity. Out of the four groups, the estimated effects on the Chronic-Severe group could be understood as providing the upper bound for the chronic illness estimates.[13] This Chronic-Severe group experiences a steep 32% drop in earnings at the year of the onset, relative to a 15% fall for the overall group. Also, there are conspicuous pre-disability trends in earnings, where earnings fall continuously by 10%-17% leading up to the onset. Lastly, all three papers differ in their treatment of zero earnings. I introduce a more natural way of including zero earnings in the analysis in Section 4.2.

As expected, the onset of disability is found to have a smaller effect on family income. For example, Stephens (2001) finds that family income drops by 7% at the time of the shock and by 14% by the sixth year. Food expenditure responses are even smaller, with no statistically significant fall at the time of the shock, and a modest 6% fall in the long run. In a similar vein, the family income response of the Chronic-Severe group of Meyer and Mok (2008) is roughly three quarters the size of the reduction in earnings at the time of the onset. Food expenditures for this group have a 2% decline at the onset, and a 9% decline by the tenth year.

[12] Unlike Stephens (2001) and Meyer and Mok (2008), Charles (2003) controls for individual-specific linear time trends. Charles (2003) report similar drops in annual earnings at the disability onset, around 10% - 17% (depending on different disability histories). Contrary to the findings of Stephens (2001) and Meyer and Mok (2008), earnings recover over the next two years. These contradictory findings, however, were later reconciled in Mok, Meyer, Charles and Achen (2008). They replicated Charles (2003) using the same dataset and found the earnings loss to be much larger and persistent than previous findings.

[13] Note that the term Chronic-Severe is based on work disability and not on the presence of chronic conditions: Those who report disability at least three times during the ten year span after the first onset and suffer from severe work limitation during half of those times.

There is not much written on the impact of health shocks per se, rather than disability, on consumption for the U.S population due to a lack of appropriate datasets that contain both consumption and health measures. Most work has been done for the developing world. Islam and Maitra (2008) and Mohanan (2011) find that households can smooth consumption against health shocks. For example, Islam and Maitra (2008) find no statistically significant effect of illness on both food and non-food consumption in Bangladesh.

3 Data

3.1 General Sample Restrictions

In this study, I use 36 waves of the Panel Study of Income Dynamics (PSID), a nationally representative longitudinal survey, spanning the years 1968-2009.[14] In 1997, the PSID changed from an annual survey to a biennial survey, which is a complication that is addressed in section 4. I restrict to male heads to avoid complications in life-cycle labor force outcomes unrelated to health. In addition, only those ages 25-65 are kept to focus on the working-age population. To eliminate severe outliers and avoid imputing top-coded data, those in the top 0.1 percentile of the distribution in any of the outcomes are excluded from the analysis.[15] Survey years 1968, 1973, 1988 and 1989 are excluded from the analysis because no food expenditure data was collected. Those with missing values for earned income, those with negative or zero before or after-tax family income, and those with zero total food expenditure are also dropped. Finally, individuals have to be present in at least three consecutive surveys to be included in the sample.

[14] For more details regarding the data used in this study, please see the attached Data Appendix.

[15] The PSID is not as heavily top-coded as other household surveys, e.g the Current Population Survey (CPS). Nevertheless, the outcome variables are top-coded and the value of these top-codes has changed over the last four decades.

3.2 PSID Health Information: Measurement of Chronic Illness

The treatment variable here is the onset of a chronic health event. Since 1999, the PSID has been asking questions regarding the type and timing of chronic conditions: Cancer, heart attack, heart disease, lung disease, stroke, arthritis, diabetes, hypertension, psychological problems, asthma, memory loss, learning disability and, since 2005, other chronic conditions. An individual is considered to have experienced a chronic health event if he was ever diagnosed with any one of the twelve chronic conditions above, including other chronic conditions from 2005 onwards. The onset of the chronic health event is defined as the year of the earliest diagnosis across these chronic conditions.[16] In section 6, the analysis is restricted to major chronic conditions since 1) the measurement error in reporting is lower; and 2) they are likely to have a larger impact on labor market income.[17]

Because the information on chronic illness is only available beginning in 1999, those who leave the survey before 1999 are excluded. Those who had the opportunity to report their health information but failed to do so for any year during 1999-2009 and those whose date of onset is missing are also excluded. To minimize the recall bias common in retrospective questions, those whose earliest chronic disease diagnosis dates back ten or more years from their first health report are excluded from the analysis.[18] The earliest chronic health event in the remaining sample occurs in 1990. Because the coefficients of interest are identified by those whose health status change from healthy to chronically ill, individuals in the chronically ill sample are required to have at least one observation prior to the onset.[19] Here the chronic health shocks of interest are the ones that occur during one's working life, therefore, those individuals who got the shock after age 65 are considered as healthy.

In the end, the sample contains a total of 2,315 men, between ages 25-65; either single or

[16] For information on the construction of onset date, see Data Appendix.

[17] Major conditions, following Smith (2003), are classified as cancer, heart disease, heart attack, lung disease and stroke.

[18] See the Data Appendix for implications of these restrictions.

[19] This implies that those who first entered the sample and experienced their first chronic health event in 2009 would be dropped automatically.

married. Out of 2,315 individuals, 1,152 men are in the chronically ill sample. The chronically ill sample is made up of those who had experienced their first chronic condition at some point within the last 10 years from their first health report, during their working life. The onset years range from 1990 to 2009. The remaining individuals are healthy: Those who have never been diagnosed with a chronic illness in their working life. The healthy are included in the analysis to help to estimate a representative counterfactual path of earnings and income in the absence of a health shock.

3.3 Measurement of Other Variables

The main outcomes analyzed are earnings,[20] after-tax family income,[21] and food expenditures. I use NBER's TAXSIM program to calculate federal income taxes for all tax years between 1968 and 2008.[22,23] Dollar values are converted to 2009 dollars using the Consumer Price Index for All Urban Consumers(CPI-U).[24] The basic control variables in the main regression are individual-fixed effects, age dummies and marital status. Finally, individual is categorized as high education if he has more than 12 years of schooling.[25]

4 Methodology

4.1 Basic Model Specification

To study the temporal effects of chronic health shocks, I estimate a fixed effects model similar to Jacobson et al. (1993), Stephens (2001) and Meyer and Mok (2008). The basic estimation equation

[20]Head's earnings is labor income, including labor income from farming and self-employment.

[21]Family income is a measure of all income sources of a household, including head's, wife's and other member's earnings, capital income, public and private transfer income, and social security income.

[22]The PSID provides a measure of federal income taxes up to 1991.

[23]Married couples are assumed to file jointly, while those in a domestic partnership are assumed to file separately. State income taxes are not taken into account.

[24]The food expenditure variable consists of the amount spent on food used at home, away from home and the net value of food stamps. Following Stephens (2001), each food component is converted into 2009 real dollars using the seasonally adjusted first quarter food component of the CPI-U. Note that in later years, a number of PSID interviews took place during other quarters.

[25]I follow Stephens (2001) in constructing an individual's years of education. The value of the education variable for each individual is the most recently reported years of completed schooling.

is as follows:

$$ln\, y_{it} = \alpha_i + \gamma_t + X_{it}^{'}\beta + \sum_{k=-p}^{K} \delta_k D_{it}^k + \varepsilon_{it}, \tag{1}$$

where $ln\, y_{it}$ is the natural log of outcome variable for individual i at time period t. The range of periods t, correspond with PSID waves; α_i represents individual fixed effects; γ_t represents calendar year effects; and X_{it} is a vector of time-varying control variables (age dummies, marital status indicator). The vector D_{it}^k is a set of event time indicator variables with $k = -p, ...0..., K$. When $k > 0$, D_{it}^k is equal to 1 if, at time t, it has been k period(s) since the health event. When $k < 0$, the event time indicator is equal to 1 if, at time t, it is k period(s) before the onset. All regressions are weighted using PSID-provided longitudinal weights. Reported standard errors are cluster robust where clustering is on the individual. This allows for heteroskedastic errors across individuals and serial correlation within individual i.

Each "period" spans two years. The PSID became a biennial survey in 1997 and began asking questions regarding health in 1999. A two-year time interval does not trace out the temporal effect in as much detail as one-year time interval, but converting the two-year interval into a one-year time interval would divide identification into separate groups for years 1999-2009: Odd post health event years would be identified by one group and the even post-event years by the other group. A two-year time interval ensures that the whole chronically ill sample is included in the estimation of each event time coefficient of interest. The omitted event time category is more than 6 years before the disease incidence and the last event time dummy includes all 7+ post onset years.

The vector of interest is δ_k. The identification of δ_k hinges on the assumption that conditional on individual fixed effects, year effects and other observable controls, the onset of chronic illness is a random health event unpredictable to an individual several periods prior to the onset.[26] This implies that had this health event not occurred, the conditional average log-outcome of the treated would have had the same "post-shock" trend as that of the always healthy (non-treated). This

[26]Specifically, conditional on the controls, the onset of first chronic illness does not have to be an unanticipated event at the time of the onset, but seven or more years prior to the onset.

identification assumption is not directly testable. However, the pre-onset trend traced out by δ_k provides information regarding the extent to which the shock was anticipated.

The above identification assumption is violated if there are confounders that are correlated both with the outcome and the health event. Since any permanent differences between individuals are absorbed by the individual fixed effects, and common systematic shocks are absorbed by calendar year effects, these confounders would have to be unobservables that vary across time and individuals. For example, changes in risk preferences across the treatment population that change both the probability of experiencing a chronic illness and work effort would lead to a bias in the estimates.[27]

As long as the identifying assumption is not violated, δ_k measures the differences in outcome growth rates for each period due to a health event: They measure the conditional average percentage change in the outcome variable k periods from the event relative to more than six years prior to the chronic illness onset. The vector δ_k can also be interpreted as identifying the deviation in the dependent variable at period k from the level the individual would have achieved had he stayed healthy.

One concern is that studying the average effects of health shocks may be less interesting if health shocks have (substantially) differential impact across individuals. For example, if those with higher education are better able to cope with chronic illness than those with lower education,[28] then those with high education may fare relatively better in the labor market after their chronic disease onset. The average effects would mask the presence of such SES heterogeneity in the effects of health shocks (see Lundborg et al. (2011)). To explore this, I estimate equation (1) for each education group: Those with more than 12 years of schooling (high) and those with equal or less than 12 years of schooling (low).

I stratify the sample (both treated and control) by education, rather than interact onset indicators with education, to control for differential time and age trends across education groups in

[27] Another example are changes in family composition, such as repeated divorce or separation. These changes may increase the risk of depression or anxiety disorders while decreasing family resources (see Richards, Hardy and Wadsworth (1997)) also biasing the estimates (upward). To address this particular concern, I include marital status in the regression as a covariate.

[28] See Goldman and Smith (2004) for the case of AIDS.

the most flexible manner. For example, changes in observed earnings inequality over time are very different between education groups (see Autor, Katz and Kearney (2008)). Hourly earnings inequality between college educated and high school graduates has grown considerably from the late 1970s to the early 1990s. Moreover, the college-high school wage differential grew faster for younger workers during this period (Card and Lemieux (2001)). These facts justify the need to consider year effects and age effects separately for each group.[29]

4.2 Alternative Specification for Earnings

When analyzing the temporal effect of health shocks on earnings, the log-linear model above necessitates excluding individuals with zero earnings, and only estimating the impact for those who participated in the labor market. If zero earnings are the result of a labor supply choice driven by a health event, then the log-linear coefficient estimates may be biased. In particular, the omission of labor force exits would bias δ_k toward zero.

The treatment of this issue in the literature varies. Stephens (2001) employs a log-linear specification to estimate the percentage changes in earnings: All those with zero earnings are simply excluded from the analysis. Charles (2003) also models only those with zero earnings. He adds a Mills ratio term to correct for sample selection. This method, however, relies on strong distributional assumptions and has not been theoretically justified in the case of short panel models with the added complication of fixed effects. Meyer and Mok (2008) include zero earners using a model in levels instead of logarithms, and compute the implied percentage change manually by dividing the estimated coefficients by the mean earnings of the treatment group in the omitted pre-disability periods.

Departing from previous work, I estimate a multiplicative fixed effects (FE) model to include those with zero earnings. Assume a multiplicative model in levels of earnings,

$$E[y_{it}|Z_{it}, v_i] = \exp(v_i + Z_{it}^{'}\theta) = \delta_i \exp(Z_{it}^{'}\theta), \qquad \text{where } \delta_i = \exp(v_i) \tag{2}$$

[29] Lundborg et al. (2011) include linear time trends that vary across education and treatment status in their regression to control for the differential pre-shock earnings trend.

Equivalently, in natural logarithms,

$$ln\,E[y_{it}|\mathbf{Z}_{it}, v_i] = v_i + Z_{it}'\theta, \tag{3}$$

where Z_{it} is a vector of time fixed effects, time-varying control variables and onset time indicators. This setup permits observations with zero earnings. The use of exponential conditional mean in yr models data that are right-skewed and heteroskedastic.[30] Furthermore, the model can accommodate fixed effects, which are eliminated by quasi-differencing, allowing the parameters of interest to be consistently estimated. Technically, the estimation procedure is the Poisson fixed effects model, but this method does not require count data and inference can be based on cluster and heteroskedastic robust standard errors. After quasi-differencing, the moment conditions become (see Blundell et al. (2002)):

$$E\left[y_{it} - \frac{\exp\left(Z_{it}'\theta\right)}{\frac{1}{T}\sum_{t=1}^{T}\exp\left(Z_{it}'\theta\right)}\bar{y}_i|\, Z_{it}\right] = 0 \tag{4}$$

where $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$. The corresponding sample moment conditions for estimating θ are,

$$\sum_{i}^{N}\sum_{t}^{T} Z_{it}\left(y_{it} - \frac{\exp\left(Z_{it}'\theta\right)}{\frac{1}{T}\sum_{t=1}^{T}\exp\left(Z_{it}'\theta\right)}\bar{y}_i\right) = 0, \tag{5}$$

which are the first-order conditions of the Poisson fixed effects MLE. As already noted, inference is based on cluster and heteroskedastic robust standard errors.

If zero earnings are a mechanical effect of chronic disease incidence, i.e. temporary exits from the labor market directly due to health shocks, then the multiplicative FE model is preferable since zero earnings are incorporated into the impact of a chronic health event. On the other hand, if zero earnings are largely the result of a behavioral response in the face of a health shock, e.g: voluntary early retirement, then including zero earnings is problematic. In this case, counterfactual changes in earnings growth for early retirees (had they continued to participate in the labor market after the event) are not observable. In order to find out whether the estimates are sensitive to early retire-

[30] See Mullahy (1998), Manning and Mullahy (2001) and Silva and Tenreyro (2006).

ment, I restrict the analysis to those under age 55.

The estimates from equation (2) capture the percentage change in expected earnings. The LSDV estimates from equation (1) capture the average percentage change in the earnings relative to more than six years prior to the health event. If we assume log earnings $\sim N(x_{it}^{'}\beta + \alpha_i, \sigma_i^2)$, then it can be shown that the two models are equivalent (in the case of only positive earners) where the differences between the two models are absorbed by the individual fixed effects.

5 Results

5.1 Background on Chronic Disease Incidence

Figure 1 compares the number of heads who are chronically ill to the number who are healthy by age group, using the 5 waves of the PSID between 1999-2009.[31] The number of those with chronic illness rises with age. Though the number of those healthy outnumbers those with a chronic illness in the age bins between ages 25-39, almost 30% have a chronic illness. This fraction quickly grows and at peak earning years, ages 45-49, there are almost as many with a chronic illness as there are healthy individuals. By the time of retirement, the number with chronic illness exceeds the healthy by approximately three fold. Not only are chronic illnesses an issue for the elderly, but they also affect those of prime working age.

Major chronic conditions are concentrated among those most likely to be retired, as exhibited in figure 2.[32] The number of those with a major illness increases with age. However, the percentage of those afflicted in each age group is relatively small (for instance, 9% in the 45-49 age bin) and the percentage in the younger age bins is negligible. Due to the small number of individuals with a major condition in the sample, the main analysis focuses on the broadest set of chronic conditions. I include a rich set of age controls to adjust for observed age effects.

[31] This is based on male heads of all ages.

[32] The major illness category is composed of cancer, heart attack, heart disease, lung disease and stroke.

The prevalence rate and the average onset age of each of the 12 chronic conditions for the male working-age population in 2009 are reported in table 1. The fraction of working-age males with at least one or more chronic condition is 53.7%; higher for less educated (56.9%) and lower for more educated (51.5%). The five most common chronic diseases are hypertension, arthritis, asthma, diabetes and psychological problems for both education groups.[33] For major chronic illnesses, the most common ailment for the less educated is lung disease (4.5%), while for those with more education, it is cancer (4.7%). Those with less education are more likely to have one of the 12 chronic conditions than those with high education, except for asthma and cancer. The average onset age for each chronic condition is also lower for those with less education with the exception of lung disease.

5.2 Summary Statistics

Table 2 presents summary statistics for those who have never experienced a chronic health event (never chronic) and those who have experienced at least one chronic health event (ever chronic) during the sample period. Surprisingly, the observed socioeconomic characteristics for the two groups are strikingly similar. They are, on average, almost identical in age, years of education, the proportion white, the percentage married, and the number of children. Similar shares of individuals are currently working. Never chronic individuals are slightly more likely to be self-employed or a salaried worker. However, the average head's earnings are higher for the never chronic, relative to those ever chronic. Family income is also approximately 7% higher for the never chronic group, but food expenditures are similar across two groups. Because the ever chronic are in the sample for a longer period of time relative to the never chronic, the ever chronic group makes up 1.5 times more observations than the never chronic (despite making up 50% of the individuals in the sample). To summarize, the observed characteristics look similar across the two groups, while the outcomes differ. For example, the differences between the average earnings of the two groups may be partly

[33] Since those whose first chronic disease incidence dates back more than 10 years are excluded from the analysis, the percentage with asthma or with a learning disability are much smaller in the final sample; 2.8% for asthma and 0.5% for learning disability.

due to permanent unobserved heterogeneity that could be controlled for by including an individual fixed effect.

5.3 Earnings

a *Main Results*

Table 3, column (1) contains the benchmark fixed-effects regressions for individual (household head) earnings. There exists a statistically significant downward trend in the growth of expected earnings prior to the health event; a 3.3% fall in 5-6 years before the illness, which increases to a 7.3% drop during 1-2 years before the onset. The largest drop in earnings does occur around the year of the event, with head's expected earnings falling by a statistically significant 12.2% relative to no disease incidence. This negative impact varies little over the next 5 years but declines further to 18% of the expected earnings in the sixth year and after. The overall post-shock effect, measured as the average percent change in the expected earnings relative to seven or more years before the onset is estimated to be a 13% drop.

There is a downward trend in earnings growth prior to the onset of chronic illness. This can be a sign that factors other than the health event are causing earnings to depart from its expected trajectory.[34] In order to find out whether the estimates are sensitive to changes in family composition beyond divorce/separation or death of a spouse, changes in the number of young kids under 6, the number of all children, and the family size are added to the main regression. The results are presented in column (2) of Table 3. The magnitude and the precision of the estimates are very similar to that of the main regression. Next, the analysis is restricted to those under 61, then to those under

[34]Some of the exhibited downward trend could be accounted for by a delay between the actual onset date and the date of diagnosis. Since the diagnosis can only be obtained after a visit to a health professional, there is a gap, however small, between the actual onset of chronic illness and its diagnosis date. A systematic delay between the two could produce an earlier drop in earnings. Even if there is no delay in diagnosis, it is unlikely that the afflicted had been in perfect health prior to the event. A more likely scenario is that the individual's health had been deteriorating for some time and chronic illness onset is just a product or by-product of this decline in health status. Self-assessed health status reports show the chronically ill sample to be, on average, in poorer health than the healthy sample even prior to the event. The chronically ill sample has a higher mean in self-reported health status (2.0) than the healthy sample (1.8). Self-reported health status ranges from 1 to 5 in the order of decreasing health status with 1 indicating "excellent" and 5 "poor." This suggests that the pre-onset coefficients may be reflecting the effects of prior health deterioration on individual's earnings.

56, to focus on those less likely to be retired. If the estimated negative impact of chronic illness in column (1) of Table 3 had been identified mostly by early retirees, we are likely to observe smaller coefficient estimates once the analysis is focused on younger cohorts. However, the estimates are robust to age restrictions. In fact, the effects appear to be slightly larger for those under 56; for those who are in their prime working age. To summarize, family structure and age cannot explain the pre-onset trends.

However, the pre-onset trend in individual earnings is only present for those with high education. Table 4 contains estimates where the sample has been stratified on education group.[35] The low education group experiences a sharp 7% earnings drop at the onset, and earnings continue to fall to 14% in the long run.[36]

Also note that some of the pooled post-onset effects are not bounded by the estimates of low and high education groups. This can be explained by the rise in between-group inequality observed from late 1970s to early 1990s.

Between the late 1970s through the early 1990s, then again between the late 1990s and the 2000s, within-group wage inequality steadily increased for those with a college degree (Lemieux (2006)).[37] Moreoever, inequality in the top end of the earnings distribution increased dramatically during this period (Piketty and Saez (2003)). These time periods coincide with my sample period and they also overlap with the range of onset years for the chronically ill sample (1990-2009). If, within the high education group, individuals with a history of low wage growth are more likely to experience a health shock than high wage growth workers, then the event time indicators would be correlated with (unobserved and time-varying) less favorable ”macro” shocks faced by the low

[35]Such disparities in the pre-trend may be attributed to between variation in the types of chronic illness across education group. For example, the less educated might be more likely to be afflicted with chronic conditions that are less predictable and sudden in nature, such as heart attack, stroke or cancer. However, table 11 indicates no marked difference in the chronic condition types between the two education groups: the three most common chronic diseases are hypertension, arthritis, and diabetes for both groups; though heart attack is more common among the less educated, stroke and cancer are more common among the more educated.

[36]A graphical representation of the estimates, plotting the coefficient estimates against event time, is given in Figure 3.

[37]For those with high school or less, within-group wage inequality did not increase as dramatically, and reached a plateau after the late 80s and even decreased after the mid 1990s (Lemieux (2006)).

wage growth earners.[38] In other words, the coefficients on event time indicators would partly reflect the negative effects of these macro shocks. When I drop the top 1% earners, the size of the observed pre-illness trend is roughly reduced by half (2.4%-3.8% drops in earnings) and the precision of the estimates increase.[39]

The differential growth in hourly earnings across education groups over this period cannot be absorbed by the (common) year fixed effects of the main regression. The year fixed effects cannot control for group-specific macro shocks when the sample contains both low and high educated individuals. If less educated individuals are more likely to experience a negative health shock than the better educated, then the effects of low education group-specific macro shocks over and beyond that of the overall population would be reflected in the coefficient estimates of the post-illness time indicators; the estimated would be again biased.

Even after stratifying on education we may still be worried about unobservable shocks that are correlated with the health event. I attempt to address these concerns using individual-specific linear trends within each education group. These results are not reported here. For the low education group, the temporal pattern of earnings, though imprecisely estimated, is essentially unchanged. For the high education group, the signs of the pre- and post-onset effects are positive and not statistically different from zero. In conclusion, due to sample size issues I cannot use linear trends to absorb unobservable shocks to earnings for the high education group, but the results for the low education group support the results above.

b *Other Labor Market Measures*

The sharp drop in head's earnings at the time of the health event is not due to workers exiting the labor market. Table 5, columns 1 through 3 contain the impact of the health event on the extensive margin (employment). The decrease in the probability of being employed at the time of the shock

[38] These shocks are low wage growth group-specific shocks that would still be present in the absence of a health shock. In this case, unless these shocks are controlled for, the model suffers from an endogenity bias through model misspecification.

[39] These results are not tabulated here. The top 1 % are chosen because, out of those in the top end of the earnings distribution, they had experienced the most dramatic earnings growth during the period (Piketty and Saez (2003)).

is small and not statistically significant for both education groups. On the other hand, there is a change on the intensive margin (hours of work, table 5, columns 4 through 6). Annual hours of work for the low educated are almost 120 hours less than the expected levels, which translates into almost 3 weeks a year if the standard 40-hour work-week is assumed for pre-onset periods. For the high educated, annual hours losses are much smaller (just over 1.5 weeks a year) and statistically insignificant.

In the long run, groups experience a decrease in employment and hours worked. In particular, there is a large decrease in employment (9 percentage points) and hours of work (242 hours, almost 1.5 months) among those in the low education group. These outcomes help to explain the long run decline in individual earnings growth.

c *Relation to Disability*

The estimates above are generally smaller than those found in the disability literature. Though Stephens (2001) finds that head's earnings fall by 10% at the time of disability onset, by the fourth year, earnings are more than 20% lower than the expected level. Meyer and Mok (2008) find earnings decline by 15% at the time of disability onset and a 20% drop by the next year. Given that earnings of the disabled men in Stephens (2001) do not exhibit a pre-shock trend except in the year before the shock, a 12% drop in earnings observed at the time of chronic illness onset is likely to be smaller relative to that induced by disability once pre-onset trends are controlled for.[40]

The smaller effects for chronic illness, relative to disability, are to be expected. A broad measure of chronic conditions is used here and more people have a chronic condition than have a disability. Disability is defined as a health-related work-limiting condition. It would have to be severe enough to interfere with an individual's labor market prospects. Chronic illness has no such restriction. Having a chronic condition may well lead to a work-limiting disability but not necessarily so. One would expect more heterogeneity in the labor market outcomes of the chronically ill than the disabled at any given point in time. Nevertheless, disability estimates could provide an

[40]In contrast, in Meyer and Mok (2008), there is a clear pre-shock trend where the biggest decline occurs between $t-1$ and $t+1$.

upper bound for the chronic disease estimates. For example, Meyer and Mok (2008) divides the disabled group, depending on the length and severity of their disability, into several categories. The estimates for their "Chronic-Severe" group may be regarded as the upper bound of the temporal effects of having a chronic disease as long as the average life span after disability onset is long enough.

In fact, chronic illness turns out to be a good predictor of disability. To formally understand the relationship between chronic disease incidence and disability, the following fixed effects linear probability model is estimated:

$$y_{it} = \alpha_i + \gamma_t + X_{it}^{'}\beta + \sum_{k=-p}^{K} \delta_k D_{it}^{k} + \varepsilon_{it}, \tag{6}$$

where y_{it} takes on the value 1, if the individual is disabled at period t, or 0, otherwise; and D_{it}^{k} are again indicators for event times of chronic disease. Estimates of δ are reported in table 6. The onset of a chronic illness raises the probability of having a disability over time. Before the health event, the probability of becoming disabled is up 3 percentage points relative to more than six years before the shock. This probability more than doubles to 8 percentage points at the time of the shock, and increases over time, reaching 12 percentage points by 6 or more years after the event. There is no pre-existing trend in the effect until one period before the event. The changes in the likelihood of becoming disabled are much greater for the low educated group.

While these findings indicate that chronic disease is one underlying mechanism driving disability, the overlap between the chronically ill group and the disabled group is modest. In fact, only 14 percent of the chronically ill sample report having a disability (defined as health-related work limitations).[41] Moreover, even in the long run, chronic disease incidence does not necessarily translate into disability: The increase in the probability of disability is 12 percentage points. The disability literature generally recognizes disability based on self-reports of narrowly-defined health-related

[41] LaPlante (1989) cites a report showing that while approximately 50% of the adult population, ages between 18-64, has at least one chronic condition, only about 15 percent of the chronically ill have "one or more activity limitations".

work limitations, while chronic conditions, though self-reported, are objective health measures defined independently of individual's work capacity. Two people with the same chronic illness with equal severity may report differently regarding work-limiting disability, depending on their occupation, work environment, and ability (see Ozminkowski et al. (2000)). These results suggest that care is needed when comparing the average impact of all those who report a work limiting disability with the chronically ill rather than a specific subgroup (e.g: the Chronic-Severe group of Meyer and Mok (2008)).

5.4 Family Income

Table 7 reports the response of before and after-tax family income to a chronic health event. After-tax income of families with afflicted heads falls by 2 to 4% prior to the event, and drops by a statistically significant 9% at the time of onset relative to more than six years prior to the onset. Family income recovers to some extent during the next four years,[42] but in the long-run, family income falls by 12% relative to the expected level. The overall post-shock effect is an 8% drop in the expected family income relative to 7+ years before the onset.[43] As was the case with individual earnings, the pre-illness trend only exists for the high education group. Post-onset, both groups exhibit a similar pattern of partial recovery then income falls again.

Figure 4 visually compares the impact of a health event across outcomes for the entire sample. The impact on family income is closer to zero relative to individual earnings.[44] Further, even if individual earnings are the only income source, a progressive tax schedule may partially insure family income against a one-to-one propagation. However, the temporal pattern of before-tax family income closely tracks that of after-tax family income suggesting this insurance effect may be

[42]The estimates are too imprecisely measured to conclude that family income recovers completely during this period. The results by education emphasizes this imprecision.

[43]As in head's earnings, this is the coefficient estimate of the post-treatment dummy, which equals 1 when the year in question is after the event and 0, otherwise. Since the pre-onset event time dummies are still included, this would measure the average post-onset effect relative to 7+ years before the onset.

[44]The average share of head's earnings in before-tax family income is approximately two-thirds.

small.[45].

The long run divergence between the responses of earnings and family income in Figure 4 suggest other sources of income, such as increased spousal labor supply and public or private transfers, may have helped protect family income against health shocks.[46] Unfortunately, impacts of a health event on spousal labor supply outcomes are imprecisely estimated and are sensitive to changes in specification (not reported here). There are no changes in taxable income of other earners in the family. Private transfers appear to increase for those in the low education group but are not statistically different from zero.

Among public transfers, social security disability insurance (SSDI) increases markedly two years after onset (see table 8). The probability of becoming a SSDI recipient increases only by 2 percentage points in the long run. Taken together, these findings cannot account for the large divergence between individual earnings losses due to health shocks and household income.

5.5 Food Expenditures

Food expenditure responses are smaller than those of after-tax family income (table 9, figure 4). The expenditures begin falling some time before the onset, reach a trough of a 4% at the time of the onset then slowly recover over the next 5 years. Given that, in the year of onset, family income drops by 8%, the 4% fall is (though slightly smaller) consistent with the literature that finds income elasticities of food expenditures to be less than 1, between 0.6-0.7 (see Stephens (2001)). What is surprising is that health shocks seem to have only a transitory effect on food expenditures, despite its negative long run impact on family income.[47]

In contrast, the disability literature finds statistically significant long run reductions both in family income and food expenditures. For example, Meyer and Mok (2008) finds that long-run

[45] The differences in the estimates are not statistically significant.

[46] For the mean married or cohabiting households, the combined earnings of head and "spouse" account for 87% of before-tax family income.

[47] This is the opposite of the retirement food consumption puzzle, where food expenditures decrease despite retirement being a predictable event (see Hurst (2008)).

food consumption drops between 6 and 7% while income by 10-14%.[48] This is also consistent with the income elasticities of food. To find out whether the Supplemental Nutrition Assistance Program (SNAP), a federal food assistance program formerly known as Food Stamps, has enabled families to protect their food expenditures against negative health shocks in the long run, the analysis was done without including the food stamps amount. The temporal pattern of food expenditure responses to a chronic health event remains unchanged. SNAP cannot explain the recovery in food expenditures.

As was the case with earnings above, the impact of a health event differs across education group for other outcomes. Figures 5 & 6 display earnings, before and after tax family income and food expenditures separately for low and high education groups. Family income and earnings diverge for both groups, however the divergence occurs after the health event for those in the low education group and before the event for those in the high education group. Indeed, family income recovers almost completely before falling again for those in the low education group, contrasting sharply with those in the high education group. At the same time, the drop in food expenditures observed in the pooled sample is completely driven by those within the low education group. This suggests that even though the health event has a smaller negative impact on family income among low education individuals relative those with more education, the latter group is better able to keep food expenditures at a pre-event level, perhaps by dissaving. Unlike earnings and family income however, the reduction in food expenditures for those with low education begins before the health event. This may be a signal of remaining temporal differences across groups (not absorbed by individual linear trends), though a more comprehensive set of expenditure measures are necessary to examine this in more detail.

[48] Similar results are found in Stephens (2001).

6 Major Chronic Conditions

In the analysis above, twelve distinct chronic conditions are analyzed as one broad health measure, averaging across heterogeneity induced by specific chronic diseases. In this section, I focus on a subset of chronic conditions that are more likely to have a severe impact on health, given the sample size constraint. These conditions consists of cancer, heart attack, heart disease, lung disease and stroke; chronic conditions that may be severely detrimental to individual health and are relatively less likely to be subject to measurement error problems in the reporting of the onset date. Repeating the analysis using only major illness may address pre-event trends in outcomes. Those with chronic conditions other than major illnesses are dropped from the analysis.

Table 10 displays the coefficient estimates that quantify the effects of major illness on earnings, family income, food expenditures and the probability of being disabled (figure 7). Although the estimated impact on earnings is larger than in table 4, earnings growth is already down by 11% 3-4 years before the onset. Even when individual-specific time trends are included in the regression, the pre-illness trend remains.[49]

The clear pre-trend in earnings may suggest a deterioration in health condition prior to the onset of a severe chronic condition. Indeed, severe chronic conditions seem to be a better predictor of the probability of being disabled relative to all chronic conditions (column 4 table 10).Smith (2007) finds that prior chronic conditions are strong predictors of a future chronic condition.[50] If this is indeed the case, the estimated long-run effect of all chronic conditions could be driven by those whose minor chronic condition(s) develops into a more severe chronic condition(s) later on. When I drop individuals with a pre-existing minor health condition at the time of the onset, the pre-illness trend in earnings seems to disappear as well as the large post-onset effects (column 3 of table 10). The precision of the estimates is such that one cannot reject that these results are different from the column 1.

[49]The size of the pre-onset trend seems smaller but the estimates are much more imprecisely estimated than the ones without individual-specific time trends. The overall post-shock effect results are similar to each other.

[50]Specifically, three most common minor chronic conditions-arthritis, diabetes and hypertension-are good predictors of the future major illness onset.

As was the case for all chronic conditions, family income and food expenditure responses are muted relative to those of earnings (columns 5-6 table 10). For major illness, after-tax family income is nearly 16% below expected levels at the time of the event and remains so in the long run. The discrepancy in the response between earnings and family income across time suggests that families are able to partially buffer their income against earnings shock. Again, the response of food expenditures is smaller: A drop of 5% at the time of the shock, but unlike the benchmark case, it continues to drop in the long run. This suggests that for those with major illnesses, food expenditures are more difficult to insure against persistent income shocks.

7 Conclusion

This paper quantifies the economic impact of chronic conditions over time for the working-age male population.[51] Almost one half of the sample experiences at least one chronic disease.[52] The first onset occurs, on average, at age 44, during prime earning years. Thus, the impact on the labor market outcomes are expected to be non-negligible.

I find a substantial and robust negative effect of chronic illness on earnings and family income. Earnings drop by a statistically significant 12% at the time of the event and this effect increases to 18% in the long run. The impact on family income is roughly three quarters the size of the fall in head's earnings; a decline of 8% around the time of the event and 12% in the long-run. Food expenditure responses are smaller: A 4% drop at the time of the onset and a trend towards zero in the long run.

These results are qualitatively similar to the findings of the disability literature (see Stephens (2001) and Meyer and Mok (2008)), though are smaller in magnitude. That is not to say that the

[51] The list of chronic conditions considered in the paper are arthritis, asthma, cancer, diabetes, emotional/psychiatric disorders, heart attack, heart diseases, hypertension, learning disability, lung diseases, memory loss, stroke and others.

[52] The most common chronic conditions are hypertension (27.1%), arthritis (13.8%), diabetes (7.3%), heart diseases (4.8%), and cancer (4.6%). The fractions in parentheses are measured as a percentage of those with the corresponding chronic illness-the incidence of which must have occurred less than 10 years ago-in a sample of male heads ages between 25-65, present at least in three consecutive surveys between 1968-2009.

temporal effects of chronic conditions are merely representing those of disability. First, chronic diseases affect a much wider segment of the population than disability. Although the onset of chronic conditions increases the probability of becoming disabled, only a modest fraction of those with a chronic illness are disabled. For example, in 2009, only 14% of those with a chronic condition had a disability. The effects of chronic conditions are distinct from those of disability and therefore, it is important to understand the impact of chronic conditions separately.

There are a few caveats. First, the identification of the effects rest upon the exogenity of the first onset conditional on observables and fixed unobservables. Second, the main results of this paper measure the average impact of chronic disease onset. However, the observed differential response by education group, and condition emphasizes the important of heterogenous effects of chronic illness across population distribution. Further research is needed. This paper can be viewed as a first step in understanding the economic consequences of chronic diseases with the current data limitations. However, to formulate specific policies, it is important that each condition is studied separately as soon as enough data is available.

References

Adams, P., Hurd, M., McFadden, D., Merrill, A. and Ribeiro, T. (2003), 'Healthy, wealthy, and wise? tests for direct causal paths between health and socio- economic status', *Journal of Econometrics* **112**, 3–56.

Autor, D. H. and Duggan, M. G. (2003), 'The rise in the disability rolls and the decline in unemployment', *The Quarterly Journal of Economics* **118**(1), 157–205.

Autor, D. H., Katz, L. F. and Kearney, M. S. (2008), 'Trends in u.s. wage inequality: Revising the revisionists', *The Review of Economics and Statistics* **90**(2), 300–323.

Baker, M., Stabile, M. and Deri, C. (2004), 'What do self-reported, objective, measures of health measure?', *Journal of Human Resources* **39**(4).

Blundell, R., Griffith, R. and Windmeijer, F. (2002), 'Individual effects and dynamics in count data models', *Journal of Econometrics* **108**(1), 113–131.

Card, D. and Lemieux, T. (2001), ‘Can falling supply explain the rising return to college for younger men? a cohort-based analysis’, *The Quarterly Journal of Economics* **116**(2), 705–746.

Centers for Disease Control and Prevention (2009), Chronic diseases : The power to prevent, the call to control, Technical report, National Center for Chronic Disease Prevention and Health Promotion, CDC.

Charles, K. K. (2003), ‘The longitudinal structure of earnings losses among work-limited disabled workers’, *Journal of Human Resources* **38**(3).

Chatterji, P., Joo, H. and Lahiri, K. (2010), Beware of unawareness: Racial/ethnic disparities in awareness of chronic diseases, NBER Working Papers 16578, National Bureau of Economic Research, Inc.

Crichton, S., Stillman, S. and Hyslop, D. (2005), Returning to work from injury: Longitudinal evidence on employment and earnings, IZA Discussion Papers 1857, Institute for the Study of Labor (IZA).

Cutler, D. M. and Lleras-Muney, A. (2006), Education and health: Evaluating theories and evidence, NBER Working Papers 12352, National Bureau of Economic Research, Inc.

Cutler, D. M., Lleras-Muney, A. and Vogl, T. (2008), Socioeconomic status and health: Dimensions and mechanisms, NBER Working Papers 14333, National Bureau of Economic Research, Inc.

Dano, A. M. (2005), ‘Road injuries and long-run effects on income and employment’, *Health Economics* **14**(9), 955–970.

Datta Gupta, N., Kleinjans, K. J. and Larsen, M. (2011), The effect of an acute health shock on work behavior: Evidence from different health care regimes, IZA Discussion Papers 5843, Institute for the Study of Labor (IZA).

DeVol, R., Bedroussain, A., Charuwon, A., Chatterjee, A., Kim, I. K., Kim, S. and Klowden, K. (2007), An unhealthy america: The economic burden of chronic disease – charting a new course to save lives and increase productivity and economic growth, Technical report, Milken Institute.

Gallo, W., Teng, H., Falba, T., Kasl, S., Krumholz, H. and Bradley, E. (2006), ‘The impact of late career job loss on myocardial infarction and stroke: a 10 year follow up using the health and retirement survey’, *Occupational and Environmental Medicine* **63**(10), 683–7.

Goldman, D. P. and Smith, J. P. (2004), Can patient self-management help explain the ses health gradient?, HEW 0403004, EconWPA.

Hurst, E. (2008), The retirement of a consumption puzzle, NBER Working Papers 13789, National Bureau of Economic Research, Inc.

Islam, A. and Maitra, P. (2008), Health shocks and consumption smoothing in rural households: Does microcredit have a role to play?, Monash Economics Working Papers 22/08, Monash University, Department of Economics.

Jacobson, L. S., LaLonde, R. J. and Sullivan, D. (1993), 'Earnings losses of displaced workers', *American Economic Review* **83**(4), 4–29.

LaPlante, M. (1989), Disability risks of chronic illnesses and impairments, Technical report, U.S Department of Health and Human Services.

Lemieux, T. (2006), 'Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?', *American Economic Review* **96**(3), 461–498.

Lundborg, P., Nilsson, M. and Vikstrom, J. (2011), Socioeconomic heterogeneity in the effect of health shocks on earnings: evidence from population-wide data on swedish workers, Working paper series, IFAU - Institute for Labour Market Policy Evaluation.

Manning, W. G. and Mullahy, J. (2001), 'Estimating log models: to transform or not to transform?', *Journal of Health Economics* **20**(4), 461–494.

Marmot, M. G. (1999), *Multilevel approaches to understanding social determinants*, Vol. Social Epidemiology, Oxford: Oxford University Press.

Meyer, B. D. and Mok, W. K. C. (2008), 'Disability, earnings, income and consumption', *Harris School of Public Policy Studies, The University of Chicago* (Working Paper 0610).

Mok, W. K. C., Meyer, B. D., Charles, K. K. and Achen, A. C. (2008), 'A note on "the longitudinal structure of earnings losses among work-limited disabled workers"', *Journal of Human Resources* **43**(3), 721–728.

Mullahy, J. (1998), 'Much ado about two: reconsidering retransformation and the two-part model in health econometrics', *Journal of Health Economics* **17**(3), 247–281.

Piketty, T. and Saez, E. (2003), 'Income inequality in the united states, 1913-98', *Quarterly Journal of Economics* **118**(1), 1–39.

Richards, M., Hardy, R. and Wadsworth, M. (1997), 'The effects of divorce and separation on mental health in a national uk birth cohort', *Psychological Medicine* **27**(5), 1121–1128.

Riphahn, R. T. (1998), 'Income and employment effects of health shocks - a test case for the german welfare state', *Journal of Population Economics* **12**(3), 363–389.

Silva, J. M. C. S. and Tenreyro, S. (2006), 'The log of gravity', *The Review of Economics and Statistics* **88**(4), 641–658.

Smith, J. P. (1999), 'Healthy bodies and thick wallets: Dual relation between health and economic status', *Journal of Economic Perspectives* **13**(2), 145–166.

Smith, J. P. (2003), Consequences and predictors of new health events, IFS Working Papers W03/22, Institute for Fiscal Studies.

Smith, J. P. (2007), 'The impact of socioeconomic status on health over the life-course', *Journal of Human Resources* **42**(4).

Stephens, M. (2001), 'The long-run consumption effects of earnings shocks', *The Review of Economics and Statistics* **83**(1), 28–36.

Stevens, A. H. (1997), 'Persistent effects of job displacement: The importance of multiple job losses', *Journal of Labor Economics* **15**(1), 165–188.

Wu, S. (2003), 'The effects of health events on the economic status of married couples', *Journal of Human Resources* **38**(1).

Figure 1 Number Healthy vs. With a Chronic Illness by Age Group

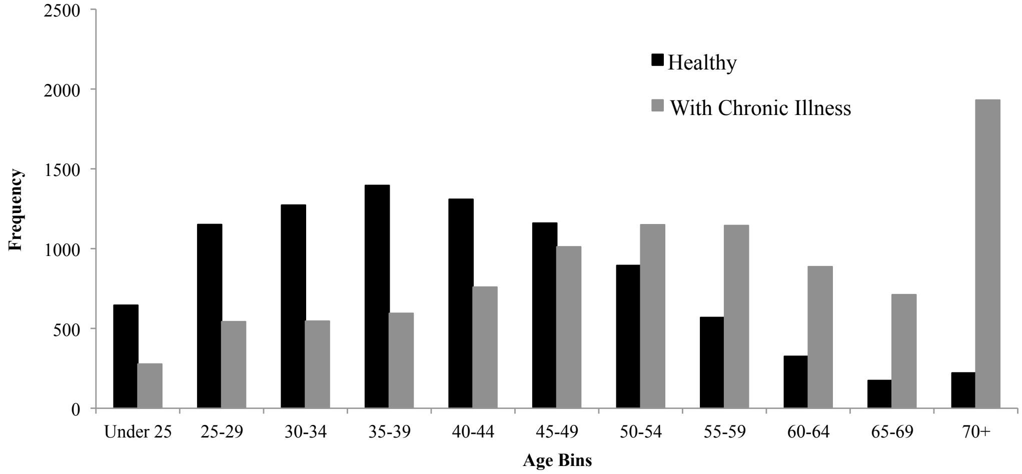
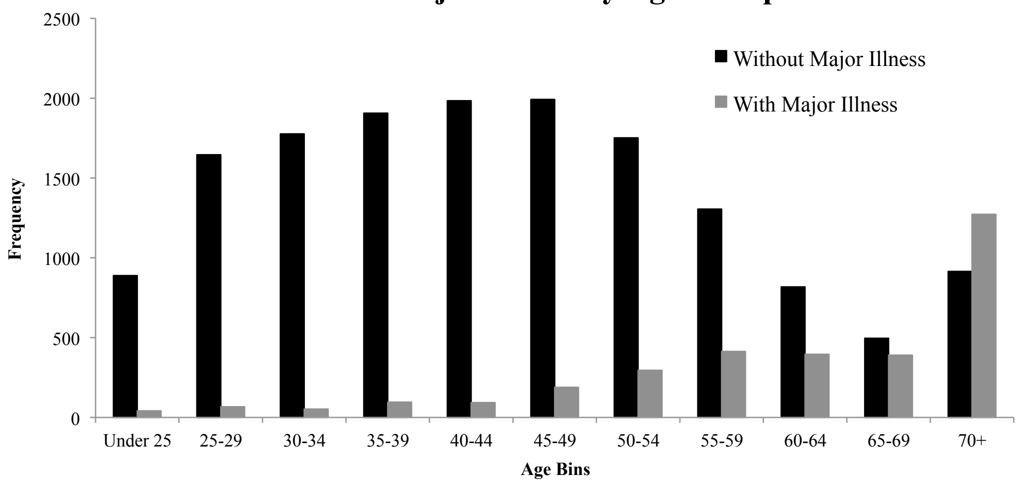


Figure 2 Number Without vs. With a Major Illness by Age Group



The sample consists of male heads between 1999 - 2009. Major illness consists of cancer, heart attack, heart disease, lung disease and stroke. The PSID-provided cross-sectional individual weights are used to calculate the frequencies.

Figure 3 Effect of Chronic Health Event on Earnings by Education Group

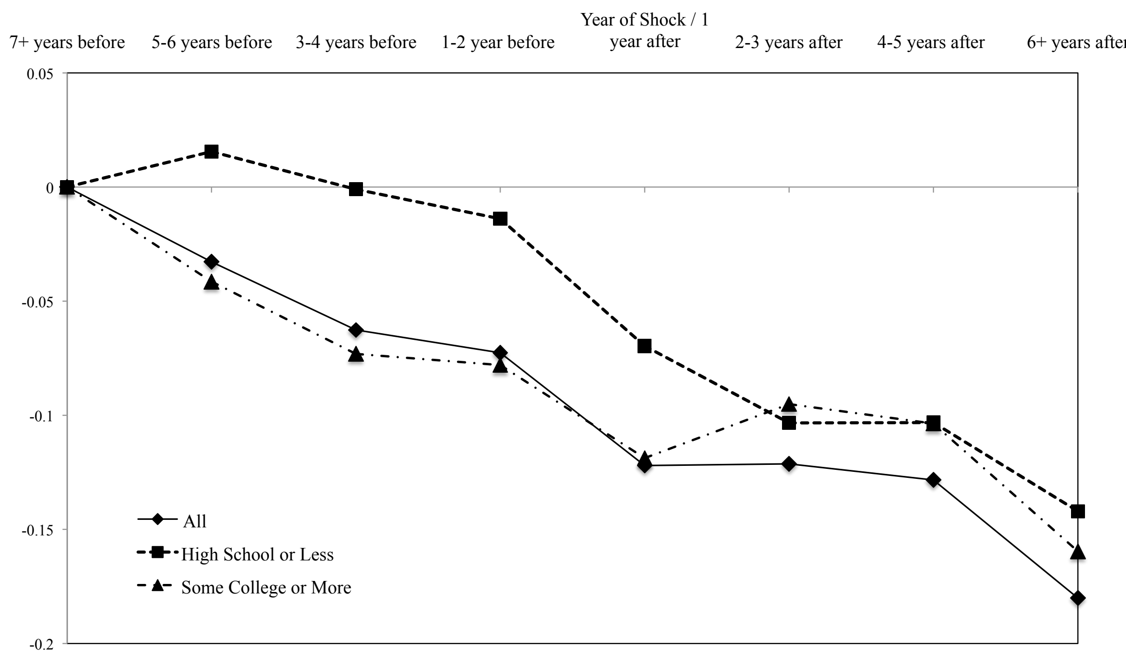


Figure 4 Effect of Chronic Health Event on Earnings, Family Income and Food Expenditures

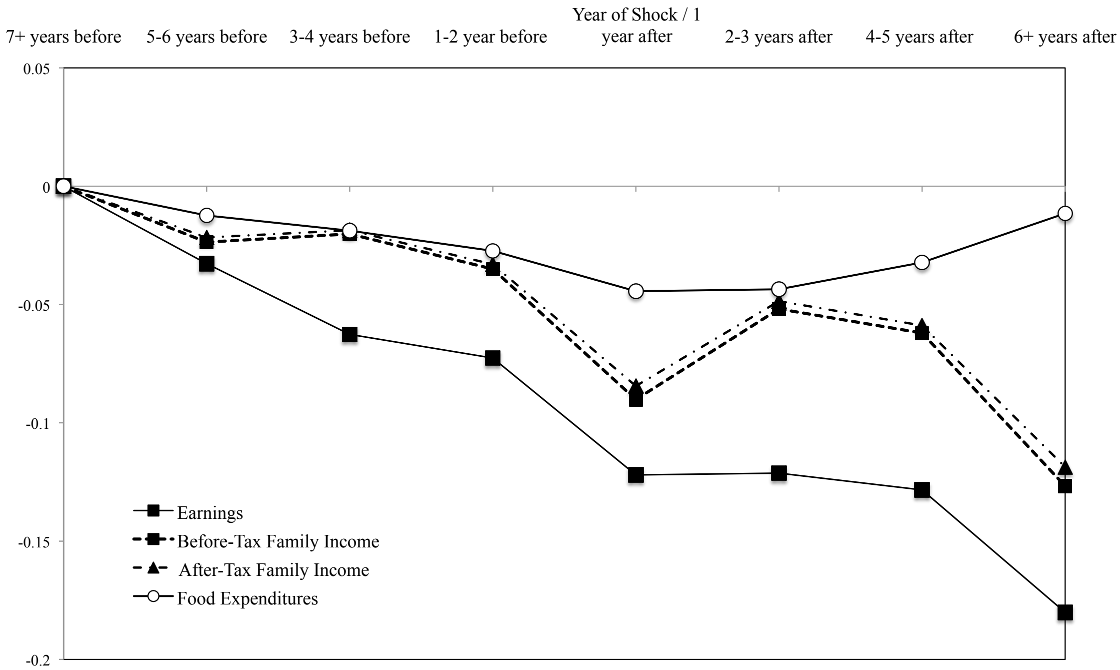


Figure 5 Effect of Chronic Health Event on Earnings, Family Income and Food Expenditures, Low Education Group Only

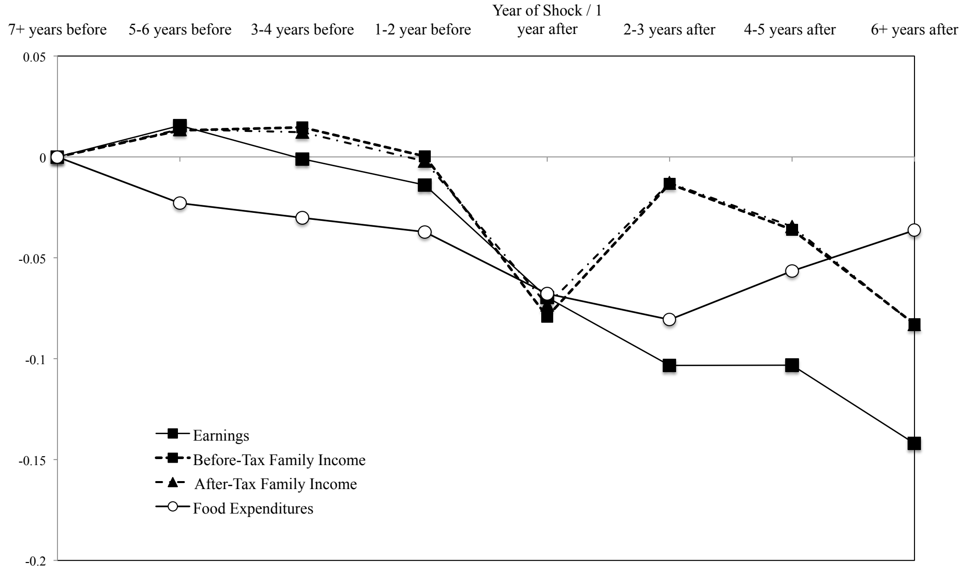


Figure 6 Effect of Chronic Health Event on Earnings, Family Income and Food Expenditures, High Education Group Only

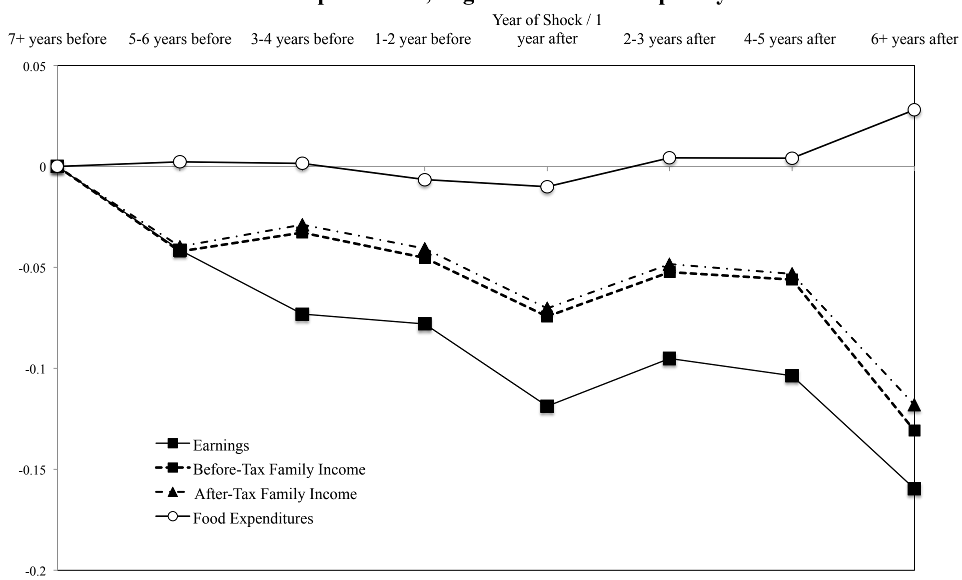


Figure 7 Effect of Major Illness on Earnings, Family Income and Food Expenditures

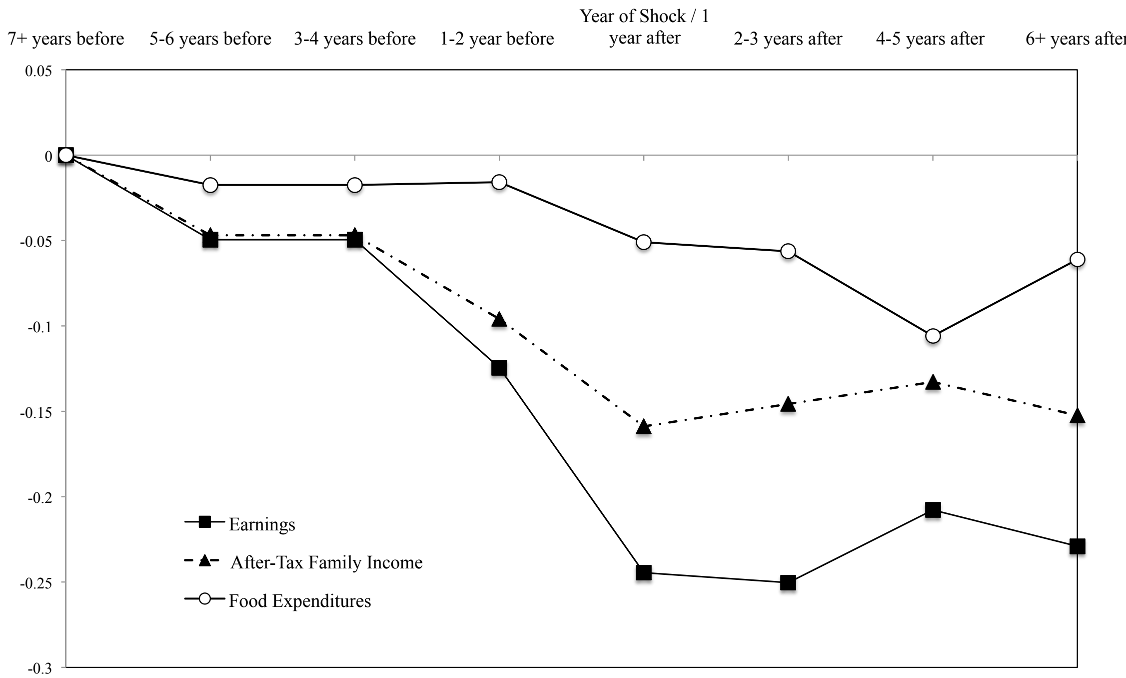


Table 1 Prevalence Rate & Average Age of Onset by Chronic Condition

	Share (%)			Average Onset Age		
	All	*HS or Less*	*Some College or More*	All	*HS or Less*	*Some College or More*
Arthritis	12.4%	14.6%	10.8%	43.5	42.3	44.6
Asthma	8.9%	8.2%	9.4%	19.4	17.6	20.5
Cancer	3.7%	2.5%	4.7%	44.9	42.8	45.7
Diabetes	8.6%	10.1%	7.4%	43.7	43.2	44.2
Emotional/Psychiatric Disorder	6.3%	7.8%	5.2%	31.3	30.6	32.1
Heart Attack	3.6%	4.0%	3.2%	46.7	43.1	50.1
Heart Diseases	4.3%	4.2%	4.3%	48.6	47.0	49.6
Hypertension	26.7%	28.6%	25.3%	43.2	42.1	44.0
Learning Disability	3.3%	4.9%	2.2%	13.7	12.1	16.0
Lung Diseases	3.2%	4.5%	2.3%	36.2	39.0	32.4
Memory Loss	1.1%	2.0%	0.5%	35.2	34.9	36.2
Stroke	1.5%	2.0%	1.1%	45.3	45.8	44.7
Other	15.3%	16.1%	14.7%	37.4	37.3	37.4
Any Chronic	53.7%	56.9%	51.5%			
Age Now		45.1			48.1	
Sample Size	2,709	1,287	1,422	1,396	713	683

The sample consists of male heads between ages 25-65 in year 2009. *HS or Less* are those with less than or equal to 12 years of education, while *Some College or More* are those with more than 12 years of education. Any chronic is the share of those who have at least one or more chronic condition. All means are weighted using the PSID provided individual cross-sectional weights for year 2009.

Table 2 Summary Statistics

	Never Chronic	Ever Chronic
Age	41.01	41.23
	(10.88)	(10.12)
Years of education	13.50	13.51
	(2.49)	(2.28)
Percentage white	0.85	0.85
	(0.36)	(0.35)
Percentage married	0.78	0.79
	(0.42)	(0.41)
Number of children	1.18	1.17
	(1.28)	(1.21)
Family size	3.04	3.03
	(1.50)	(1.43)
Years present in sample	16.94	20.94
	(8.02)	(7.50)
Percent of survey years married	0.80	0.81
	(0.32)	(0.30)
Annual earnings ($)	63,578	57,526
	(59028)	(45194)
Family income ($)	80,424	75,424
	(58540)	(49394)
Food Expenditure ($)	9,246	9,201
	(4771)	(4465)
Percentage working now	0.92	0.91
	(0.27)	(0.28)
Percentage self-employed	0.16	0.13
	(0.37)	(0.34)
Percentage salaried worker	0.45	0.44
	(0.50)	(0.50)
Percentage hourly worker	0.34	0.36
	(0.47)	(0.48)
Self-reported health status	1.79	2.15
	(0.80)	(0.92)
Wife's age if present	39.84	39.67
	(10.70)	(10.16)
Wife's education if present	13.27	13.18
	(2.46)	(2.11)
Wife's earnings if present	22,592	23,430
	(29554)	(30767)
Percentage working now (Wife)	0.66	0.70
	(0.47)	(0.46)
Fraction of Individuals	0.501	0.499
Number of Observations	16,776	25,096

The sample is restricted to male heads between ages 25-65, present at least in three consecutive surveys between 1968-2009. All means are computed based on all available data from 1968-2009, after necessary sample restrictions have been imposed (see Section 3 of the text for details). Data on percentage salaried and hourly workers are available 1977 onwards. Self-reported health status ranges from 1 to 5 in the order of decreasing health status with 1 "excellent" to 5 "poor", and is available from 1984. The PSID longitudinal family weights are used to calculate the means. All monetary values are converted to 2009 $ using the CPI-U.

Table 3 Effects of Chronic Health Event on Earnings

	(1)	(2)	(3)	(4)
	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]
5-6 years before	-0.0328*	-0.0314*	-0.0340**	-0.0360**
	[0.0171]	[0.0169]	[0.0167]	[0.0165]
3-4 years before	-0.0626***	-0.0611***	-0.0623***	-0.0638***
	[0.0185]	[0.0183]	[0.0182]	[0.0181]
1-2 year before	-0.0726***	-0.0709***	-0.0703***	-0.0831***
	[0.0224]	[0.0223]	[0.0224]	[0.0233]
Year of Shock / 1 year after	-0.1220***	-0.1195***	-0.1219***	-0.1345***
	[0.0238]	[0.0237]	[0.0238]	[0.0246]
2-3 years after	-0.1213***	-0.1193***	-0.1185***	-0.1318***
	[0.0278]	[0.0277]	[0.0280]	[0.0289]
4-5 years after	-0.1284***	-0.1264***	-0.1315***	-0.1384***
	[0.0340]	[0.0338]	[0.0339]	[0.0357]
6+ years after	-0.1801***	-0.1771***	-0.1814***	-0.1902***
	[0.0408]	[0.0406]	[0.0410]	[0.0447]
Overall postshock effect	-0.1321***	-0.1298***	-0.1314***	-0.1416***
	[0.0257]	[0.0256]	[0.0258]	[0.0267]
Sample Size	42,211	42,211	40,925	38,773

(1) A multiplicate fixed-effects model is estimated using the Poisson fixed effects estimator. The controls include year fixed effects, age dummies and marital status.
(2) Extra controls for changes in marital status and number of children are added to (1).
(3) Same as (1) but restricted to male heads ages between 25-60.
(4) Same as (1) but restricted to male heads ages between 25-55.
Dependent variable is head's earnings. All regressions are weighted using the PSID-provided longitudinal weights.
The overall post-shock effect is the average post-event effect relative to more than six years before the onset.
Standard errors are cluster-robust clustering on the individuals.

Table 4 Effects of Chronic Health Event on Earnings by Education Group

	All	*High School or Less*	*Some College or More*
	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]
5-6 years before	-0.0328*	0.0155	-0.0414*
	[0.0171]	[0.0224]	[0.0219]
3-4 years before	-0.0626***	-0.0010	-0.0731***
	[0.0185]	[0.0254]	[0.0233]
1-2 year before	-0.0726***	-0.0139	-0.0780***
	[0.0224]	[0.0313]	[0.0283]
Year of Shock / 1 year after	-0.1220***	-0.0696**	-0.1188***
	[0.0238]	[0.0346]	[0.0295]
2-3 years after	-0.1213***	-0.1034**	-0.0951***
	[0.0278]	[0.0402]	[0.0340]
4-5 years after	-0.1284***	-0.1032**	-0.1037**
	[0.0340]	[0.0488]	[0.0420]
6+ years after	-0.1801***	-0.1420**	-0.1596***
	[0.0408]	[0.0552]	[0.0497]
Overall postshock effect	-0.1321***	-0.0936***	-0.1166***
	[0.0257]	[0.0358]	[0.0315]
Sample Size	42,211	19,565	22,646

The first column are the results from benchmark regression, column (1) of Table 3. *High school or less* are defined as those with equal or less than 12 years of schooling. *Some college or more* are those with more than 12 years of schooling. All regressions include individual fixed-effects, year effects, age dummies and marital status. All regressions are weighted using the PSID-provided longitudinal weights. The overall post-shock effect is the average post-event effect relative to more than six years before the onset. Standard errors are clustered on individual.

Table 5 Effects of Chronic Health Event on Labor Maket Outcomes by Education Group

	Employment			Annual Hours of Work		
	All	*High School or Less*	*Some College or More*	All	*High School or Less*	*Some College or More*
	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]
5-6 years before	0.0071	0.0137	0.0019	4.26	22.23	-8.43
	[0.0060]	[0.0093]	[0.0080]	[23.23]	[38.01]	[28.71]
3-4 years before	0.0000	-0.0046	0.0034	-33.53	-23.84	-37.94
	[0.0079]	[0.0130]	[0.0100]	[25.88]	[38.63]	[35.05]
1-2 year before	0.0002	-0.0082	0.0071	-30.57	-42.40	-20.68
	[0.0092]	[0.0149]	[0.0117]	[30.92]	[49.49]	[39.25]
Year of / 1 year after	-0.0105	-0.0214	-0.0015	-93.42***	-117.24**	-70.51
	[0.0112]	[0.0190]	[0.0133]	[35.06]	[56.83]	[44.06]
2-3 years after	-0.0283**	-0.0333	-0.0220	-122.10***	-174.69***	-69.47
	[0.0144]	[0.0251]	[0.0168]	[39.56]	[64.86]	[48.84]
4-5 years after	-0.0372**	-0.0546*	-0.0216	-100.73**	-130.00*	-72.27
	[0.0159]	[0.0282]	[0.0181]	[44.43]	[73.44]	[54.75]
6+ years after	-0.0555***	-0.0925***	-0.0274	-175.59***	-242.81***	-119.71*
	[0.0186]	[0.0333]	[0.0208]	[51.36]	[81.30]	[65.59]
Sample Size	41,868	19,403	22,465	41,521	19,233	22,288

The impact on employment is estimated using a fixed effects linear probability model. The effect on annual hours of work are estimated using linear fixed effects models. High school or less includes those with equal or less than 12 years of schooling. Some college or more are those with more than 12 years of schooling. All regressions include individual fixed-effects, year effects, age dummies and marital status. All regressions are weighted using the PSID-provided longitudinal weights. The overall post-shock effect is the average post-event effect relative to more than six years before the onset. Standard errors clustered on individuals.

Table 6 Effects of Chronic Health Event on the Probability of Becoming Disabled

	All	*High School or Less*	*Some College or More*
	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]
5-6 years before	0.0042	0.0007	0.0073
	[0.0072]	[0.0118]	[0.0091]
3-4 years before	0.0043	-0.0009	0.0084
	[0.0075]	[0.0114]	[0.0102]
1-2 year before	0.0324***	0.0425***	0.0239**
	[0.0092]	[0.0156]	[0.0112]
Year of Shock / 1 year after	0.0813***	0.0910***	0.0735***
	[0.0129]	[0.0210]	[0.0163]
2-3 years after	0.0889***	0.1114***	0.0705***
	[0.0139]	[0.0229]	[0.0173]
4-5 years after	0.0971***	0.1281***	0.0732***
	[0.0154]	[0.0267]	[0.0177]
6+ years after	0.1245***	0.1533***	0.1026***
	[0.0158]	[0.0286]	[0.0184]
Sample Size	41,836	19,391	22,445

Fixed-Effects Linear Probability Models are used with the PSID-provided longitudinal weights. *High school or less* includes those with equal or less than 12 years of schooling. *Some college or more* are those with more than 12 years of schooling. Controls include calendar year effects, age indicators and marital status. Standard errors are clustered on individuals. The dependent variable is equal to 1 if the respondent answers "yes" to the following question: *Do you (HEAD) have any physical or nervous condition that limits the type of work or amount of work you can do?*

Table 7   Effects of Chronic Health Event on Before- and After-Tax Family Income

	Before-Tax Family Income			After-Tax Family Income		
	All	*High School or Less*	*Some College or More*	All	*High School or Less*	*Some College or More*
	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]
5-6 years before	-0.0236	0.0131	-0.0421*	-0.0216	0.0136	-0.0397*
	[0.0196]	[0.0298]	[0.0248]	[0.0184]	[0.0286]	[0.0228]
3-4 years before	-0.0202	0.0146	-0.0328	-0.0186	0.0122	-0.0288
	[0.0196]	[0.0283]	[0.0259]	[0.0182]	[0.0269]	[0.0236]
1-2 year before	-0.0350	0.0003	-0.0453	-0.0330	-0.0021	-0.0407
	[0.0230]	[0.0319]	[0.0315]	[0.0213]	[0.0300]	[0.0288]
Year of Shock / 1 year after	-0.0901***	-0.0792**	-0.0743**	-0.0844***	-0.0733**	-0.0703**
	[0.0265]	[0.0381]	[0.0346]	[0.0244]	[0.0354]	[0.0315]
2-3 years after	-0.0519*	-0.0134	-0.0523	-0.0486*	-0.0129	-0.0484
	[0.0280]	[0.0382]	[0.0384]	[0.0257]	[0.0354]	[0.0349]
4-5 years after	-0.0621*	-0.0360	-0.0560	-0.0589**	-0.0345	-0.0532
	[0.0324]	[0.0459]	[0.0433]	[0.0298]	[0.0431]	[0.0392]
6+ years after	-0.1267***	-0.0829	-0.1308**	-0.1187***	-0.0832*	-0.1181**
	[0.0393]	[0.0521]	[0.0540]	[0.0364]	[0.0493]	[0.0495]
Overall postshock effect	-0.0792***	-0.0524	-0.0741**	-0.0743***	-0.0494	-0.0689**
	[0.0259]	[0.0345]	[0.0355]	[0.0238]	[0.0321]	[0.0323]
Sample Size	41,868	19,403	22,465	41,868	19,403	22,465

Dependent variables are in natural logs. Taxes are calculated the NBER TAXSIM calculator. *High school or less* includes those with equal or less than 12 years of schooling. *Some college or more* are those with more than 12 years of schooling. All regressions include individual fixed-effects, year effects, age dummies and marital status. All regressions are weighted using the PSID-provided longitudinal weights. The overall post-shock effect is the average post-event effect relative to more than six years before the onset. Standard errors are clustered on individuals.

Table 8 Effects of Chronic Health Event on SSDI

	SSDI	*Probability of Becoming a SSDI recipient*
	Coeff. [Se]	Coeff. [Se]
5-6 years before	-2.68	-0.0011
	[31.55]	[0.0020]
3-4 years before	-25.92	-0.0018
	[31.28]	[0.0021]
1-2 year before	-68.14	-0.0056*
	[43.68]	[0.0029]
Year of Shock / 1 year after	-58.27	-0.0032
	[49.67]	[0.0036]
2-3 years after	206.26**	0.0132**
	[100.75]	[0.0058]
4-5 years after	191.09*	0.0138**
	[113.34]	[0.0069]
6+ years after	301.86***	0.0203***
	[93.58]	[0.0053]
Sample Size	30,790	30,790

Social Security Disablity Insurance (SSDI) has been measured in the PSID since 1986. The amount of the received benefits are mesasured in 2009 real dollars and is estimated using a linear fixed-effects model; while the *Probability of Becoming a SSDI Recipient* is estimated using a fixed-effects linear probability model. All regressions include individual fixed-effects, year effects, age dummies and marital status. All regressions are weighted using the PSID-provided longitudinal weights. Standard errors are clustered on individuals.

Table 9 Effects of Chronic Health Event on Food Expenditures

	All	*High School or Less*	*Some College or More*
	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]
5-6 years before	-0.0124	-0.0229	0.0023
	[0.0146]	[0.0220]	[0.0188]
3-4 years before	-0.0188	-0.0302	0.0015
	[0.0157]	[0.0224]	[0.0213]
1-2 year before	-0.0273	-0.0372	-0.0065
	[0.0172]	[0.0263]	[0.0221]
Year of Shock / 1 year after	-0.0444**	-0.0678**	-0.0101
	[0.0196]	[0.0302]	[0.0249]
2-3 years after	-0.0436*	-0.0807**	0.0043
	[0.0228]	[0.0347]	[0.0292]
4-5 years after	-0.0322	-0.0566	0.0041
	[0.0251]	[0.0378]	[0.0325]
6+ years after	-0.0115	-0.0364	0.0280
	[0.0287]	[0.0440]	[0.0362]
Overall postshock effect	-0.0377*	-0.0668**	0.0024
	[0.0197]	[0.0294]	[0.0256]
Sample Size	41,868	19,403	22,465

Linear fixed-effects models are estimated. *High school or less* includes those with equal or less than 12 years of schooling. *Some college or more* are those with more than 12 years of schooling. All regressions include individual fixed-effects, year effects, age dummies and marital status. All regressions are weighted using the PSID-provided longitudinal weights. The overall post-shock effect is the average post-event effect relative to more than six years before the onset. Standard errors are clustered on the individual.

Table 10 Effects of Major Illness on Earnings, Family Income, Food Expenditures and the Probability of Becoming Disabled

	(1)	(2)	(3)	(4)	(5)	(6)
	Earnings	Earnings	Earnings	Disability	Family Income	Food Expenditures
	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]	Coeff. [Se]
5-6 years before	-0.0495*	-0.0610	0.0147	0.0297*	-0.0469	-0.0175
	[0.0282]	[0.0635]	[0.0537]	[0.0165]	[0.0290]	[0.0250]
3-4 years before	-0.1107***	-0.0787	0.0015	0.0314*	-0.0469	-0.0175
	[0.0305]	[0.0830]	[0.0546]	[0.0180]	[0.0305]	[0.0233]
1-2 year before	-0.1244***	-0.0281	-0.0351	0.0603***	-0.0959***	-0.0158
	[0.0322]	[0.0934]	[0.0572]	[0.0192]	[0.0353]	[0.0255]
Year of Shock / 1 year after	-0.2445***	-0.1961*	-0.0689	0.1313***	-0.1588***	-0.0510*
	[0.0382]	[0.1164]	[0.0639]	[0.0238]	[0.0344]	[0.0291]
2-3 years after	-0.2503***	-0.3151**	-0.0523	0.1784***	-0.1457***	-0.0564*
	[0.0434]	[0.1425]	[0.1128]	[0.0278]	[0.0378]	[0.0324]
4-5 years after	-0.2078***	-0.2486	-0.1665*	0.1845***	-0.1328***	-0.1059***
	[0.0503]	[0.1653]	[0.0900]	[0.0299]	[0.0465]	[0.0408]
6+ years after	-0.2290***	-0.3541*	-0.0759	0.2173***	-0.1524***	-0.0611
	[0.0700]	[0.1971]	[0.1135]	[0.0296]	[0.0463]	[0.0406]
Individual-specific linear time trends	No	Yes	No	No	No	No
Exclude those with pre-existing minor conditions	No	No	Yes	No	No	No
Overall postshock effect	-0.2368***	-0.2084*	-0.0816	0.1685***	-0.1499***	-0.0631**
	[0.0397]	[0.1171]	[0.0697]	[0.0219]	[0.0322]	[0.0271]
Sample Size	31,432	31,227	16,363	31,204	31,227	31,227

A multiplicative fixed-effects model using the Poisson fixed-effects estimator is estimated for column (1); while linear fixed-effects models are estimated for columns (2), (3), (5) and (6). A fixed-effects linear probability model is estimated for column (4). All regressions include individual fixed-effects, year effects, age dummies and marital status. All regressions are weighted using the PSID-provided longitudinal weights. The overall post-shock effect is the average post-event effect relative to more than six years before the onset. Standard errors are clustered on individuals.

Table 11 Chronic Disease Prevalence by Education Group

	Share (%)		
	All	*High School or Less*	*Some College or More*
Arthritis	13.8%	16.0%	12.1%
Asthma	2.8%	2.6%	2.9%
Cancer	4.6%	3.8%	5.1%
Diabetes	7.3%	7.8%	7.0%
Emotional/Psychiatric Disorders	4.3%	4.1%	4.5%
Heart Attack	3.7%	4.5%	3.2%
Heart Diseases	4.8%	4.5%	5.0%
Hypertension	27.1%	29.0%	25.6%
Learning Disability	0.5%	0.4%	0.6%
Lung Diseases	2.5%	2.8%	2.2%
Memory Loss	0.7%	1.2%	0.3%
Stroke	1.4%	1.2%	1.6%
Other	15.6%	14.5%	16.4%
Healthy	50.0%	48.3%	51.6%
Number of Individuals	3,199	1,514	1,685

The sample consists of male heads between ages 25-65, present at least in three consecutive surveys between 1968 - 2009. Those who experienced their first chronic condition more than 9 years prior to their first health report are dropped. *High school or Less* are those with less than or equal to 12 years of education, while *Some College or More* are those with more than 12 years of education. Any chronic is the share of those who have at least one or more chronic condition. All means are weighted using the PSID-provided longitudinal weights.

A Data Appendix

A.1 Background on PSID

The PSID is a comprehensive longitudinal survey of a nationally representative sample of the U.S. population that contains information on a variety of socio-economic topics. The survey began in 1968, with a nationally representative cross-section of about 3,000 families and a sample of about 2,000 low-income families (commonly known as the Survey Research Center (SRC) sample and the Survey of Economic Opportunity (SEO) sample, respectively). It followed individuals from families in the core sample as well as their descendants on an annual basis until 1996. In 1997, the PSID became a biennial survey and its sample underwent a major change: the SEO subsample was reduced by two-thirds, though 609 families were later reinstated, and 441 immigrant families were newly added to keep up with the changing U.S population structure.[53]

A.2 Measurement of Chronic illness

a Construction of the Onset Date

This section defines a chronic health event and describes how an onset date is determined. Since 1999, the PSID has been asking questions regarding a set of chronic conditions: Cancer, heart attack, heart disease, lung disease, stroke, arthritis, diabetes, hypertension, psychological problems, asthma, memory loss, learning disability and since 2005, other chronic conditions. The survey also asks about the date of first diagnosis. For example, in the case of heart attack, the PSID asks the respondent, *"Has a doctor ever told you that you have or had a heart attack?"* If the response is positive, then the follow-up question about the date of first diagnosis is asked. During the 1999-2003 waves, the PSID asked *"How long have you had this condition?"* and measured this in days, weeks, months and years. In the 2005 wave, the question was changed to *"How old were you the first time you had a heart attack?"*

[53] In 1968, the sample size was 4,082 families, it grew to 10,764 families in 1994, reduced to 6,747 in 1997 and finally as of 2009, comprised 8,690 families.

An individual is considered to have experienced a chronic health event if he was ever diagnosed with any one of the thirteen chronic conditions mentioned above, including other chronic conditions from 2005 onwards. The onset of chronic illness is defined as the year of the earliest diagnosis. It may be that individuals are afflicted with multiple chronic illnesses over time. Then, the earliest diagnosis across all thirteen conditions is deemed to be the onset date. Specifically, for each chronic condition in each wave during the period 1999-2003, the year of diagnosis is computed by first constructing the number of days the respondent had the condition, and subtracting it from the interview date. For the period 2005-2009, the difference between the current age and the age of first diagnosis, which gives us roughly the number years since diagnosis, is subtracted from the current calendar year for each condition in each wave. This procedure gives us the year of diagnosis for each chronic condition in each wave.

Since the PSID asks the same question in each wave, individuals report their onset date multiple times, as long as they have stayed in the sample for more than a wave during the period 1999-2009. In order to create a single earliest onset date across chronic conditions for each afflicted individual, I take the following approach. For each chronic disease in the 1999-2003 surveys, the reported date is regarded as the true onset if the reported number of days since the onset is less than 365 days. For 2005-2009 surveys, if the age at interview is the same as the reported onset age, then the corresponding calendar year is considered as the true onset date. If an individual still has more than one onset date after this step for each chronic condition, the earliest one is taken as the true onset date. For others whose date of first diagnosis falls out of the above one-year range, the mode of the reported dates is deemed to be the true onset date, and if there is no mode, then the earliest reported date. Finally, the earliest reported onset date across the thirteen conditions is obtained as the date of the event. Another method of computing the onset date is to simply take the mode of the repeatedly reported dates of first diagnosis for each chronic condition, then take the earliest date across the conditions.[54]

[54]The results using this method is not reported in the paper, but are available upon request.

b *Implications*

The questions described above have been asked since 1999. But, for those who report a chronic condition between 1999 and 2009, the PSID contains a rich set of historical labor supply information (potentially stretching back decades). This study should be understood as investigating the economic lives of those who participated in the 1999-2009 waves of the PSID with their past economic and chronic illness history available.

In addition to dropping those who attrit out of the survey before 1999, those whose health information are missing in any year between 1999-2009 are dropped as well. To reduce recall bias, an individual whose earliest diagnosis dates back ten or more years from their first health report is excluded from the analysis. For those whose onset dates are in earlier years, say the early 1990s, these restrictions would require them to be survivors. For example, if an individual gets his first chronic disease in 1990, he must survive at least the next nine years (and stay in the sample) to report his disease incidence in 1999. This may lead to an over-sampling of survivors. If the chronic illness in question has a low survival rate, and the severity in labour productivity decline is negatively correlated with the survival rate - the shorter the survival span, the steeper the productivity decline over the span, then the estimates would be dampened by the over-representation of survivors. If the survivors have special qualities that enable them to cope with a catastrophic event better than the non-survivors, then the estimates would be again biased upwards. This could be further exacerbated if the survival rate of chronic illness has changed over the years due to medical advances in early detection and treatment. Considering that chronic illness is generally a long-term disease that needs to be managed over time, this is more likely to be a problem when the analysis is restricted to acute illnesses. In general, due to such changing effectiveness of chronic illness care over time and the business cycles of the economy, the calendar year in which a worker experiences his chronic illness may greatly influence the recovery or stabilization rate of his illness, which in turn would affect his labour market outcome. However, because calendar year dummies and event indicators in the regression are perfectly collinear with the onset year, the onset year effect cannot be controlled for.