Who bears the cost of workers' health-related presenteeism and absenteeism?

Atsuko Tanaka*

April 23, 2016

First Version: March 16, 2015

Abstract

With an aging population and a rising prevalence of chronic conditions in the United States (U.S.), it is important to understand what happens when workers suffer unanticipated reductions in productivity. This paper investigates who pays for the loss caused by labor productivity reductions—a phenomenon often described as "presenteeism" or "absenteeism"—due to a stroke. Using the Health and Retirement Study (HRS) data, I find that, in the case of older workers, the employer often pays through higher costs of labor, rather than the worker through lower wages, because wages and earnings remain at the level before the worker had a stroke despite reduced hours. The existence of such rigidity in the employment contract translates to an increase in calculated hourly wages. Thus, this study warns that wages, earnings, or salaries cannot be clearly interpreted as accurate values of the marginal product of labor.

Keywords: Aging-associated diseases; Presenteeism; Absenteeism; Workplace accommodation; Hourly wages; Measurement error
JEL code: I00, J00, J33

*Department of Economics, University of Calgary. Email: atanaka@ucalgary.ca. University of Calgary 2500 University Drive NW Calgary AB Canada T2N1N4 (403) 220-6709. I would like to thank Yukiko Asai, Herbert Emery, Xiaodong Fan, John Kennan, Carl Sanders, Hitoshi Shigeoka, and Chris Taber for discussion and suggestions. All remaining errors are mine. The previous version of this paper was circulated under the title "Earnings Effects of Productivity Reduction: Incidence of Stroke and Higher Hourly Wages."

1 Introduction

Older workers participate in the labor force more than ever. In the U.S. between 1977 and 2007, employment of workers aged 65 and older increased by 101% while the total employment (16 and over) increased by 59%.[1] With an aging population and a rising prevalence of chronic conditions (such as stroke, heart disease, and diabetes), it is important to understand what happens when workers suffer unanticipated reductions in productivity. Among commonly reported problems due to aging are energy decline, muscle function decline, and concentration lapses (Koolhaas et al. (2012)). The literature has mostly focused on the cost incurred by the government that supports workers with social security assistance (since these workers withdraw from the labor force or work very little), and little is known about who pays the cost of lost productivity when workers continue to be employed. If workers with health limitations remain employed, who bears the cost of lower worker productivity? Are employers more flexible in terms of their expectations about hours worked? Do workers who experience severe productivity reductions accept lower wages?

Many theoretical papers point out that firms would bear the risk of productivity uncertainty when workers are more risk-averse than firms (e.g. Azariadis, 1988; Michelacci and Quadrini, 2005; 2009; Bernhardt and Timmis, 1990; Burdett and Coles, 2003). The resulting contractual arrangements would insure workers against unexpected productivity reductions by offering wages that do not depend on realization of productivity shocks due to a fixed salary (e.g. Baily, 1974; Gordon, 1974; Azariadis, 1975; Danziger, 1988). Despite abundant discussions in the theoretical literature, there is relatively little empirical evidence that shows how flexible or inflexible wage responses are to productivity shocks. Guiso et al. (2005) and Cardoso and Portela (2005) are among the few empirical papers that estimate how much firms insure workers against transitory productivity shocks resulting from demand changes. Guiso et al. (2005) use Portuguese data and Cardoso and Portela (2005) use Italian data. Both papers find that firms have wage responses less sensitive to permanent productivity shocks and find evidence that firms protect workers from risks as a form of wage insurance. The importance of implicit insurance is reconfirmed in the study by Guiso et al. (2013). In contrast, Kátay (2008) find little evidence of wage insurance using the data from Hungary, noting the importance of heterogeneity in the European labor market.

Unlike the existing empirical papers, this paper looks at the productivity shocks due to workers' health, not due to changes in market demand, and helps us understand how workers and employers are impacted by workers' unanticipated health shocks. Using the Health and Retirement Study (HRS) data, I look at workers who remain in the workforce after health shocks due to a stroke and find that they often continue receiving their pre-stroke salary even though their productivity is reduced. With unobserved heterogeneity taken into account, I find evidence that employers bear much of the cost of the productivity shock since wages and earnings remain at their pre-shock level despite reduced hours. The finding that employers bear the cost of lower worker productivity due to health shocks goes along with the theory of implicit employment contract in the economics literature mentioned above.

Among the available indicators for abrupt health reductions, I focus on health shocks from a stroke because a stroke is a common disease that suddenly strikes workers and often leaves them with severe work limitations. A stroke occurs when blood circulation to the brain fails. Decreased blood flow

[1] This increase was not merely driven by the baby-boom population because those baby-boomers who were born between 1946 and 1964 had not reached the age of 65 in 2007. The proportion of those who worked full-time increased from 48% to 56% (Data source: U.S. Bureau of Labor Statistics). This increase has accelerated since individuals in the baby-boom generation began reaching the age of 65 in 2011. As of 2014, workers aged 65 and older accounted for 5.4% of the civilian labor force, and those aged 55 and older accounted for 21.7%.

results in a lack of oxygen, leading to death of brain cells. As such, a stroke permanently damages brain function, often leaving a part of the body paralyzed, deteriorating cognitive skills, and/or impairing vision.[2] While stroke is the second most common cause of death worldwide after heart disease during the past decade (World Health Organization, 2012), the survival rate after the first stroke is relatively high, making it a major cause of morbidity. Survivors are often left with lasting brain damage and/or permanent disabilities and thus require long-term rehabilitation and care. Among stroke symptoms are sudden weakness; paralysis or numbness of the face, arms, or legs; trouble speaking or understanding speech; and trouble seeing.[3] Stroke symptoms typically start suddenly, over seconds to minutes, and a stroke is unlikely to be anticipated by an individual. As such, the incidence of stroke likely induces abrupt and substantial changes in labor productivity that may affect employment contract as well as workers' work decisions.

My empirical analysis sits on three pillars. First, I investigate how a stroke affects workers' productivity by looking at changes in their health indicators and labor supply before and after a stroke. I also study differences in health changes after a stroke between workers who continue to work and those who withdraw from the labor force.

Second, I investigate what happens to workers' labor supply and earnings after a stroke using the HRS data. To be specific, I estimate the impact of health shocks caused by a stroke on days worked, hours worked, job switch, and pecuniary compensation for the workers who continue to work following a stroke. The main strategy is in the spirit of the differences-in-differences method. That is, I will have individuals who have suffered from a stroke in a certain age-sex category as a treated group and those who have not suffered from a stroke in the corresponding age-sex category as a control group. I look at the same individual before and after the first incidence of stroke and compare the difference in labor market outcomes with individuals with the same age-sex range in a control group. In doing so, I estimate the average treatment effect of having had a stroke on labor market outcomes.

Last, I conduct a robustness check using different sub-samples and different estimation methods. For example, I restrict the sample to those younger than 55 in order to avoid having my estimates affected by early retirement decisions. To complement the main analysis, I use the fixed-effect estimator to estimate the difference in the labor market outcomes before and after the stroke.

The finding of this paper sheds light on the labor market arrangement that workers with sudden health reduction face. Studies point out that the large cost incurs when workers go to work when they cannot fully function due to a medical illness (presenteeism) and when sick employees are absence from work (absenteeism). For example, Weaver (2010) estimated that the annual cost of presenteeism in 2010 in the U.S. was $180 billion and that the cost of absenteeism was $118 billion. Compared to other industrialized countries, presenteeism is thought to be more serious in the U.S., which is the only developed country without federal sick leave insurance. Susser and Ziebarth (2016) study the 2011 American Time Use Survey and find that the average presenteeism rate is 2.1%, which amounts to three million U.S. employees. They find that presenteeism is more likely to be observed for female workers and those who work at wages lower than $30 per hour. Böckerman and Laukkanen (2010) find some evidence that the rates of presenteeism and absenteeism are negatively correlated with one another and presenteeism is more prevalent for those who are engaged in permanent full-time work.

[2] The HRS data offers the information of seven other diseases as well as stroke, but the other diseases do not strike the workers as suddenly as stroke and the data show that both hourly and weekly earnings decrease as the individuals report one of each diseases, which is different from the impact of stroke. Therefore, I do not report the effects of diseases other than stroke.

[3] Source: http://www.nhlbi.nih.gov/

By focusing on absenteeism and presenteeism associated with stroke among elderly workers, this paper examines the labor market conditions those stroke survivors face.

In addition to answering who pays the cost of lower worker productivity, this paper points out that when economists use hourly wages in their analyses, it appears that they are not measuring productivity itself but measuring productivity cost increases borne by the employer. As such, the findings in this paper imply that wage variables naively created by dividing labor earnings by work hours could contain significant measurement errors for evaluating older workers' productivity. While many papers document errors in self-reported studies due to the fact that respondents often inflate or round hourly wage rates (e.g., Bound et al. (2001); Baum-Snow and Neal (2009)), few papers have examined health-induced errors in the measurement of hourly wages. The results in this paper serve as a cautionary note for researchers who construct hourly wage data by using information from earnings and hours worked. Thus, This paper contributes to the literature by providing a broader implication about measurement errors of productivity.

This paper proceeds as follows. Section 2 summarizes the data and presents descriptive statistics. Section 3 outlines the empirical strategy and presents the results. Section 4 examines a set of hypotheses and discusses the interpretation of the results. Section 5 concludes.

2 Data

I use the University of Michigan Health and Retirement Study (HRS) from 1992 to 2008, which is designed to explore the changes in labor force participation and the health transitions of elderly individuals. The HRS is a biennial longitudinal panel study that interviews a representative sample of Americans born before 1953 and their spouses. This paper's estimates are based on the RAND-HRS data, a consolidated version of the HRS, which the RAND Corporation has produced with consistent definitions of variables across the nine survey waves. The data has over 150,000 responses including multiple observations of the same individuals. The initial sample in the 1992-2008 HRS data contains observations whose age ranges from 22 to 109. The average age of the whole sample in the first wave is 55.

2.1 Sample Universe

The sample for the main analysis consists of data taken from men and women under the age of 70 who have participated in the HRS surveys.[4] This sample restriction based on age reduces the sample size to 190,545 person-year observations. In conducting a robustness check, I later restrict the sample to those younger than or equal to 55-years-old in order to avoid having my estimates affected by early retirement decisions.[5] There are 22,687 unique individual observations in the first wave. Among the individuals in my sample, 9% of them permanently drop out from the sample and 21% of them pass away in the middle of the panel survey. I exclude those who had a stroke before they entered the survey in order to evaluate sudden changes in health and productivity that occurred to healthy workers.[6]

[4] Those over age 70 are excluded, because in the data, few people in this age cohort work; less than 10% males and less than 5% of females work in their 70s in the 1992-2008 HRS data.

[5] There are 54,980 person-year observations for those under age 55.

[6] Studying workers with aftereffects of stroke is important but complicates the scheme of the analysis by adding more factors to consider. For example, those stroke survivors may receive disability insurance, which substantially affects their labor supply decisions. Such analysis is beyond the scope of this paper, and I leave it to the future work.

Table 1: Statistics of Demographics in the HRS data

		Age 50-54	Age 55-59	Age 60-64	Age 65-69
Variable					
Gender	Male	48.8%	45.8%	47.2%	48.8%
	Female	51.2	54.2	52.8	51.2
Race	Caucasian	79.5%	81.1%	82.3%	83.6%
	Black	13.5	14.0	13.8	13.0
	Other race	7.0	4.9	3.8	3.5
Stroke	Never had	95.9%	93.6%	90.6%	87.5%
	First time survivor	3.5	3.6	4.2	5.0
	Stroke survivor	0.6	2.7	5.2	7.5
# Observation		9,597	26,537	24,219	20,947

Source: the 1992-2008 HRS data.
Notes: The number of observations is the number of unique individuals observed at least once in the corresponding age period.

Table 1 shows summary statistics. In my sample, 46% are male and 54% are female, and 80% are White and 17% are African American, which is roughly proportional to the demographic composition of the U.S. population aged over 50. The principal measure used for health shocks is the incidence of stroke. The data tells us whether the survey respondents and their spouse had a major illness, including a stroke, since the last survey interview.[7] The data also contains the health and employment status at the time when the survey was conducted. Since the HRS data is collected every other year, I set the unit of analysis to two years and look at the average treatment effects of having a stroke in the past two years on labor market outcomes at the time of surveys.

The key outcome variables in this analysis are health conditions, labor force participation, hours worked, weekly earnings, and hourly earnings. All these variables are recorded in the RAND-HRS data set. The workers' health conditions at the time of survey are self-reported by respondents in the HRS data. For example, the data tells us about workers' disability status and what workers are capable of doing (e.g., carrying heavy things, climbing up stairs, etc.). The labor force participation variable is recorded as "RwLBRF." RwLBRF records the labor force status for the respondent at each survey panel as working full-time, working part-time, unemployed, partly retired, retired, disabled, or not in the labor force.[8] Hours variables are "RwJHOURS," which is the answer to the survey questions about usual hours worked per week if respondents answer that they work. The variables of reported hourly and weekly wages for currently employed individuals are recoded as "RwWGIHR" and "RwWGIWK." The wages are calculated using the usual hours worked per week, usual weeks worked per year, and pay rate. These questions are asked in all survey waves. Wages do not include non-labor earnings such as disability income or other government transfers.

The incidence of a stroke is frequently observed in the data since most respondents in the HRS data are older people. The proportion of the respondents who answered that they had had a stroke during

[7]For the incidence of stroke, I use the data on the raw response to the question regarding whether or not a doctor has told the respondent had a stroke since their last interview. In addition, I use the data on whether the respondent reported to have had a stroke in the past.

[8]Similarly, SwWGIHR, SwWGIWK, SwJHOURS, and SwLBRF summarize the labor market outcomes for the respondent's spouse or partner.

the survey period between 1992 and 2008 is a relatively high 18%.[9] When we look at the incidence of stroke by age, 3.7% of uniquely observed individuals had their first stroke between age 55 and 59. Among those who never had a stroke until age 59, 4.2% had a stroke between age 60 and 64, and 5.2% had their first stroke between age 65 and 69. The number of stroke survivors in my sample aged under 70 is 4,673 which accounts for 18% of the total observations and is sufficiently large to conduct statistical analysis.

2.2 The Treated and the Control Groups

The respondents are classified into a treated group if they had a stroke for the first time in a certain age range and a control group if they have never had a stroke in the corresponding age range.[10] To estimate the effect of the first stroke, I compare the treated group and the control group during the corresponding age period after controlling for observable differences in health and productivity in the data.

Table 2 shows how the treated group and the control group are different in observable characteristics before treatment (having a stroke). On average, those in the treated group have worse health conditions, have lower household assets, and are less likely to be working even before they had their first stroke than those in the control group. These differences in the initial conditions between these two groups will be addressed in the main estimation analysis of a differences-in-differences method.

2.2.1 Health Outcomes and Presenteeism due to a Stroke

Having a stroke for the first time results in substantial changes in an individual's life by leaving a part of the body paralyzed, damaging brain function, and/or deteriorating cognitive skills. As such, a stroke leads to presenteeism when stroke survivors continue to work because these workers cannot function as well as before.

The raw data exhibits a substantial deterioration in the health measures of the respondents after their first stroke. These health indicators serve as a more direct proxy of workers' physical health. Figure 1 plots trends in physical health changes for those who have had a stroke. For presentation purpose, I present the changes in three different health indicators among the three health measures available in the data: difficulty in lifting things, difficulty in climbing up staircases, and difficulty in picking up a dime off of a table. These variables take one if the respondents answer that they have difficulty in the activity asked in the survey and zero otherwise. The vertical axis plots the following three measures. The greater these indicators are, the worse the health conditions of the respondent. The horizontal axis shows the time elapsed around the first-time stroke. The 0 point on the horizontal axis represents the point at which the individuals had a stroke for the first time in their lives. Since the survey is conducted every two years, the unit is two years. For example, -4 on this axis indicates 7 to 8 years before the stroke, while 4 indicates 7 to 8 years after the stroke. Regardless of the measures used, the changes in health conditions are large. The proportion of those who have difficulty in daily activities increases by 12 percentage points from 25% to 37% for lifting things, by 9 percentage points from 21% to 30% for climbing one stair, and by 7 percentage points from 6% to 13% for picking up

[9]For those who are aged between 50 and 59, the proportion of stroke survivors is 16.9%. For those between age 60 and 69, the proportion is 23.6%.

[10]The data contains the two variables: whether the respondent reported a stroke in the survey year and whether he/she has reported having a stroke.

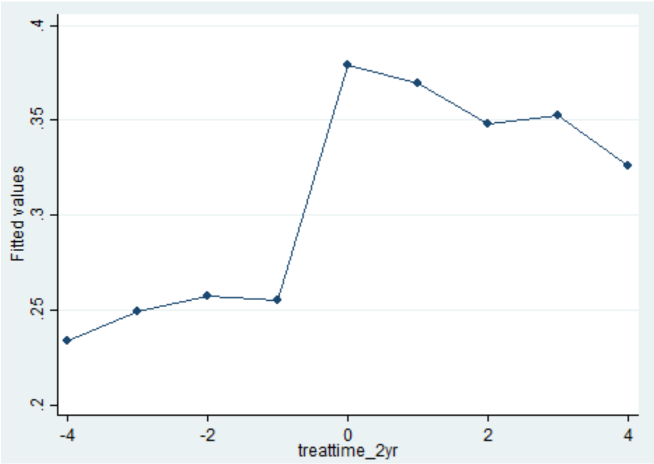
Table 2: Worker Characteristics by the Incidence of Stroke

Sample: Those who had never had a stroke at the time of the first survey (1992)						
Age in 1992	50 to 54			55 to 59		
Whether had a stroke 1993–2008	Control	Treated		Control	Treated	
Marital Status (%)						
Currently Married	74.9	79.9	***	72.6	80.1	***
	(0.5)	(0.9)		(0.8)	(1.6)	
Never Married	3.5	2.4	**	3.9	2.1	**
	(0.2)	(0.3)		(0.4)	(0.6)	
Labor Market						
Currently Working (%)	65.4	51.5	***	66.3	54.7	***
	(0.5)	(1.1)		(0.9)	(2.0)	
Hourly Wage	17.6	36.5	*	17.8	12.2	
	(1.9)	(23.1)		(3.3)	(0.5)	
Weekly Wage	623.9	541.4		614.6	479.9	**
	(21.2)	(39.7)		(28.2)	(19.5)	
Hours worked	40.7	39.6	**	40.5	39.5	
	(0.2)	(0.4)		(0.3)	(0.7)	
Weeks worked	49.5	49.5		49.6	49.5	
	(0.1)	(0.2)		(0.2)	(0.4)	
Household Asset	89021.9	71350.2	***	90097.4	74727.9	***
	(1080.2)	(1834.4)		(1984.7)	(3396.1)	
Other Disease (%)						
High Blood Pressure	37.3	51.4	***	37.4	53.7	***
	(0.5)	(1.1)		(0.9)	(2.0)	
Diabetes	9.7	17.5	***	9.8	18.2	***
	(0.3)	(0.9)		(0.6)	(1.5)	
Cancer	5.4	5.9		5.6	6.2	
	(0.2)	(0.5)		(0.4)	(1.0)	
Lung Disease	7.5	11.7	***	7.9	10.0	*
	(0.3)	(0.7)		(0.5)	(1.2)	
Heart Disease	11.6	22.0	***	11.9	19.1	***
	(0.3)	(0.9)		(0.6)	(1.6)	
Arthritis	37.1	45.7	***	39.4	48.0	***
	(0.5)	(1.1)		(0.9)	(2.0)	
Mental Disease	9.8	15.5	***	9.7	14.5	***
	(0.3)	(0.8)		0.6	1.4	
No. of observations	7,179	1,463		1,928	597	

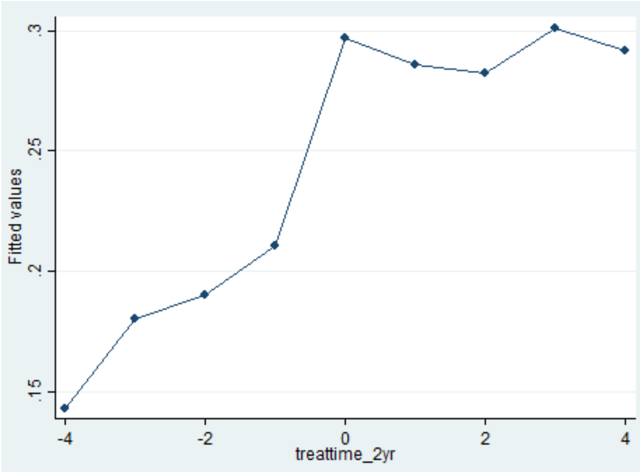
Notes: All sample are those who had not had a stroke in 1992, at the time of the first survey. The two age cohorts (50-54, 55-59) in the first survey are divided into the "ever had stroke" group and the "never had stroke" group. That is, the sample in the first wave is stratified depending on whether they experienced a stroke sometime between 1993 and 2008. Regardless of the age cohort, I find that the two groups are different in the proportion of married individuals, the proportion of employed workers, household asset levels, and the incidence of other diseases at the time of the first interview. However, I cannot find a statistically significant difference in wage and work hours from the raw data statistics. *, **, and *** indicate that the difference between two statistics is different from zero at the 1, 5, and 10 percent significance levels, respectively. More asterisk marks mean that the corresponding characteristics are more significantly different between the two groups. Standard errors are presented below each set of statistics. The variables for other diseases indicate 1 if the respondent has ever suffered from the symptom prior to the first interview and 0 if otherwise.

Figure 1: The Impact of the First-time Stroke on Physical Health

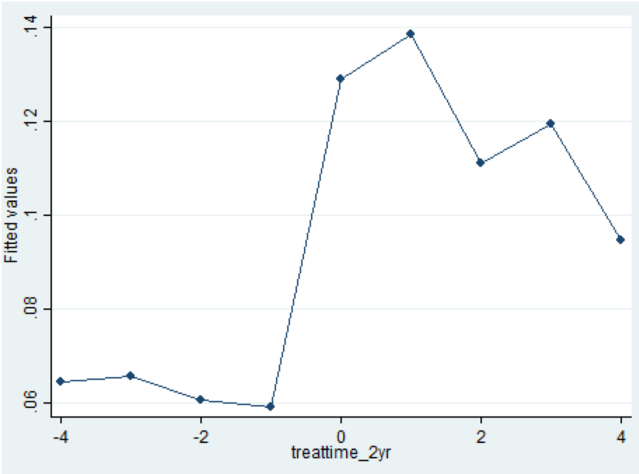
Difficulty in lifting things



Difficulty in claiming one stair



Difficulty in picking up a dime



Source: the 1992-2008 HRS data.

Notes: The graph shows the proportion of the respondents who reported the difficulty on the vertical axis and the time elapsed with unit of two years on the horizontal axis. The treated group consists of those who had their first stroke between ages 55 and 59. The control group consists of those who had never had a stroke until age 59. For the treated group, I plot the health measures over age from eight years before a stroke to eight years after a stroke for stroke survivors. For the control group, I set age 57 as the time equal to zero.

a dime. A comparison of the fraction of individuals in the average health conditions before and after stroke shows that having a stroke abruptly deteriorates health conditions to a great extent.

To show how people's health changes before and after their first stroke, I first break individuals into two broad age categories: age 50 to 59 and age 60 to 69. Summary statistics of health indicators for individuals in their 50s are reported in Table 3 and those for individuals in their 60s are reported in Table 13 in the Appendix. I divide each of these age categories into the treated and the control group, respectively. For example, the treated group for the age period 50-59 are those who had a stroke between 50 and 59; the control group includes those who did not have a stroke in the corresponding age period. The treated and the control groups are further divided into age categories: 50 to 54 and 55 to 59. The treated age 50-54 column is individuals aged 50 to 54 who had not had a stroke at the time of observation. The treated age 55-59 column is individuals aged 55 to 59 who had had a stroke at the time of observation. The control group is the same age category for individuals who never had a stroke. The same is true for age 60 to 69 group, presented in the Appendix. The statistics from the corresponding control group are presented side-by-side for comparison purposes. The control group for the age period 50-59 is those in the same age period as the treated group and did not have a stroke between the ages of 50 and 59. Before columns show that magnitude of difficulty that individuals had performing these three tasks before their first stroke. After columns in contrast show the difficulty after their first stroke.

Table 3: Summary Statistics for the Treated and the Control Groups

Sample	Men 50 - 59					
	Treated			Control		
Age	50-54	55-59		50-54	55-59	
Variable	before	after	diff.	no stroke	no stroke	diff.
Lift things	0.16	0.30	0.14	0.08	0.10	0.02
Climb up stairs	0.34	0.52	0.18	0.19	0.24	0.05
Pick up a dime	0.08	0.12	0.04	0.03	0.03	0.00
No. Obs.	372	1285		11729	21913	

Sample	Women 50 - 59					
	Treated			Control		
Age	50-54	55-59		50-54	55-59	
Variable	before	after	diff.	no stroke	no stroke	diff.
Lift things	0.30	0.39	0.09	0.17	0.20	0.03
Climb up stairs	0.47	0.65	0.18	0.36	0.40	0.04
Pick up a dime	0.05	0.12	0.07	0.03	0.04	0.01
No. Obs.	773	1236		18623	25952	

Source: the 1992-2008 HRS data.
Notes: Treated Group: Those who had the first stroke at age 50 to 59. Control Group: Those who never had a stroke till age 59. Statistics for age 50-54 in the treated group are from those aged 50 to 54, had not had a stroke at the time of the survey, but had (would have) their first stroke at age between 50 and 59. Statistics for age 55-59 in Treated group are from those for aged 55 to 59 who already had their first stroke between 50 and the time of survey age. Statistics for the Control group in the corresponding ages are calculated from those who did not have (would not have) their first stroke at age between 50 and 59. LFP stands for Labor Force Participation. No. Observations is the total number of observations in each category group. Some of the variables (e.g. earnings) are missing.

The health statistics in this table confirm that physical health deteriorates more for those individuals who had a stroke (treated group) than those individuals who did not have a stroke (control group). For example, 30% of the treated group, respondents in which experienced their first stroke in their 50s, reported difficulty after a stroke; however, only 10% of the control group, respondents in which did not have a stroke in their 50s, reported difficulty. Decline in physical health is observed using the other two physical health indicators.

2.2.2 Labor Supply and Absenteeism due to a Stroke

Partly due to a reduction in health and limitation in mobility, a stroke substantially affects worker's labor supply. Some of them stop working, switch from full-time to part-time, or reduce hours of work. The reduction in hours worked per week implies that worker productivity decreases when we take a time span longer than one week to measure the productivity. Thus, a stroke leads to absenteeism in a sense that employees are more likely to be absence from work due to their limited health after having a stroke. From raw data, we observe that hours worked per week decrease for those who had a stroke and continue to work.

While some stroke survivors withdraw from the labor force, many stroke survivors continue working after their first stroke. Table 4 presents the proportion of workers by employment status, which is calculated from the raw data, before and after their first stroke. As in Table 3, this table presents changes of the treated group before and after the first stroke and corresponding changes of the control group side by side for comparison purposes. For example, for people in their early 50s, 77% in the treated group are in the labor force, while 86% of the control group are in the labor force. The labor force participation (LFP) rate for individuals in their late 50s drops by 29 percentage points for the treated group (stroke survivors) from 77% to 48%, while the rate drops by 10 percentage points for the control group. Although the decrease is more substantial for the treated group, the remarkable feature is that about a half of the stroke survivors remain in the labor force after surviving a stroke. When we look at their employment status, 42% of the stroke survivors are working full-time, 4% work part-time, and 2% are unemployed. Similar features are observed for those in their 60s (reported in Table 14).

Noticeable changes can be found in the years of job tenure, which indicates years during which the individual works in the current firm at the time of survey. For men in their 60s and women, the treated group's job tenure does not grow as much as that of the control group. Such a data feature implies that the percentage of workers who change their job (conditional on working in the previous period) is higher in the treated group than the control group. For men in their 50s, however, the years of job tenure grow more for stroke survivors than the control group. This feature implies two possibilities: one explanation is that workers who have a longer tenure are more likely to continue working after a stroke; the other explanation is that conditional on working in their late 50s, male stroke survivors are less likely to switch their job.

The last three variables in the table look at statistics for workers who work full-time. Conditional on working, reduction of work hours are especially notable for those had a stroke, but changes in earnings do not exhibit coherent patterns between the treated group and the control group.

Table 5 presents the details of the results from analysis that regresses the incidence of stroke on hours worked (conditional on working full-time) using the fixed-effects estimator. By looking at the same individual who works full-time before and after his or her first stroke, I find that weekly hours of work decreased by 1.3 to 1.8 hours after suffering a stroke. When I restrict the sample to male

Table 4: Summary Statistics for the Treated and the Control Groups

Sample	Men 50 - 59					
	Treated			Control		
Age	50-54	55-59		50-54	55-59	
Variable	before	after	diff.	no stroke	no stroke	diff.
LFP	0.77	0.48	- 0.29	0.86	0.76	- 0.10
(full-time)	0.67	0.42	- 0.25	0.77	0.68	- 0.09
(part-time)	0.01	0.04	0.03	0.05	0.04	- 0.01
(unemployed)	0.02	0.02	0.00	0.03	0.02	- 0.01
Job Tenure (yrs)	11.8	13.3	1.5	14.0	14.8	0.8
Hours (full-time)	46.8	43.2	-3.6	45.6	44.3	- 1.3
Wkly Wage (full)	623	738	115	1012	1054	42
Hrly Wage (full)	13.5	17.0	3.5	22.5	29.4	6.9
No. Obs.	372	1285		11729	21913	

Sample	Women 50 - 59					
	Treated			Control		
Age	50-54	55-59		50-54	55-59	
Variable	before	after	diff.	no stroke	no stroke	diff.
LFP	0.61	0.40	- 0.21	0.69	0.59	- 0.10
(full-time)	0.43	0.30	- 0.13	0.50	0.42	- 0.08
(part-time)	0.15	0.09	- 0.06	0.16	0.12	- 0.04
(unemployed)	0.03	0.02	- 0.01	0.02	0.02	0.00
Job Tenure (yrs)	10.0	10.4	0.04	10.6	11.4	0.08
Hours (full-time)	38.9	36.6	- 2.3	37.7	36.3	- 1.4
Wkly Wage (full)	509	456	- 53	587	593	6
Hrly Wage (full)	13.9	15.4	1.5	16.3	17.8	1.5
No. Obs.	773	1236		18623	25952	

Source: the 1992-2008 HRS data.
Notes: Treated Group: Those who had the first stroke at age 50 to 59; Control Group: Those who never had a stroke until age 59. Before columns show that magnitude of difficulty that individuals had performing these three tasks before their first stroke. After columns in contrast show the difficulty after their first stroke. LFP stands for Labor Force Participation. No. Observations is the total number of observations in each category group. Some of the variables (e.g. earnings) are missing.

Table 5: Effect of Stroke on Hours Worked for Full-time Workers

Sample	All			Male		
Dep. Var.	Hours Worked			Hours Worked		
Variables						
Stroke	-1.37**	-1.86**	-1.34*	-1.75**	-2.70***	-3.27***
	(0.63)	(0.74)	(0.77)	(0.84)	(1.00)	(1.04)
Post-treat		-1.86**	-1.43		-2.80**	-3.25***
(1-2 yrs)		(0.88)	(0.91)		(1.23)	(1.25)
Post-treat		0.04	0.46		-1.05	-1.32
(3-4 yrs)		(1.01)	(1.02)		(1.48)	(1.49)
Post-treat		-0.53	0.07		-1.30	-1.34
(5-6 yrs)		(1.19)	(1.22)		(1.78)	(1.80)
Age	4.28***	4.28***	4.14***	5.33***	5.34***	5.07***
	(0.19)	(0.19)	(0.19)	(0.33)	(0.33)	(0.34)
Age^2	-0.04***	-0.04***	-0.04***	-0.05***	-0.05***	-0.05***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Industry Dummies	No	No	Yes	No	No	Yes
Occupation Dummies	No	No	Yes	No	No	Yes
No. Obs.	55,974	55,974	52,195	26,562	26,562	25,006
No. ID	15,133	15,133	14,433	7,238	7,238	6,979
R^2	0.10	0.10	0.10	0.13	0.13	0.14

Notes: the sample is those who had a stroke and continue to work full-time and those who did not have had a stroke and work full-time. Additional controls include marital status, household asset, and years of job tenure.
Standard errors, in parentheses, and clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1

workers who continue to work full-time, I find that weekly hours of work decreased by 1.8 to 3.3 hours following a stroke.

3 Estimation Method and Results

3.1 Differences-in-Differences

The main empirical strategy to estimate the impact of stroke is in the spirit of a differences-in-differences approach. The first difference is before and after the stroke, and the second difference is over age, which is a key determinant for workers' productivity, labor supply, and health conditions. As defined above, the treated group consists of those who had their first stroke within a certain age period ,and the control group consists of those who had never had their stroke during the corresponding age period.

For each age group, I implement the differences-in-differences comparison by running the following regression:

$$y_{it} = \alpha + \gamma Treated_t + \delta Stroke_{it} + \lambda_t + X'_{it}\beta + \epsilon_{it} \tag{1}$$

where i denotes the individual, t denotes age (with one unit equal to two years), and y_{it} is the outcome variable of interest (e.g., health indicator, labor market outcomes). The variable $Treated$ is a dummy for treatment group and takes 1 if the person belongs to the treatment group (had their first stroke during the age period) and 0 if the person belongs to the control group. The variable $Stroke$ is a dummy for having had the first stroke, and it takes 1 if the person already had his or her first stroke and 0 otherwise. λ_t is a vector of time-specific fixed effects to control for economic conditions. The vector X_{it} is a set of individual characteristics to control for observable differences that affect both the likelihood of receiving health shocks and outcome variables, such as other health indicators, occupation, industry, education, and asset/wealth levels. These variables are highly correlated with a worker's productivity and his/her ability to continue working after a stroke. These variables are also correlated with how firms accommodate workers after a stroke. Accommodation after a stroke may be very different depending on occupation and role in a company. Without these controls, the estimated effects will be biased if these omitted variables are correlated with an incidence of stroke. Finally, ϵ_{it} represents an error term assumed to be uncorrelated with the treatment dummy $Stroke_{it}$, conditional on the control variables. That is, for the differences-in-differences estimator to be valid, the following assumption needs to hold:

$$E(\epsilon_{it}|Treated_i, Stroke_{it}, \lambda_t, X_{it}) = E(\epsilon_{it}|Treated_i, \lambda_t, X_{it}). \tag{2}$$

I discuss this assumption in Subsection 3.3.

3.2 Estimation Results

I present the two sets of results using all workers who continue working full-time and using only male samples, respectively. Examining the effects conditional on gender gives us a more comprehensive picture because men and women may respond very differently to strokes, especially in terms of the labor market responses. The sample wages are from those who had a stroke and continue working full-time (treated) and those who did not have a stroke and work full-time. Important wage determinants such as occupation and/or industry are included as control variables in the regression analysis.

Regardless of the model specification, the estimation results show that the incidence of stroke does not reduce hourly wage or weekly wage. Tables 6 and 7 present how stroke affects earnings of workers who continue working following a stroke for all workers and for male workers only, respectively. The first three columns in each table look at the effects on hourly wage, and the last three columns in each table look at the effects of weekly wage. The coefficient of our interest, δ_0, is often estimated as significantly positive when the sample includes all workers. For example, in Column (2), the regression results indicate that, after controlling for age and other observable characteristics, workers who had a stroke within the past two years receive an hourly wage that is 10 percentage points higher than those who did not have a stroke. The positive coefficients of δ_s mean that hourly wages are on average higher than hourly wages of non-stroke survivors several years after the first stroke. This estimate does not mean stroke survivors are earning more than non-stroke survivors because the base level for stroke survivors is captured by the separate term γ and estimated as negative. That is, the estimates show that, on average, those who have a stroke receive less hourly wages than those who did not have a stroke, but their average hourly wage increases after a stroke. Column (3) uses a slightly different model specification by additionally controlling for age dummies (instead of quadratic terms), occupation dummies, and industry dummies and shows almost the same patterns as the ones in Column (2). Similar results are found when we look at weekly wages. When we restrict our sample to male workers, some of the results turn out to be insignificant, but the sign of the estimates remain the same as before.

Table 6: Labor Earnings and the Incidence of Stroke (Differences-in-Differences Method)

Sample	All					
Dep. Var.	log Hourly Wage			log Weekly Wage		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Stroke (δ_0)	0.03	0.10**	0.12***	0.02	0.09**	0.12***
(0-1 yrs)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Post-treat (δ_1)		0.22***	0.18***		0.24***	0.20***
(2-3 yrs)		(0.04)	(0.04)		(0.05)	(0.04)
Post-treat (δ_2)		0.17***	0.12**		0.20***	0.15***
(4-5 yrs)		(0.05)	(0.05)		(0.05)	(0.05)
Post-treat (δ_3)		0.22***	0.22***		0.25***	0.24***
(6-7 yrs)		(0.08)	(0.07)		(0.08)	(0.07)
Treat (γ)	-0.18***	-0.25***	-0.23***	-0.18***	-0.25***	-0.23***
	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)
Female	-0.26***	-0.26***	-0.25***	-0.35***	-0.32***	-0.32***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age	0.06***	0.06***		0.06***	0.06***	
	(0.01)	(0.01)		(0.01)	(0.01)	
Age^2	-0.00***	-0.00***		-0.00***	-0.00***	
	(0.00)	(0.00)		(0.00)	(0.00)	
Age dummy	No	No	Yes	No	No	Yes
Occupation	No	No	Yes	No	No	Yes
Industry	No	No	Yes	No	No	Yes
No. Obs.	63,861	63,861	60,469	63,870	63,870	60,478
R^2	0.04	0.04	0.28	0.06	0.06	0.30

Notes: the sample is those who had a stroke and continue to work full-time and those who did not have had a stroke and work full-time. All wages are measured by log(wage). Additional controls include marital status, gender dummy, race dummy, education years, and years of job tenure. For the analysis using the whole sample, I also include the interaction term of gender and marital status to capture the gender-specific role of marriage that could affect wages.
Standard errors, in parentheses, and clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1

The findings that wages after a stroke are on average higher than wages before a stroke are counter-intuitive since a stroke adversely affects workers' health. To see how robust the data features are, I conduct the analysis for subset of the sample.

Table 7: Labor Earnings and the Incidence of Stroke (Differences-in-Differences Method)

Sample	Male					
Dep. Var.	log Hourly Wage			log Weekly Wage		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Stroke (δ_0)	0.02	0.10	0.12*	0.00	0.08	0.12*
(0-1 yrs)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)
Post-treat (δ_1)		0.18***	0.17***		0.19***	0.20***
(2-3 yrs)		(0.06)	(0.06)		(0.06)	(0.05)
Post-treat (δ_2)		0.08	0.07		0.10	0.09
(4-5 yrs)		(0.07)	(0.07)		(0.07)	(0.07)
Post-treat (δ_3)		0.18*	0.17*		0.20*	0.20*
(6-7 yrs)		(0.11)	(0.10)		(0.11)	(0.10)
Treat (γ)	-0.11***	-0.19***	-0.22***	-0.10**	-0.18***	-0.22***
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
Female	–	–	–	–	–	–
	–	–	–	–	–	–
Age	0.08***	0.08***		0.08***	0.08***	
	(0.01)	(0.01)		(0.01)	(0.01)	
Age^2	-0.00***	-0.00***		-0.00***	-0.00***	
	(0.00)	(0.00)		(0.00)	(0.00)	
Age dummy	No	No	Yes	No	No	Yes
Occupation	No	No	Yes	No	No	Yes
Industry	No	No	Yes	No	No	Yes
No. Obs.	33,504	33,504	31,919	33,512	33,512	31,927
R^2	0.01	0.01	0.21	0.01	0.01	0.22

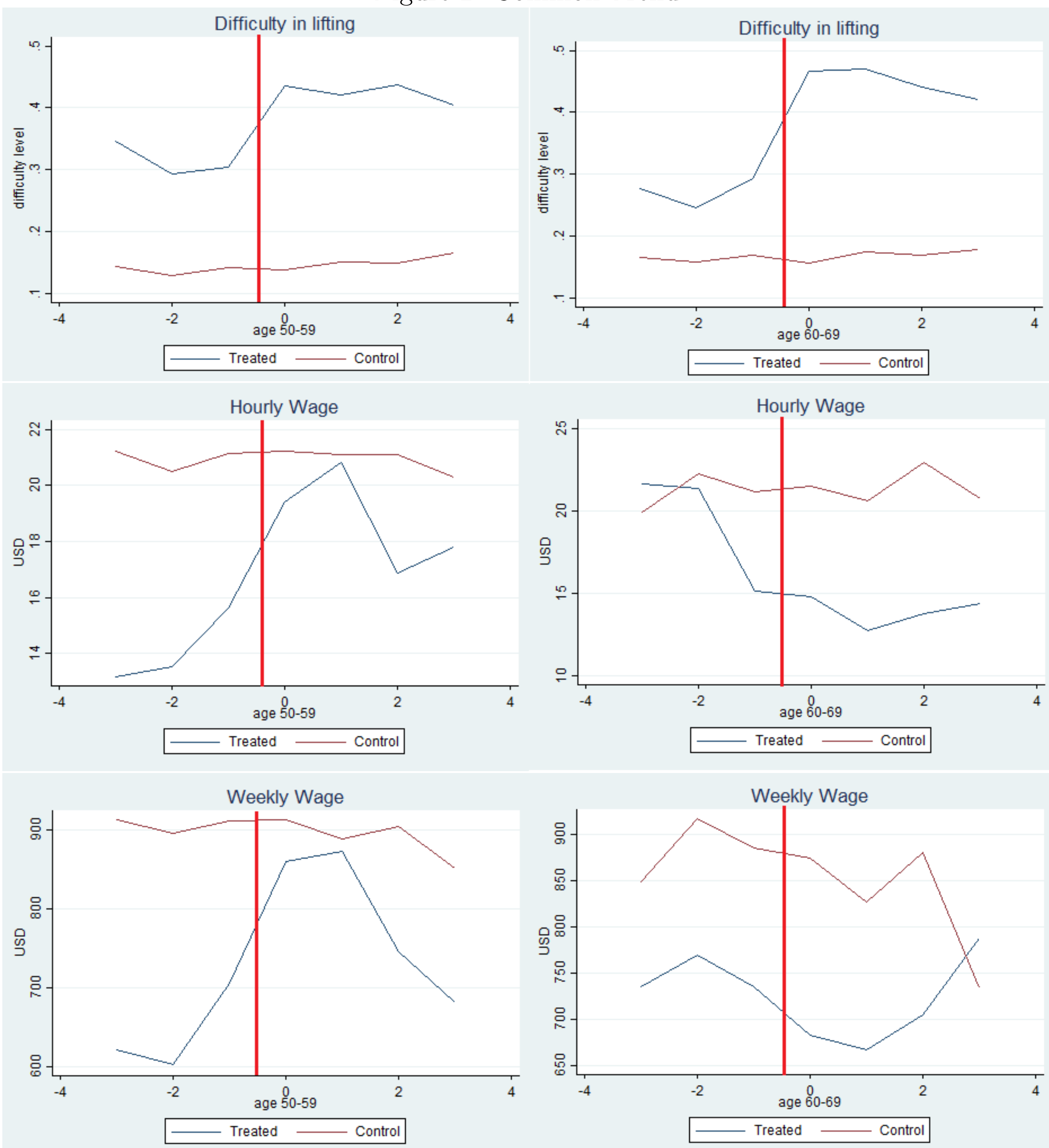
Notes: the sample is those who had a stroke and continue to work full-time and those who did not have had a stroke and work full-time. Wages are log(wage). Additional controls include marital status, gender dummy, race dummy, education years, and years of job tenure. For the analysis using the whole sample, I also include the interaction term of gender and marital status to capture the gender-specific role of marriage that could affect wages.
Standard errors, in parentheses, and clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

3.3 Endogeneity Issues

Before presenting the results by subgroup, I discuss whether or not the assumptions introduced in Section 3.1 hold. To be specific, I address endogeneity concerns by checking if Equation (2) holds. For this equation to hold, three important assumptions must be satisfied: (i) a common time trend between the treated group and the control group, (ii) no effects of stroke preceding its occurrence, and (iii) no self-selection conditional on observables.

The first key assumption is common trend assumption, which is assumes that a time effect is common across comparison groups. If the treated group and the control group have different time trend, estimation will be biased. Figure 2 presents the time trend to check if the trend of the treated

Figure 2: Common Trend



Source: Own calculations, based on the HRS data.

group is parallel to that of the control group. Here, in order to illustrate the trend line graphs, I plot the health indicators for males who had a stroke in a certain age range and those who never had a stroke in the corresponding age range.[11] These graphs in Figure 2 compare the trend differences between the control and treated groups. In general, the time trend lines run parallel between the groups before the incidence of a stroke (denoted by the vertical line). For the trend of health indicators, the common trend is obvious from the graphs: the level of difficulty in lifting heavy things runs parallel between the groups except at time 0 on the horizontal axis, which represents the timing when the treated group had a stroke. For the trend of labor market outcomes, however, we observe non-trivial changes for the treated group in time -1, which is one to two years prior to the incidence of a first stroke. For example, hourly wages and weekly wages slightly increase for the treated group aged 50 to 59, and hourly wages decrease for the treated group aged 60 to 69. Weekly wages for those aged 60 to 69 show a parallel trend except for time 3, in which we have few observations.

The second key assumption is that there are no effects of stroke preceding its occurrence. Although a stroke usually occurs suddenly without preceding symptoms, the health of an individual might deteriorate gradually prior to the incidence of a stroke. To investigate if having a stroke affects health and labor market outcomes over time (pre-treatment effects and post-treatment effects), Equation (1) is generalized by adding lagged and leading terms of the stroke variable:

$$y_{it} = \alpha + \gamma Treated_{it} + \sum_{s=t-4}^{t+4} \delta_s Stroke_{is} + \lambda_t + X'_{it}\beta + \epsilon_{it} \tag{3}$$

I estimate Equations (1) and (3) for all individuals and then for all male individuals. The test results for the overall sample and the male sample are reported in the Appendix in Tables 15 and 16, respectively. Columns (3) and (6) in each table show the regression results of Equation 3. The coefficients on lagged terms *(0-1 yrs before* and *2-3 yrs before* in Columns (3) and (6) in Tables 15 and 16) are both insignificant in the regression. These results indicate that there are no time trend specific to the treated group before the incidence of a stroke.

The third key assumption is no self-selection of having a stroke—also referred to as random assignment assumption. As shown in Table 2, there are differences in observable characteristics between the treated and control groups. This fact implies that a stroke does not arrive randomly; individuals with certain characteristics are more likely to experience a stroke. If a treatment (having a stroke) is not assigned randomly, the estimation can be contaminated due to omitted variable bias. Omitted variables bias may exist because people who earn higher wages might be at a greater risk of suffering a stroke due to relatively high stress. For example, a job with higher pay could give workers more duties and tasks or be associated with a higher risk of lay-offs, which create a more stressful environment for workers. Since stress triggers stroke or other serious diseases, we would observe a positive correlation between hourly wage and the incidence of stroke, not because having a stroke increases wage payments, but because higher wage payments result in a higher chance of a stroke. Since the HRS data ask individuals about hours worked per year, we cannot exclude the possibility that workers had a stroke following increasing work load or transferring to a more stressful job in exchange for higher wage compensation.

As a simple way of investigating if strokes follow a wage increase or vice versa, I check whether stress levels are correlated with stroke. By applying a fixed-effects model, I regress the incidence of each disease on job stress in the previous period. The results are shown in the Appendix in Table 17 .

[11] For labor market outcomes, the sample consists of male individuals who work full time.

The results show that job stress is correlated with all diseases except stroke and cancer. Such findings suggest that the correlation between the previous work condition and the incidence of stroke is not detected. Thus, the main findings cannot be attributed to the reverse causality that high hourly wages result in a higher chance of stroke.

Another way to address omitted variable bias is to add health indicators as covariates (X) in regression analysis. In doing so, I provide further statistical and institutional (mechanical) evidence that my main findings do not significantly suffer from omitted variable bias and rather capture the effect of a stroke itself. To be specific, I include in the regression analysis the health variables that can predict the incidence of a stroke as control variables X_{it}.

Scientific research shows that the greatest risk factor for a stroke is hypertension (high blood pressure). Hypertension often results in a blockage of a blood vessel in the brain or neck and is responsible for about 80 percent of strokes. This type of stroke is called an ischemic stroke.[12] According to the National Institute of Neurological Disorders and Stroke (NINDS), other risk factors also associated with hypertension are diabetes, cigarette smoking, and common heart disorders such as coronary artery disease, valve defects, irregular heart beat (atrial fibrillation), and enlargement of one of the heart's chambers. Studies also show the risk of stroke depends on age, gender, race, and family history.[13] As such, I assume that the likelihood of having a stroke is correlated with blood pressure, diabetes, lung disease, heart disease, arthritis, age, gender, and race, but it is random after controlling for these health indicators and other factors.

By controlling for these risk factor variables, I regress hourly wages on the propensity of having a stroke. The results of the regression analysis are reported in the Appendix in Tables 15 and 16 for the overall sample and the male sample, respectively. Risk factor covariates are included in all specifications except the baseline specifications presented in Columns (1) and (4). Compared to the baseline results in Column (1), the results in Column (2) indicate very similar results with respect to the post-treatment effects of stroke on hourly wages. Column (3) further includes the pre-treatment terms (*0-1 yrs before* and *2-3 yrs before*) discussed above. The estimated effects are not significant any more, but the direction of effects remains positive. These findings are similar when I regress weekly wages instead of hourly wages and when I use the male sample only, as reported in the tables.

Based on these combined results, which suggest that omitted variable bias is not a serious concern in my analysis, I believe that conditional independence assumption, specified by Equation (2) holds after controlling for the risk factors for a stroke.

3.4 Do the Effects of a Stroke Vary with Age or Job Tenure?

Tables 8 to 10 present estimates of the effects from a stroke on earnings for different subsamples of male individuals. First, I divide the sample by age group and present the results in Tables 8 to 9. In this analysis, I compare the average wage of stroke survivors before and after their first stroke and that of workers who did not have a stroke, assuming a common time (age) trend in earnings after controlling for observable characteristics. While the estimates become insignificant, which may be partly due to

[12] There are two broad categories of stroke: those caused by a blockage of blood flow (called ischemic stroke) and those caused by bleeding into the brain (called hemorrhagic stroke). (Source: http://www.ninds.nih.gov/disorders/stroke/preventing_stroke.htm#Warning Signs)

[13] The NINDS reports that men have a higher risk of stroke, but more women die from stroke. Studies show that the age-adjusted incidence of stroke in the United States is approximately twice as high in African Americans and Hispanic Americans as in Caucasians. The findings of these studies are consistent with the data features presented in Table 2.

the small sample size, the sign of the earnings effects mostly remain positive regardless of the model specification for any age group.

For each age group and each dependent variable, I run regression with three different specifications. For example, using the sample of males aged 50 to 54, I run regression of hourly wage with different model specifications and reported in Columns (1) to (3). Regardless of specifications, important wage determinants such as marital status dummy, gender dummy, race dummy, education years, and years of work experience are included as controls. In addition to these controls, the model presented in Column (1) includes as covariants the concurrent effects of stroke term and treated group dummy as well as age and age squared terms. Additionally, the model in Column (2) includes post-treated terms δ_1, which captures how having a stroke affects hourly wages in the post-stroke period (2-3 years after the stroke). The model in Column (3) adds more control variables to the second specification, including occupation dummies and industry dummies. It also uses age dummies (four binary dummies for age 51, 52, 53, and 54) instead of treating age as a continuous variable. As a result, these regression results show a higher R-squared value (0.21), meaning that the model has more explanatory power. Similarly, I run regression of weekly wage and reported in Columns (4) to (6). The results are the qualitatively same as the regression of hourly wage.

To further investigate what drives non-negative effects of a stroke on wage, I allow potentially different effects of a stroke across workers with different work history. I divide the sample into five groups based on their years of job tenure, which indicates how many years the respondent stays in the current firm. According to this analysis, I find negative wage effects of a stroke for those workers who change their job after a stroke. The results from a series of analysis lead to the conclusion that data features favor the hypothesis that non-decreasing weekly/monthly earnings reflect that employers cover the cost of productivity reduction and pay the same as before.

3.5 Robustness Check

To examine how sensitive the results are to selection of estimation method, I conduct an analysis to estimate the effect of a stroke in a slightly different way than before. Instead of the differences-in-differences method, I apply Heckman's two-step method combined with the Instrument Variable (IV) method. These methods complement the differences-in-differences methods by providing estimates using a different model assumption and presenting data features of our interest from different perspectives.

3.5.1 Fixed-effect estimators

Instead of looking at changes in differences between the control group and treated group, the fixed-effect estimators allow us to look at differences within the same individual before and after he or she had a stroke. To look at changes within the same individual, I use the fixed-effects estimator as follows:

$$y_{it} = \alpha + \delta Stroke_{it} + \lambda_t + X_{it}'\beta + \eta_i + \epsilon_{it} \tag{4}$$

where i denotes the individual, t denotes time, and y_{it} is the outcome variable of interest (e.g., health indicator, labor market outcomes). The variable $Stroke$ is a dummy for having had the first stroke and takes 1 if the person already had his or her first stroke and 0 otherwise. λ_t is a vector of time-specific

Table 8: Labor Earnings and the Incidence of Stroke for Male by Age

Sample	Male Age 50 - 54						Male Age 55 - 59					
Dep. var.	log (Hourly Wage)			log (Weekly Wage)			log (Hourly Wage)			log (Weekly Wage)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Stroke	0.11	0.14	0.15*	0.08	0.10	0.11	-0.05	0.03	0.06	-0.06	0.02	0.08
(0-1 yrs)	(0.11)	(0.12)	(0.08)	(0.12)	(0.12)	(0.11)	(0.07)	(0.10)	(0.10)	(0.08)	(0.10)	(0.09)
Post		0.21*	0.13		0.13	0.09		0.22**	0.25**		0.23**	0.27**
(2-3 yrs)		(0.12)	(0.10)		(0.12)	(0.10)		(0.10)	(0.09)		(0.10)	(0.09)
Post								0.14	0.13		0.10	0.09
(4-5 yrs)								(0.10)	(0.10)		(0.10)	(0.10)
Post								0.08	0.15		0.05	0.14
(6-7 yrs)								(0.14)	(0.12)		(0.15)	(0.12)
Treat	-0.14**	-0.17**	-0.18**	-0.09	-0.11*	-0.11**	-0.11**	-0.19**	-0.23**	-0.10**	-0.18**	-0.22**
	(0.06)	(0.06)	(0.04)	(0.06)	(0.07)	(0.05)	(0.05)	(0.08)	(0.07)	(0.05)	(0.08)	(0.07)
Age	-0.78*	-0.77*		-0.85*	-0.85*		0.01	0.02		-0.30	-0.29	
	(0.44)	(0.44)		(0.45)	(0.45)		(0.34)	(0.34)		(0.34)	(0.34)	
Age^2	0.01*	0.01*		0.01*	0.01*		-0.00	-0.00		0.00	0.00	
	(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)	
Age dum.	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Occup.	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Industry	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
No. Obs.	7,971	7,971	7,685	7,973	7,973	7,687	7,971	7,971	7,685	7,973	7,973	7,687
R^2	0.00	0.00	0.21	0.00	0.00	0.22	0.00	0.00	0.21	0.00	0.00	0.22

Notes: Wages are log(wage). Additional controls include marital status, gender dummy, race dummy, education years, and years of job tenure. For the analysis using the whole sample, I also include the interaction term of gender and marital status to capture the gender-specific role of marriage that could affect wages. The terms that capture post-treatment effects are not included since the sample contains very few observations that recorded having had a stroke 4-10 years prior to the age range 50-54. This lack of observations is due to the sample design of the HRS data. Since there are no retrospective data, the data on having had a stroke four or more years prior to ages 50-54 require that individuals need to enter the survey earlier than 45-49. The HRS data started collecting the data on the cohorts who were aged 50-54 and their partners in the first survey wave, and thus the sample contains few observations of individuals in their 40s. Standard errors, in parentheses, are clustered at the individual level. ** p<0.05, * p<0.1.

fixed effects to control for economic conditions. The vector X_{it} is a set of individual characteristics to control for observable differences that affect both the likelihood of receiving health shocks and outcome variables, such as occupation, industry, education, and asset/wealth levels. Unlike the previous specification, I include individual fixed-effects η_i. In doing so, I can control for unobserved differences across individuals as long as the differences are time-invariant. That is, wage differences due to time-invariant wage determinants such as gender and race are taken into account as a part of individual fixed-effects. The treatment effect of interest will be captured by the estimates of δ, the coefficient of a stroke dummy.

I find that statistical association between the incidence of stroke and earnings is insignificant. Table 11 summarizes the regression results. I find that hourly wages increase and weekly wage decrease, but all estimates are not significantly different from zero. The results reinforce the previous findings that wages do not decrease following a stroke and provide evidence for inert nature of employment contract. As documented before, I find that hours worked are significantly reduced after a stroke. The findings

Table 9: Labor Earnings and the Incidence of Stroke for Male by Age

Sample	Male Age 60 - 64						Male Age 65 - 69					
Dep. var.	log (Hourly Wage)			log (Weekly Wage)			log (Hourly Wage)			log (Weekly Wage)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Stroke	0.06	0.09	0.09	0.02	0.13	0.14*	-0.07	0.31	0.36	-0.13	0.20	0.23
(0-1 yrs)	(0.07)	(0.08)	(0.08)	(0.07)	(0.09)	(0.09)	(0.15)	(0.22)	(0.24)	(0.15)	(0.18)	(0.18)
Post		0.10	0.00		0.24**	0.18*		0.42*	0.59**		0.31	0.45**
(2-3 yrs)		(0.12)	(0.11)		(0.11)	(0.11)		(0.24)	(0.27)		(0.19)	(0.21)
Post		-0.12	-0.13		0.04	0.04		0.66**	0.79**		0.62**	0.73**
(4-5 yrs)		(0.12)	(0.13)		(0.13)	(0.14)		(0.22)	(0.24)		(0.19)	(0.19)
Post		0.31**	0.21*		0.43**	0.34**		0.33	0.56		0.33	0.51
(6-7 yrs)		(0.14)	(0.11)		(0.13)	(0.11)		(0.38)	(0.42)		(0.36)	(0.38)
Treat	-0.11**	-0.15*	-0.17**	-0.11**	-0.22**	-0.25**	-0.15	-0.53**	-0.76**	-0.11	-0.44**	-0.65**
	(0.05)	(0.08)	(0.07)	(0.05)	(0.08)	(0.08)	(0.11)	(0.20)	(0.22)	(0.11)	(0.15)	(0.16)
Age	1.39**	1.37**		1.14**	1.14**		-0.90	-0.79		-1.39	-1.32	
	(0.56)	(0.56)		(0.56)	(0.56)		(1.08)	(1.08)		(1.07)	(1.07)	
Age^2	-0.01**	-0.01**		-0.01**	-0.01**		0.01	0.01		0.01	0.01	
	(0.00)	(0.00)		(0.00)	(0.00)		(0.01)	(0.01)		(0.01)	(0.01)	
Age dum.	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Occup.	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Industry	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
No. Obs.	8,254	8,254	7,964	8,258	8,258	7,968	2,735	2,735	2,496	2,735	2,735	2,496
R^2	0.00	0.00	0.22	0.00	0.00	0.22	0.01	0.01	0.20	0.01	0.01	0.20

Notes: Wages are log(wage). Additional controls include marital status, gender dummy, race dummy, education years, and years of job tenure. For the analysis using the whole sample, I also include the interaction term of gender and marital status to capture the gender-specific role of marriage that could affect wages.
Standard errors, in parentheses, and clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

altogether imply that weekly or monthly salary rates may decrease following a stroke because workers reduce their intensive labor supply, but conditional on working, their wage rates do not decrease.

3.5.2 Heckman's Two-step Method

Applying a fixed-effect model is not sufficient to estimate the unbiased effects of stroke, however. As is well known among economists, earnings regression estimation suffers from endogeneity due to self-selection bias. In the context of this study, self-selection bias occurs because I do not observe labor earnings for all the treated workers (stroke survivors), but only those stroke survivors who are healthy enough to continue being employed or lucky enough to receive a high wage that compensates their discomfort from stroke symptoms. A fixed-effects estimator, which controls for time-invariant individual heterogeneity, can address the sample bias that arises due to unobservable characteristics across individuals, but given that a stroke changes characteristics within the same individual, we cannot consider that the individuals are the same before and after a stroke. Rather, they must experience something different besides a reduction in health.

Table 10: Labor Earnings and the Incidence of Stroke for Male by Tenure

Job Tenure	0		1-2		2-5		5-15		16+	
Dep Var	Hourly	Weekly	Hourly	Weekly	Hourly	Weekly	Hourly	Weekly	Hourly	Weekly
Variables										
Stroke (δ_0)	-0.60**	-0.76*	0.25	0.23	0.24	0.14	-0.01	0.01	0.11	0.09
(0-1 yrs)	(0.29)	(0.45)	(0.17)	(0.17)	(0.16)	(0.17)	(0.19)	(0.18)	(0.10)	(0.10)
Post-treat (δ_1)	-0.21	0.34	0.27	0.34*	0.08	-0.09	0.31**	0.29**	0.17*	0.14
(2-3 yrs)	(0.24)	(0.28)	(0.18)	(0.18)	(0.09)	(0.12)	(0.12)	(0.12)	(0.10)	(0.09)
Post-treat (δ_2)	1.47***	1.42***	0.19	0.19	-0.09	0.01	0.35***	0.25**	0.15	0.06
(4-5 yrs)	(0.26)	(0.27)	(0.18)	(0.19)	(0.22)	(0.21)	(0.11)	(0.11)	(0.10)	(0.10)
Post-treat (δ_3)			0.08	0.10	-0.48***	-0.43***	0.48***	0.49***	0.06	-0.05
(6-7 yrs)			(0.21)	(0.19)	(0.09)	(0.09)	(0.14)	(0.14)	(0.16)	(0.17)
Treat (γ)	-0.40**	-0.58**	-0.19***	-0.22***	-0.12	-0.04	-0.18***	-0.14***	-0.20**	-0.14*
	(0.16)	(0.23)	(0.02)	(0.02)	(0.08)	(0.08)	(0.06)	(0.05)	(0.08)	(0.08)
No. Obs.	82	82	2,766	2,766	2,594	2,594	5,435	5,435	9,282	9,286
R^2	0.29	0.30	0.21	0.23	0.22	0.23	0.25	0.28	0.20	0.21

Notes: Wages are log(wage). Additional controls include marital status, gender dummy, race dummy, education years, and years of job tenure. For the analysis using the whole sample, I also include the interaction term of gender and marital status to capture the gender-specific role of marriage that could affect wages.
Standard errors, in parentheses, and clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

To address the self-selection bias issue, I apply the Heckman two-step correction with labor supply decisions instrumented by potential government transfers. Using Heckman's method with potential government transfers as an instrument, I can look at two almost identical workers with significantly different potential public transfer. For example, one of the workers receives slightly lower labor earnings than the eligibility threshold (or a threshold for a higher amount of government transfer), and the other receives slightly higher labor earnings than the threshold. The former worker has less incentive to work, while these two workers are otherwise the same. The detailed implementation, including discussion and test for validity of the application, is explained in the appendix.

I report the estimation results with and without instruments. Table 12 presents the coefficient estimates of effects on hourly wage with and without instruments in Columns (1) and (2), respectively. Similarly, the estimated effects of stroke on weekly wage are presented in Columns (3) and (4). I then restrict the sample to male workers and conduct a similar analysis. The results are shown in Columns (5) to (8).

The results from Heckman's correction model are comparable to the Ordinary Least Squared (OLS) results presented as follows. The IV estimate of the impact of stroke on both hourly wages and weekly wages has large standard errors and is insignificant at the 10 percent level. The comparable results from the OLS estimator, which does not take self-selection into account, show significantly negative association between the incidence of stroke and hourly wages, but the estimated impact of stroke is five percentage points reduction and not as large as the reduction in health indicators. Besides, the correlation is weak and not significant at the 5% level. Similar results are confirmed when I restrict the sample to male workers.

The results from IV method and corresponding OLS estimator are documented to present estimates of the impact of stroke under different estimation assumptions. While the IV method would address

Table 11: Labor Earnings and the Incidence of Stroke (Fixed-Effects Estimation)

Sample	All				Male			
Dep. Var.	Hourly Wage		Weekly Wage		Hourly Wage		Weekly Wage	
Stroke (δ_0)	0.04	0.03	-0.01	-0.01	0.04	0.03	-0.01	-0.09
(0-1 yrs)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.07)
Post-treat (δ_1)		-0.04		-0.09		-0.04		-0.20**
(2-3 yrs)		(0.05)		(0.06)		(0.07)		(0.08)
Post-treat (δ_2)		-0.04		-0.01		-0.05		-0.10
(4-5 yrs)		(0.05)		(0.06)		(0.09)		(0.09)
Post-treat (δ_3)		-0.00		-0.02		-0.04		-0.11
(6-7 yrs)		(0.07)		(0.07)		(0.10)		(0.11)
Treat (γ)		-0.01		-0.06		0.04		-0.12
(7-8 yrs)		(0.08)		(0.09)		(0.13)		(0.14)
Age	0.13***	0.11***	0.29***	0.27***	0.29***	0.26***	0.09***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Age^2	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
No. Obs.	50,095	45,439	50109	50109	50,305	45,624	23,536	21,585
No. ID	14,601	13,584	14625	14625	14,633	13,614	6,971	6,561

Notes: Additional controls include marital status, education years, race dummy, occupation dummies, industry dummies, and years of job tenure.
Standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

endogeneity concerns and therefore would provide a causal argument about the relationship between the incidence of stroke and earnings, a caveat of this approach is that the estimation results are driven by sub-group of the samples, and the inference is limited when the effects are different across individuals. In this analysis, the results are driven by the difference in earnings between stroke survivors and those who participate in the labor force because of their lower amount of potential government transfer and those who withdraw from the labor force because of their higher amount of potential government transfer. As is often the case with the IV method, the standard errors are larger than those in OLS, and thus we cannot reject the null hypothesis that the impact of stroke on earnings is significantly different from zero for both hourly earnings and weekly earnings.

4 Discussion

Using a differences-in-differences method, a fixed-effects estimator, and Heckman's two-step method, this paper finds workers' health is deteriorated and their hours worked are reduced while their weekly or monthly salary rates do not decrease following a stroke. I investigate possible explanations for this observation and show that the data favors the hypothesis that earnings are inelastic with respect to productivity shocks caused by a stroke. That is, my hypothesis is that an increase in hourly wage after a stroke is due to inertia in payment schemes or employment contracts that are effective for a fixed-period. Owing to the presence of employment contracts or an equivalent arrangement, past wage payments might not respond to a sudden productivity reduction due to health shocks.

Table 12: Labor Earnings and the Incidence of Stroke (The second-stage IV Results)

Dep Var	Hourly Wage		Weekly Wage		Hourly Wage		Weekly Wage	
Sample	All				Male			
Method	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stroke	-0.05*	-0.04	-0.04	-0.03	-0.08**	-0.06	-0.07*	-0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)
Age	0.04***	0.04***	0.04***	0.05***	0.05***	0.05***	0.05***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age^2	-0.0003***	-0.0003***	-0.0004***	-0.0004***	-0.0005***	-0.0004***	-0.0005***	-0.0005***
	(5.6e-05)	(6.9e-05)	(5.7e-05)	(7.1e-05)	(9.7e-05)	(0.0001)	(9.9e-05)	(0.0001)
Edu Yrs	-0.02***	-0.07***	-0.03***	-0.08***	-0.01	-0.07***	-0.01	-0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Edu^2	0.004***	0.006***	0.004***	0.006***	0.003***	0.01***	0.004***	0.01***
	(0.0002)	(0.0004)	(0.0002)	(0.0004)	(0.0003)	(0.00)	(0.0003)	(0.00)
Female	-0.22***	-0.25***	-0.28***	-0.31***	–	–	–	–
	(0.01)	(0.01)	(0.01)	(0.01)	–	–	–	–
Married	0.03***	-0.01	0.04***	-0.00	0.10***	0.01	0.12***	0.03*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
No. obs.	56,468	125,298	56,761	125,298	26,521	54,920	26,692	54,920

Notes: The instrument is the amount of potential government transfer calculated based on the observable characteristics of workers following the government payment scheme. In addition to the instrument (potential amount of government transfer), age and household asset variables are included in estimating the first-stage selection model. The control variables include occupation dummies and industry dummies, but are not reported. Wages are wages conditional on earning positive amount. Wages for workers who do not work (recorded as zero in the data) would not enter our sample in OLS regression analysis or in the second stage of the IV regression analysis. The number of observations reported for the IV estimation is the total observations used for the estimation. Among 125 thousand in the total sample, 77 thousand are censored in the second stage of estimation. Among 55 thousand in the male sample, 32 thousand are censored.
Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, *$p<0.1$.

The finding of inertia in payment schemes highlights the fact that a fixed monthly salary or an annual employment contract is a prevalent form of payment scheme. While a fixed salary is well studied as a wage contract in theoretical papers (e.g., Holmström and Tirole (1993); Levin (2003)), many empirical papers ignore the presence of an employment contract and assume a perfectly competitive labor market in which wages are flexible and wage compensation thus equals worker productivity.

4.1 Who pays the cost of lower productivity due to health reduction?

Given that wage payment does not significantly decrease after productivity decline following a stroke, the analysis supports the hypothesis that employers pays lost worker production through higher costs of labor, not the worker through lower wages. As seen in the analysis of raw data, we observe that workers who continue working after a stroke surely suffer from reduction in their physical functionality and reduce the number of work hours. Nonetheless, wages and earnings remain at their pre-shock level despite reduced hours.

The hypothesis is further supported by the observation that the inert nature of wage payment is stronger for workers who stay with the same employers. As shown in the analysis by job tenure, workers with a longer job tenure are more likely to receive the same salary levels as before as stroke. In contrast, we observe the wage significantly decreased for workers who switch their jobs after a stroke.

The findings from this paper using the U.S. data indicate the importance of wage insurance, as in Guiso et al. (2005) in the Portuguese context and Cardoso and Portela (2005) in the Italian context. A caveat of my analysis is that the sample consists of workers aged older than 50. The compensation is found to be rigid maybe because these older workers are involved with tasks that cannot be replaced by other workers. Thus, we have to keep it in our mind that, if we conduct the same analysis for young workers, the results may differ.

4.2 Measurement Errors of Productivity

This results of this paper also have a broader implication about measurement errors of productivity. Since the existence of such rigidity in the employment contract translates to an increase in calculated hourly wages, this study also warns that wages, earnings, or salaries cannot be clearly interpreted as accurate values of the marginal product of labor. As a matter of fact, I find that the non-decreasing earnings along with the reduction in hours worked results in an increase in hourly wage rates constructed from reported hours and weekly earnings. When viewed in terms of neoclassical economics, this finding is a surprising data feature that appears as if productivity increased following a stroke because wage rates are often considered to reflect worker productivity.

In fact, many seminal papers drew their conclusions using hourly wage to represent productivity in studying the labor supply decisions of older individuals (e.g., French (2005); Low and Pistaferri (2014)). Thus, a reconsideration of this approach has important implications for the field of labor economics and future research.[14] When economists use hourly wages in their analyses of health and labor supply, it appears that they are really measuring productivity cost increases borne by the employer.

5 Concluding Remarks

Using the HRS data, I find that if workers stay with the same employers, hours worked are reduced while weekly or monthly salary rates do not decrease following a stroke. I investigate possible explanations for this observation and show that the data favors the hypothesis that earnings are inelastic with respect to productivity shocks caused by a stroke. The finding contrasts with an alternative hypothesis of self-selection, namely that high wages are driven by workers who remain in the labor market and actually become more productive or receive a positive wage/salary shock after having had a stroke. Thus, the findings of this paper indicate that employers bear much of the cost of the productivity shock since wages and earnings remain at their pre-stroke level despite reduced work hours. The limitation of this study is that the results only hold for older workers due to data availability. The findings of rigid employment contracts with respect to compensation may not be generalized to understand the labor market that most young workers face.

The findings in this paper convey two important messages. First, little change in labor earnings after a stroke casts doubt on the notion that wage determination mechanisms are as prevalent as neoclassical economists presume. The inert nature of wage payment supports alternative hypotheses; among possible explanations for inert wage schedule are sticky price theory, in which prices are resistant to change, and an implicit insurance mechanism that insures risk-averse workers against earnings

[14] Both French (2005) and Low and Pistaferri (2014) report that they computed wages as annual earnings divided by hours, using the PSID data in their empirical analysis.

shocks. More importantly, the implied higher wage per hour is in fact an increase in labor cost borne by the employer due to reduced productivity.

Second, since that existence of such rigidity in the employment contract translates to an increase in calculated hourly wage, this study also warns that wages, earnings, or salaries cannot be clearly interpreted as clear values of the marginal product of labor. Thus, this paper points out that wage variables naively created by dividing labor earnings by work hours could contain significant measurement errors for evaluating older workers' productivity. The results in this paper serve as a cautionary note for researchers who construct hourly wage data by using information from earnings and hours worked.

Appendix:

A Descriptive Data

Table 13 shows how people's health changes before and after a stroke for those in their 60s. Each of gender categories is divided into the treated and the control group, respectively. The treated and the control groups are further divided into age categories: 60 to 64 and 65 to 69. The treated age 60-64 column is individuals aged 60 to 64 who had not had a stroke at the time of observation. The treated age 65-69 column is individuals aged 65 to 69 who had had a stroke at the time of observation. The control group is the same age category for individuals who never had a stroke. Before columns show that magnitude of difficulty that individuals had performing these three tasks before their first stroke. After columns in contrast show the difficulty after their first stroke.

Table 13: Summary Statistics for the Treated and the Control Groups

Sample	Men 60 - 69					
Sample	Treated			Control		
Age	60-64	65-69		60-64	65-69	
Variable	before	after	diff.	no stroke	no stroke	diff.
Lift things	0.17	0.27	0.10	0.09	0.10	0.01
Climb up stairs	0.41	0.56	0.15	0.27	0.31	0.04
Pick up a dime	0.07	0.11	0.04	0.04	0.04	0.00
No. Obs.	1050	1571		20120	16941	

Sample	Women 60 - 69					
	Treated			Control		
Age	60-64	65-69		60-64	65-69	
Variable	before	after	diff.	no stroke	no stroke	diff.
Lift things	0.31	0.45	0.14	0.22	0.24	0.02
Climb up stairs	0.56	0.69	0.13	0.46	0.49	0.03
Pick up a dime	0.07	0.11	0.04	0.05	0.06	0.01
No. Obs.	1353	1450		21692	17144	

Source: the 1992-2008 HRS data.
Notes: The sample consists of those aged 60 to 69. Treated Group: Those who had the first stroke at age 60 to 69. Control Group: Those who never had a stroke till age 69. Statistics for age 60-64 in the treated group are from those aged 60 to 64, had not had a stroke at the time of the survey, but had (would have) their first stroke at age between 60 and 69. Statistics for age 65-69 in Treated group are from those for aged 65 to 69 who already had their first stroke between 60 and the time of survey age. Statistics for the Control group in the corresponding ages are calculated from those who did not have (would not have) their first stroke at age between 60 and 69. LFP stands for Labor Force Participation. No. Observations is the total number of observations in each category group. Some of the variables (e.g. earnings) are missing.

In their late 60s, 10% of the stroke survivors work full-time while 19% of those without stroke work full-time. Thus, a stroke greatly affects labor supply, but not always to the extent that individuals cannot work completely.

Table 14: Summary Statistics for the Treated and the Control Groups

Sample	Men 60 - 69					
	Treated			Control		
Age	60-64	65-69		60-64	65-69	
Variable	before	after	diff.	no stroke	no stroke	diff.
LFP	0.45	0.13	- 0.32	0.51	0.22	- 0.29
(full-time)	0.38	0.10	- 0.28	0.45	0.19	- 0.24
(part-time)	0.06	0.04	- 0.02	0.05	0.04	- 0.01
(unemployed)	0.01	0.004	- 0.006	0.01	0.003	- 0.007
Job Tenure (yrs)	13.7	9.6	- 4.1	14.7	11.9	- 2.8
Hours (full-time)	40.2	32.1	- 8.1	40.6	33.3	- 7.3
Wkly Wage (full)	748	829	81	950	809	- 141
Hrly Wage (full)	22.8	31.5	8.3	38.7	37.0	- 1.7
No. Obs.	1050	1571		20120	16941	

Sample	Women 60 - 69					
	Treated			Control		
Age	60-64	65-69		60-64	65-69	
Variable	before	after	diff.	no stroke	no stroke	diff.
LFP	0.31	0.08	- 0.23	0.35	0.14	- 0.21
(full-time)	0.21	0.06	- 0.15	0.25	0.09	- 0.16
(part-time)	0.09	0.01	- 0.08	0.13	0.05	- 0.08
(unemployed)	0.01	0.001	- 0.009	0.01	0.002	- 0.008
Job Tenure (yrs)	11.6	9.0	- 2.6	11.6	11.0	- 0.6
Hours (full-time)	33.4	27.0	- 6.4	33.2	27.5	- 5.7
Wkly Wage (full)	383	456	73	523	386	- 37
Hrly Wage (full)	11.5	13.1	1.6	16.4	15.1	- 1.3
No. Obs.	1353	1450		21692	17144	

Source: the 1992-2008 HRS data.
Notes: The sample consists of those aged 60 to 69. Treated Group: Those who had the first stroke at age 60 to 69; Control Group: Those who never had a stroke until age 69. Statistics for age 60-64 in Treated group are from those who are aged 60 to 64, had not had a stroke at the time of the survey but had (would have) their first stroke at age between 60 and 69. Statistics for age 65-69 in Treated group are from those for aged 65 to 69 who already had their first stroke between 60 and the time of survey age. Statistics for the Control group in the corresponding ages are calculated from those who did not have (would not have) their first stroke at age between 60 and 69. LFP stands for Labor Force Participation. No. Observations is the total number of observations in each category group. Some of the variables (e.g. earnings) are missing.

B Other Estimation Results

Table 15 reports the regression results that investigate endogeneity issues. The analysis based on these results is discussed in Section 3.3. The sample consists of those who had a stroke and continue to work full-time and those who did not have a stroke and continue working full-time. All wages are measured by logarithm of wage. Additional controls include marital status, gender dummy, race dummy, education years, and years of job tenure. For the analysis using the overall sample, I also include the interaction term of gender and marital status to capture the gender-specific role of marriage that could affect wages. Table 16 presents the regression results that examine endogeniety issue using the male sample only. The sample restriction is the same as the one in Table 15.

Table 17 reports the correlation between the incidence of each disease and self-reported stress levels in the previous period. The results show that the propensity of having a stroke is not significantly associated with the stress level in the previous period.

C Heckman Two-step Method

The implementation of estimation here follows Heckman (1979) and can be explained in two steps: the selection process in the first stage and the earning relationship in the second stage. First, I describe the selection stage. I assume that the determination process of labor supply is approximated by the following reduced-form model:

$$U_{it}^* = \mathbf{X}_{it}'\beta_x + \beta_z Z_{it} + \epsilon_{it} \tag{5}$$

where U_{it}^* is the latent variable that represents the utility from working. Worker i works if $U_{it}^* > 0$. The last term ϵ_{it} captures taste for work, which is unobserved by researchers. The vector Z_{it} includes exclusion restrictions. They affect the likelihood of participating in the labor force, but do not directly affect the wage, conditional on $\mathbf{X}_{it}$.

In the first stage, a model of labor force participation (Equation 5) is estimated using Probit regression. After estimating (5), the inverse Mills ratio is calculated as $\lambda(\mathbf{X}, Z) = \phi(\mathbf{X}, Z)/\Phi(\mathbf{X}, Z)$, where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the p.d.f. and c.d.f. of the standard normal distribution, respectively. Using the inverse Mills Ratio, I estimate the wage equation as

$$\ln w_{it} = \mathbf{W_{it}'}\alpha + \phi 1\{stroke_{it} = 1\} + \sigma\lambda(\mathbf{X}, Z) + \upsilon_{it}$$

where $\mathbf{W}$ is a vector of wage determinants, *stroke* is an indicator of having had a stroke, σ is the correlation between unobserved determinants of propensity to work and unobserved determinants of wage offers υ_{it}, using the sample of workers who work at time t. The coefficient of my interest is ϕ, which captures the average treatment effect of surviving stroke on labor earnings.

As in Low and Pistaferri (2014) and the empirical public finance literature, I use potential government transfers such as food stamps and government transfers including EITC as the exclusion restrictions for the estimation of the wage parameters.[15] Here, I briefly discuss the two main conditions that must be met for the IV method to be effective in identification in our context: exclusion

[15] The EITC is a refundable income tax credit for low-income working individuals and families. Among the determinants and eligibility of EITC are employment history, income, family size, and the number of children.

Table 15: Labor Earnings and the Incidence of Stroke (Differences-in-Differences Method)

Sample	All					
Dep. Var.	Hourly Wage			Weekly Wage		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Main Specification	Main Specification with risk factors	Include years prior to a stroke	Main Specification	Main Specification with risk factors	Include years prior to a stroke
Stroke (δ_0)	0.12***	0.12***	0.04	0.12***	0.13***	0.04
(0-1 yrs)	(0.04)	(0.05)	(0.08)	(0.04)	(0.05)	(0.07)
Post-treat (δ_1)	0.18***	0.20***	0.11	0.20***	0.24***	0.14*
(2-3 yrs)	(0.04)	(0.05)	(0.08)	(0.04)	(0.05)	(0.07)
Post-treat (δ_2)	0.12**	0.11**	0.02	0.15***	0.14**	0.04
(4-5 yrs)	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.07)
Post-treat (δ_3)	0.22***	0.19***	0.11	0.24***	0.21***	0.12
(6-7 yrs)	(0.07)	(0.07)	(0.09)	(0.07)	(0.07)	(0.08)
Treat (γ)	-0.23***	-0.23***	-0.15**	-0.23***	-0.23***	-0.13**
	(0.03)	(0.04)	(0.07)	(0.03)	(0.04)	(0.07)
0-1 yrs before			-0.07			-0.08
			(0.09)			(0.08)
2-3 yrs before			-0.07			-0.09
			(0.08)			(0.08)
Female	-0.25***	-0.24***	-0.24***	-0.32***	-0.31***	-0.31***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
High Blood Pressure		0.03***	0.03***		0.02***	0.02***
		(0.01)	(0.01)		(0.01)	(0.01)
Diabetes		-0.02	-0.02		-0.03**	-0.03**
		(0.01)	(0.01)		(0.01)	(0.01)
Lung Disease		-0.08***	-0.08***		-0.08***	-0.08***
		(0.02)	(0.02)		(0.02)	(0.02)
Heart Disease		-0.03**	-0.03**		-0.03**	-0.03**
		(0.01)	(0.01)		(0.01)	(0.01)
Arthritis		-0.01	-0.01		-0.01	-0.01
		(0.01)	(0.01)		(0.01)	(0.01)
Age dummy	Yes	Yes	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
No. Obs.	60,469	54,377	54,377	60,478	54,386	54,386
R^2	0.28	0.28	0.28	0.30	0.30	0.30

Standard errors, in parentheses, and clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1

Table 16: Labor Earnings and the Incidence of Stroke (Differences-in-Differences Method)

Sample	Male					
Dep. Var.	Hourly Wage			Weekly Wage		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Main Specification	with risk factors	Include years prior to a stroke	Main Specification	with risk factors	Include years prior to a stroke
Stroke (δ_0)	0.12*	0.13*	0.11	0.12*	0.14**	0.09
(0-1 yrs)	(0.06)	(0.07)	(0.11)	(0.06)	(0.07)	(0.10)
Post-treat (δ_1)	0.17***	0.20***	0.18*	0.20***	0.24***	0.20**
(2-3 yrs)	(0.06)	(0.07)	(0.10)	(0.05)	(0.07)	(0.10)
Post-treat (δ_2)	0.07	0.08	0.06	0.09	0.11	0.06
(4-5 yrs)	(0.07)	(0.07)	(0.10)	(0.07)	(0.08)	(0.10)
Post-treat (δ_3)	0.17*	0.17	0.15	0.20*	0.20*	0.15
(6-7 yrs)	(0.10)	(0.11)	(0.13)	(0.10)	(0.10)	(0.12)
Treat (γ)	-0.22***	-0.24***	-0.23**	-0.22***	-0.23***	-0.19**
	(0.05)	(0.05)	(0.10)	(0.05)	(0.05)	(0.09)
0-1 yrs before			0.02			-0.02
			(0.12)			(0.11)
2-3 yrs before			0.02			-0.04
			(0.12)			(0.11)
Female	–	–	–	–	–	–
	–	–	–	–	–	–
High Blood Pressure		0.04***	0.04***		0.03**	0.03**
		(0.01)	(0.01)		(0.01)	(0.01)
Diabetes		-0.00	-0.00		-0.01	-0.01
		(0.02)	(0.02)		(0.02)	(0.02)
Lung Disease		-0.09***	-0.09***		-0.08***	-0.08***
		(0.02)	(0.02)		(0.02)	(0.02)
Heart Disease		-0.02	-0.02		-0.02	-0.02
		(0.02)	(0.02)		(0.02)	(0.02)
Arthritis		-0.01	-0.01		-0.02	-0.02
		(0.01)	(0.01)		(0.01)	(0.01)
Age dummy	Yes	Yes	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
No. Obs.	31,919	28,618	28,618	31,927	28,626	28,626
R^2	0.21	0.21	0.21	0.22	0.22	0.22

Standard errors, in parentheses, and clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1

restriction and non-weak correlation with labor force participation. Following the literature, I argue that the first condition is satisfied by showing that our instrument (potential government transfer) is non-weakly correlated with the effect of our interest (labor force participation). To support non-weak correlation, Table 18 presents the F-statistics that show how strong the correlation between the amount of potential transfers and labor force participation probability is. Throughout the exercise, standard errors are clustered at the individual level. All the F-statistics in the first stage of the IV exceed the Stock and Yogo (2005) 10% critical value. The coefficient of the interaction term shows that workers are 2–3% less likely to participate in the labor force when the amount of potential transfers is higher by 1%.

Second, I argue that the second condition necessary for a valid IV method, the exclusion restriction, is thought to be plausible in this study. Generally speaking, potential government transfers are considered as exogenous and can serve as a "simulated IV." Indeed, "realized" public income transfers are endogenous given that the individuals take-up decision is a choice. However, the amount of benefits individuals are potentially eligible for (before any take up decision is made) can be considered as exogenous since payment schemes in public programs are determined outside the worker's choice model. Thus, I assume that the key conditions for the IV method hold, and apply the IV method to estimate how the incidence of stroke affects worker's earnings on average.

Table 17: Job Stress and Physical Health

Dep. Var.	Stroke	Psych	High blood	Diabetes	Cancer	Lung disease	Heart prob.	Arthritis
Age	-0.002***	0.017***	0.030***	0.006***	-0.011***	0.005***	-0.008***	0.036***
	(0.0007)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Age^2	0.0000***	-0.0001***	-0.0001***	0.0000	0.0001***	-0.0000	0.0001***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Hours	-0.0002***	-0.0009***	-0.0006***	-0.0002***	-0.0004***	-0.0003***	-0.0004***	-0.0011***
	(0.0000)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Job stress	0.0007	0.016***	0.008***	0.003**	0.002	0.003***	0.004**	0.013***
	(0.0006)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Constant	0.0484**	-0.4082***	-0.9764***	-0.2884***	0.2203***	-0.1493***	0.0752	-1.1926***
	(0.022)	(0.082)	(0.067)	(0.045)	(0.037)	(0.035)	(0.050)	(0.071)
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	61,885	61,480	59,664	61,436	61,657	61,600	61,928	59,750
# ID	16,226	16,218	16,180	16,209	16,221	16,217	16,224	16,171

Notes: Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 18: The First Stage Parameter Estimates

Dependent Variable	LFP	LFP
	(1)	(2)
Log (Potential Gov. Transfer)	-0.02***	-0.02***
	(0.003)	(0.003)
Edu Yrs	0.01***	0.01***
	(0.002)	(0.002)
$(\text{Edu Yrs})^2$	0.0003***	0.0001
	(9.25e-05)	(9.30e-05)
Female	-0.004	-0.001
	(0.005)	(0.005)
Married	0.05***	0.04***
	(0.0048)	(0.0048)
Married X Female	-0.12***	-0.12***
	(0.0058)	(0.0058)
Log (household asset)	0.02***	0.01***
	(0.0015)	(0.0015)
Additional Control Variables	No	Yes
Observations	122,287	114,573
R-squared	0.32	0.36
F statistics for weak identification	3423.84	2501.11

Sample: Respondents in HRS 1992–2008.
Note) Standard errors in parentheses. *** p<0.01, ** p<0.05, *p<0.1.
Additional Control Variables include health conditions and the number of small children.

References

C. Azariadis. Implicit contracts and underemployment equilibria. *Journal of Political Economy*, 83 (6):1183–1202, 1975.

C. Azariadis. Human capital and self-enforcing contracts. *The Scandinavian Journal of Economics*, 90(4):507–528, 1988.

M. Baily. Wages and employment under uncertain demand. *The Review of Economic Studies*, 51, 1974.

N. Baum-Snow and D. Neal. Mismeasurement of usual hours worked in the census and acs. *Economics Letters*, 102(1):39 – 41, 2009.

D. Bernhardt and G. C. Timmis. Multiperiod wage contracts and productivity profiles. *Journal of Labor Economics*, 8(4):529–563, 1990.

J. Bound, C. Brown, and N. Mathiowetz. Chapter 59 - measurement error in survey data. volume 5 of *Handbook of Econometrics*, pages 3705 – 3843. Elsevier, 2001.

K. Burdett and M. Coles. Equilibrium wage-tenure contracts. *Econometrica*, 71(5):1377–1404, 2003.

P. Böckerman and E. Laukkanen. What makes you work while you are sick? evidence from a survey of workers. *The European Journal of Public Health*, 2010.

A. R. Cardoso and M. Portela. The provision of wage insurance by the firm: Evidence from a longitudinal matched employer-employee dataset. Institute for the Study of Labor Discussion Paper No. 1865. Available at SSRN: http://ssrn.com/abstract=868992, 2005.

L. Danziger. Real shocks, efficient risk sharing, and the duration of labor contracts. *The Quarterly Journal of Economics*, 103(2):435–440, 1988.

E. French. The effects of health, wealth, and wages on labour supply and retirement behaviour. *Review of Economic Studies*, 72(2):395–427, 2005.

D. F. Gordon. A neo-classical theory of keynesian unemployment. *Economic Inquiry*, 12(4):431–59, 1974.

L. Guiso, L. Pistaferri, and F. Schivardi. Insurance within the firm. *Journal of Political Economy*, 113 (5):1054–1087, 2005.

L. Guiso, L. Pistaferri, and F. Schivardi. Credit within the firm. *Review of Economic Studies*, 80:211 – 247, 2013.

J. J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47(1):153–161, 1979.

B. Holmström and J. Tirole. Market liquidity and performance monitoring. *Journal of Political Economy*, 101(4):pp. 678–709, 1993.

G. Kátay. Do firms provide wage insurance against shocks? evidence from hungary. 2008. Magyar Nemzeti Bank.

W. Koolhaas, J. J. van der Klink, J. W. Groothoff, and S. Brouwer. Towards a sustainable healthy working life: associations between chronological age, functional age and work outcomes. *The European Journal of Public Health*, 22(3):424–429, 2012.

J. Levin. Relational incentive contracts. *The American Economic Review*, 93(3):pp. 835–857, 2003.

H. Low and L. Pistaferri. Disability insurance and the dynamics of the incentive-insurance tradeoff. *Working Paper*, 2014.

C. Michelacci and V. Quadrini. Borrowing from employees: Wage dynamics with financial constraints. *Journal of the European Economic Association P&P*, 3, 2005.

C. Michelacci and V. Quadrini. Financial markets and wages. *Review of Economic Studies*, 76, 2009.

J. Stock and M. Yogo. *Testing for Weak Instruments in Linear IV Regression*, pages 80–108. Cambridge University Press, New York, 2005.

P. Susser and N. R. Ziebarth. Profiling the u.s. sick leave landscape: Presenteeism among females. *Health Services Research*, pages 1–3, 2016.

R. Weaver. Cost of presenteeism surpasses absenteeism. Retrieved from http://www.examiner.com/human-capital-in-detroit/cost-of-presenteeism-surpasses-absenteeism, 2010.