Understanding the Aggregate Effects of Disability Insurance: Safety Net for the Unhealthy and Its Consequences on the Healthy*

Soojin Kim† Serena Rhee‡

March 14, 2018

Abstract

This paper studies the output and productivity effects of disability insurance: the direct effect from foregone labor; and the indirect effect from the potential complementarity between those who exit the labor force and the existing workforce. To do so, we estimate the impact of disability on a worker's human capital–(pure) labor and experience–endowments. We find that the amount of experience increases over the life-cycle relative to the value of labor. Moreover, disability has a lower impact on a worker's experience than it does on his labor. These results, coupled with high share of older DI recipients imply a significance loss in the amount of experience supplied in the labor market. Secondly, we estimate and find that at the aggregate level, labor and experience are complementary inputs. Thus, DI has pecuniary externality, through changes in the relative price of experience. Using a quantitative general equilibrium model, we find that DI reduces relative supply of experience by 8.9%, lowering the output by 3%.

JEL Codes: J31, J24, E24, J11
Keywords: disability, wage risk, skill complementarity

*For valuable comments and discussion, we would like to thank seminar and conference participants at numerous institutions. We are grateful to Iourii Manovskii for helping us generating labor market experience variables using Panel Study of Income Dynamics data. This paper was written in part while Rhee was visiting the RAND Corporation and USC Schaeffer Center for Health Policy and Economics. Rhee is grateful to them for their hospitality during her visits. Rhee thanks the National Institutes of Health under award number 2T32AG000244 for support. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the National Institutes of Health.

†Purdue University, 403 W. State Street, West Lafayette, Indiana 47907. soojink@purdue.edu.

‡University of Hawaii at Manoa, 2424 Maile Way, Saunders Hall Room 542, Honolulu, Hawaii 96822. rhees@hawaii.edu.

1 Introduction

How substantial are the output and productivity effects of disability insurance (DI)? With the recent increase in DI beneficiaries, fiscal and labor supply effects of the policy have drawn great interests among economists, and their significance and size are well-established in the literature.[1] On the other hand, studies of its aggregate effects on output and productivity are limited.[2] This paper aims at analyzing the aggregate effects of DI that consists of first, the loss from lower labor supply, the direct effect; and second, the productivity impact driven by the imperfect substitutability of DI recipients (the exiters from the labor market) and workers in the labor market (the stayers), the indirect effect.

In order to measure the direct productivity effects from foregone labor, we need to understand how disability affects a worker's productivity. While many agree that disability adversely affects a worker's productivity (wage), it is more difficult to understand its role on various aspects of his human capital. For example, while health shocks may have significant impact on effective labor a workers provides, it may have little impact on his accumulated experience. Thus, we estimate wage (productivity) process with assumption that workers are endowed with (pure) labor and experience. This generalized framework allows us to capture potential heterogeneous impacts of disability on a worker's productivity. Second, to measure how the withdrawal of DI recipients from the labor market affects other workers (the spillover or indirect effect), we estimate the imperfect substitutability between labor and experience at the aggregate level. Then, we build a general equilibrium life-cycle model, in which households face health and income risks and make endogenous labor supply decisions. Lastly, we calibrate the model to match key U.S. empirical facts, and use it to quantify the labor supply, output, and productivity effects of the current DI system.

Our empirical analysis for understanding the individual-level productivity follows and extends that of Jeong, Kim and Manovskii (2015). We use Panel Study of Income Dynamics (PSID), which contains work history (years of experience) and health status data, to estimate the productivity of workers. Workers' endowments–(effectiveness) of labor and non-labor ("experience") inputs–vary over the life-cycle, and health statuses of workers can have differential impacts on these inputs.[3] Our findings suggest that (i) the share of experience in a worker's total amount of human capital (labor and experience) increases over the life-cycle; (ii) disability lowers a worker's effective labor by 30% among high school graduates (44% for college-

[1] In 2015, total benefit payments for Social Security Disability Insurance and Medicare for qualified beneficiaries exceeded \$220 billion (5.8% of the federal budget). With the aging of population, the Congressional Budget Office projects that the DI trust fund will be exhausted in 2022 (Congressional Budget Office, 2016). Maestas, Mullen and Strand (2013) and French and Song (2014) are among the few recent studies estimating the labor supply disincentive effects of DI.

[2] Kitao (2014) focuses on the employment effects of DI, but abstracts from fully analyzing the output and productivity effects.

[3] We control for various individual characteristics in estimating the effective endowments of human capital, and allow for education-dependence on the effects of health. Moreover, we use Heckman's two-step procedure to control for selection.

educated); and (iii) the impact on experience of a disability is smaller at 26% for high school graduates (8% for college-educated). Since 75% of DI recipients are older than 50, whose human capital is primarily in experience, the amount of experience we lose from DI is quite large. Moreover, since the DI program disproportionately affects experience compared to labor at the aggregate level, this change can cause indirect effects through the change in factor prices.

To evaluate the macro effects of the DI program, we construct the aggregate supply of labor and experience using our micro-level estimates. We assume that labor and experience are inputs with constant elasticity of substitution (CES) in aggregate production. Our framework follows Jeong, Kim and Manovskii (2015), and is a generalized version of previous studies.[4] Instead of taking age as an index for human capital, we explicitly establish the linkage between experience and wage, because health affects the accumulation of experience and the value from the accumulated experience. Our estimation, which uses the time-series variation in the relative supply of experience and labor, finds that two inputs are complementary with the elasticity of substitution of -2.1. As older workers are more experienced than the young, we find that the composition effects are especially pertinent with the current DI program.

Next, we build a life-cycle general equilibrium model to conduct counterfactual analyses. In it, finitely-lived households are subject to health (disability and mortality), medical expenditure, and productivity shocks, and make endogenous labor supply decisions until they reach a mandatory retirement age. We also model the current DI program in the U.S. and other government policies that interact with DI in household decisions, such as Social Security and Medicare. We calibrate the model to match the key aspects of labor market outcomes by age and health.

Using the estimated wage processes and the calibrated structural model, we evaluate the impact of introducing DI on individual's decision to work, aggregate price of labor and experience, and the aggregate output. In our calibrated economy, a generous DI program reduces work incentives of young and old workers, thereby lowering the aggregate supply of both labor and experience. However, the decrease is not proportional, as the experience-abundant old are more likely to withdraw from the labor force. Moreover, even though the majority of these old workers are unhealthy, as health shocks have less detrimental effects on experience (compared to labor), they are still productive, especially in the amount of experience they supply. In the aggregate, an introduction of DI, therefore, increases the relative price of experience by 8.5%. This equilibrium affects all workers in the labor market, including the young and the healthy. In the economy with DI, the employment rates of disabled and non-disabled workers decrease by 5.5 and 2.4 percentage

[4] There have been studies that consider different generations as heterogeneous inputs in labor market (e.g. Card and Lemieux, 2001; Gruber and Milligan, 2010; and Munnell and Wu, 2012) and found imperfect substitutability between them.

points, respectively. Overall, output decreases by 3% and aggregate efficiency, (as defined by output divided by composite labor) increases by 1%.

Related Literature This paper is related to several strands of literature studying (i) the role of heterogeneous inputs in production and their interactions in the labor market; (ii) the disincentive effects of DI on the labor supply; and (iii) the effects of social insurance policies in structural models with heterogeneous agents.

First, we build on the literature that studies heterogeneous inputs in production. Empirically, a few papers study the relationship between young and old workers in the labor market. In Gruber and Milligan (2010), unemployment rates of youth (those aged between 20 and 24) and prime-aged (between 25 and 54) workers drop when elderly (between 55 and 64) employment rates increases for males, and the effects are statistically significant for the prime-aged workers. Munnell and Wu (2012) uses a state-year-age mortality rate as an instrumental variable to find that there is no evidence of crowd-out effect of young workers when the employment of old workers increases. Instead, they find that the increased employment rate of the old leads to an increased employment rate of the prime-aged workers.

These are empirical evidence of complementary between young and old workers, consistent with our finding that labor and experience (thus implicitly young and old) are complementary in production. Relatedly, other papers in the literature estimate the degree of substitutability across heterogeneous inputs in production using empirical data, mostly assuming a Constant Elasticity of Substitution production function (e.g., Card and Lemieux, 2001; Krusell et al., 2000; and Karabarbounis and Neiman, 2014).[5]

In terms of methodology, we are most closely related to Jeong, Kim and Manovskii (2015), which estimates the efficiency of labor and experience, two distinct inputs (human capital) a worker is endowed with, using work experience data along with individual-level characteristics from the PSID. We expand their wage process to allow for the health impact on both labor and experience endowment of workers over the life-cycle. There are two contributions from our analysis in this regard. First, we use this new approach to understand the role of disability on productivity (income). Most papers with health- or disability-dependent productivities (e.g., Low and Pistaferri, 2015) assume a one-dimensional impact of health on productivity (wage), while we allow for two channels (labor and experience) in which health can affect market income. Importantly, we find differential impact of health on labor and experience, and thus our paper can be a step towards understanding and decomposing the impact of health on worker inputs. Secondly, we use the

[5] Card and Lemieux (2001) uses a CES production function with labor inputs from different skill and age to explain college premium; Krusell et al. (2000) shows that capital-skill complementarity can explain the rise of the skilled labor and the skill premium; and Karabarbounis and Neiman (2014) estimates the elasticity of substitution between Information Technology (IT) and labor to explain the decline of the labor share.

estimated wage (productivity) process to conduct policy analysis, and find that capturing the heterogeneity across health and age is important for evaluating social insurance policies.

Secondly, the paper builds on and expands the studies of labor supply disincentive effects of DI, which has long been a topic of interest starting with Bound (1989). Recently, Maestas, Mullen and Strand (2013) and French and Song (2014) use random assignment of disability examiners and judges to estimate the disincentive effects of disability insurance on labor supply of workers. Both papers find a strong disincentive effect of disability insurance. Maestas, Mullen and Strand (2013) finds that the SSDI causes a 28 percentage point (pp) decline in employment rate. For those between the ages of 40-49 and 50-59, the impact is larger at 35.5pp and 29pp, respectively. Given that these age groups comprise 28.4 and 34.9% of the total applicant pool in their data, we see that a large fraction of workers affected by the SSDI are in these age groups. French and Song (2014) also finds a 25.4pp effect among those aged 45-54, and 24.8pp among 55-59 year olds. While these papers use econometric approach to study individual behavior, Kitao (2014) and Low and Pistaferri (2015) are among the few that develop a life-cycle model to analyze the effects of DI. Kitao (2014) focuses on the interaction between DI and unemployment insurance, as well as other policies, including Social Security and Medicare. By endogenizing the enrollment decision to various social programs and employment decision of households, she quantitatively shows the aggregate effects of policy reforms. On the other hand, Low and Pistaferri (2015) focuses on the incentive and insurance trade-off that DI creates, but abstracts from aggregate effects of DI, which is a central focus of this paper.

Finally, this paper also contributes to the literature analyzing the effects of social insurance policies especially with respect to health or medical expense risks (e.g., Hubbard, Skinner and Zeldes, 1995; and Attanasio, Kitao and Violante, 2011) . Some of the recent papers in the literature include De Nardi, French and Jones (2016) and Braun, Kopecky and Koreshkova (2017) both of which analyze the role of social insurance policies for the old (e.g., Medicaid and for the latter both Medicaid and Supplemental Security Income). Using a richly calibrated model, they measure the welfare gains from these means-tested social insurance programs in the presence of health and medical expenditure risks. We, on the other hand, study the role of DI, and aim at measuring the efficiency loss in a general equilibrium model framework.

The organization of the paper is as follows. Section 2 outlines our empirical estimation of the productivity of workers with different health statuses, and the elasticity of substitution between labor and experience. Section 3 develops a general equilibrium model with DI, which serves as a laboratory for evaluating the effects of DI. In Section 4, we use the calibrated model to conduct counterfactual analyses. We conclude in Section 5.

2 Empirical Analysis: The Health-Specific Value of Labor and Experience

In this section, we use data from the Panel Study of Income Dynamics (PSID) to first estimate the health-specific value of (pure) labor and experience, using a generalized framework of Jeong, Kim and Manovskii (2015) which decomposes wage into return from a worker's (pure) labor and that from his experience. Using the micro-level estimates to construct the aggregate supply of labor and experience, and its time-series variation, we estimate the elasticity of substitution between the two inputs.

The hourly wage rate (or productivity) of an individual consists of returns from both labor and experience. Labor represents deterministic age- and health-dependent physical ability, while experience can potentially grow over time through labor market participation. We can interpret the experience as a proxy for human capital acquired on the job.

An individual's labor (λ_L) is deterministic with respect to age j and health status h, while the amount of experience is a product of both deterministic component (λ_E) and his past work history ($g(e)$). Denoting the factor price of labor and experience at time t as $R_{L,t}$ and $R_{E,t}$, the hourly wage rate of a worker with age j, health status h is

$$w_{j,t}(h,e) = R_{L,t}\lambda_L(j,h) + R_{E,t}\lambda_E(j,h)\,g(e).$$

For simplicity, the deterministic components of labor and experience are represented by polynomial functions of age j:

$$\tilde{\lambda}_X(j) = \exp\left(\lambda_{X,0}(j) + \lambda_{X,1}(j)\,j + \lambda_{X,2}(j)\,j^2\right),$$

with $X = L$ and E. Each coefficient can vary depending on individual characteristics (x).[6] Under the functional form assumptions, the relative efficiency of experience compared to labor is given as

$$\frac{\tilde{\lambda}_E(j)}{\tilde{\lambda}_L(j)} = \exp\left(\lambda_{E/L,0}(j) + \lambda_{E/L,1}(j)\,j + \lambda_{E/L,2}(j)\,j^2\right),$$

where $\lambda_{E/L,k}(j) \equiv \tilde{\lambda}_{E,k}(j) - \tilde{\lambda}_{L,k}(j)$. Moreover, the role of accumulated experience on effective experience is captured by $g(e) = \sum_{n=1}^{4} \theta_n e^n$ with $\theta_1 \equiv 1$.

Health proportionately affects these labor and experience profiles by a factor ϕ_x^h so that $\lambda_X(j,h) = \phi_X(h)\,\tilde{\lambda}_X(j)$ for $X = L, E$. Without loss of generality, we normalize the profile for the healthy individuals

[6] In our benchmark analysis, we allow the profile to differ by gender and education. For the sake of brevity, we suppress these notations.

to $\phi_X\left(h^H\right) = 1$. Under these model specifications, the log-wage equation can be expressed as

$$
\begin{aligned}
& \ln w_{it}(j,h) \qquad (1)\\
= \; & \ln R_{L_t} + \ln \phi_L^h(x_{it}) + \ln \tilde{\lambda}_L(j, x_{it}) + \ln\left(1 + \Pi_{E_t}\phi_{E/L}^h(x_{it})\lambda_{E/L}(j, x_{it}) g(e_{it})\right) + \ln z_{it} \\
= \; & \ln R_{L_t} + \ln \phi_L^h(x_{it}) + \left\{\lambda_{L,0}(x_{it}) + \lambda_{L,1}(x_{it}) j + \lambda_{L,2}(x_{it}) j^2\right\} + \ln z_{it} \\
& + \ln\left[1 + \Pi_{E_t}\frac{\phi_E^h(x_{it})}{\phi_L^h(x_{it})}\exp\left(\lambda_{E/L,0}(x_{it}) + \lambda_{E/L,1}(x_{it}) j + \lambda_{E/L,2}(x_{it}) j^2\right) g(e_{it})\right],
\end{aligned}
$$

where $\Pi_{E_t} = \frac{R_{E_t}}{R_{L_t}}$ is the relative price of experience. The regressor z_{it} includes time-specific dummies for education, gender, region, and race. In estimation, we augment the equation (1) with standard measurement error term denoted by $\varepsilon \sim N\left(0, \sigma_\varepsilon^2\right)$.

2.1 Wage Offer Distribution: Selection Bias and Identification Strategy

One challenge of estimating the equation (1) is that we only observe wages of selected individuals; workers, especially those who are participating in the labor market despite their disabilities, might be systematically different from those who exit the labor force. Therefore, the estimated health effects on labor and experience efficiencies without correcting for this selection can be biased. Thus, we estimate the wage process using a standard two-stage procedure described in Heckman (1979):

$$
\ln w_{it}^{\text{offer}}(j,h) = \begin{cases} \ln w_{it}(j,h) + \varepsilon_i & \text{if } l > 0 \\ \text{not available} & \text{if } l = 0. \end{cases}
$$

Following Low and Pistaferri (2015), we address this selection bias by instrumenting the underlying labor force participation decision problem with the generosity of the local government's welfare programs. The scale of welfare programs is measured based on the potential government's transfer benefits on "representative" earners.[7] Our identification strategy relies on the assumption that (i) these potential benefits affect the labor market participation decision of individuals, but not their wage rates, and (ii) the effects of welfare program vary across health types. Further details of the estimation process can be found in Appendix A.2.

In the estimation, we define those who report to have a work limitation, as "disabled", and others as "non-disabled" workers. Table 1 presents the results from the first-stage probit regression. We observe that work limitation has a significant impact on employment probability, and it lowers the probability by 16.7 *pp* for a marginal worker and 13.5 *pp* for an average worker.

[7] Therefore, the size of welfare transfers used in our estimation is not the actual amount each individual received.

1

TABLE 1: FIRST-STAGE PROBIT REGRESSION RESULTS

Independent Variables	Coefficients	Effects on probability of employment	
		Marginal effects	Average effects
Work Limitations	-1.362^{***}	-0.167^{***}	-0.135^{***}
	(0.100)	(0.028)	(0.022)
Number of obs.		68, 332	
Pseudo R^2		0.274	

Note: Table 1 reports the first stage probit regression results of Heckman's two-stage estimation for selection correction. The dependent variable is the employment status of an individual. Independent variables include individual characteristics such as age, experience, years of schooling, marital status, states as well as time-varying year dummies, male dummies, and race dummies. We use state-level generosity of welfare program and tax credits as exclusion restriction. The complete list of estimated coefficients for these variables are reported in Appendix B. We uses individual-level survey weights for our analysis. Standard errors clustered at individual level and reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Tables 2 and 3, as well as Figure 1 summarize our findings. First, in Table 2, we present the effect of a work limitation on wage before and after controlling for selection. While using only the observed wage rates of labor market participants implies a 15 *pp* decrease in a worker's wage, after controlling for selection, we see that the effect is larger at 26 *pp*.

TABLE 2: THE EFFECTS OF DISABILITIES ON WAGE ESTIMATION WITH AND WITHOUT SELECTION CONTROL

Independent Variables	Coefficient	
	No Control	Selection Controlled
Work Limitations	-0.151^{***}	-0.258^{***}
	(0.018)	(0.033)
Inverse Mills Ratio		0.297^{***}
		(0.073)
Number of obs.	116, 068	56, 840
R^2	0.276	0.266

Note: Table 2 reports the second stage linear regression results of Heckman's two-stage correction, and contrasts its outcomes from the estimation without selection control. The dependent variable is the log wage of an employed individual. The right hand variables also include age, age square, experience, experience square, experience cube, dummy variables for gender, race, marital status and years. The estimation with control has extra term, the inverse Mills ratio, which is constructed from the first stage estimation. The complete list of estimated coefficients are reported in Appendix B. We uses individual-level survey weights for our analysis. Standard errors clustered at individual level and reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Using these estimates, we construct and compare the average wage of the employed workers, and the average wage offer from the estimated wage offer distribution for both disabled and non-disabled workers (see Table 3). While the non-disabled exhibits almost no difference in mean between the average wage and the average wage offer, those who experience health problems show discrepancy between potential wage

TABLE 3: ACCEPTED WAGE VS. WAGE OFFER DISTRIBUTIONS

Disability Status	Average wage of the employed	Average wage offer from estimation	Ratio of offer-to-accepted wages (%)
Non-Disabled	2.123 (0.771)	2.113 (0.387)	99.0
Disabled	1.979 (0.810)	1.819 (0.451)	91.9

Note: Table 3 reports the mean and standard variation of log hourly wage of the employed workers. Based on the PSID, the hourly wage variable is computed by dividing the total labor income variable with the total annual working hours. We use sample periods from 1984 to 2011, and converted wage variables into 2011 US dollars using the Consumer Price Index (CPI). Observations include both man and woman with any education level in working-age between 25 and 65, and considered being employed if their reported more than 500 annual working hours. Estimated wage offer statistics are constructed based on the coefficients from the wage equation with Heckman's two-stage linear regression. Both mean and standard deviations are weighted using the individual weights.

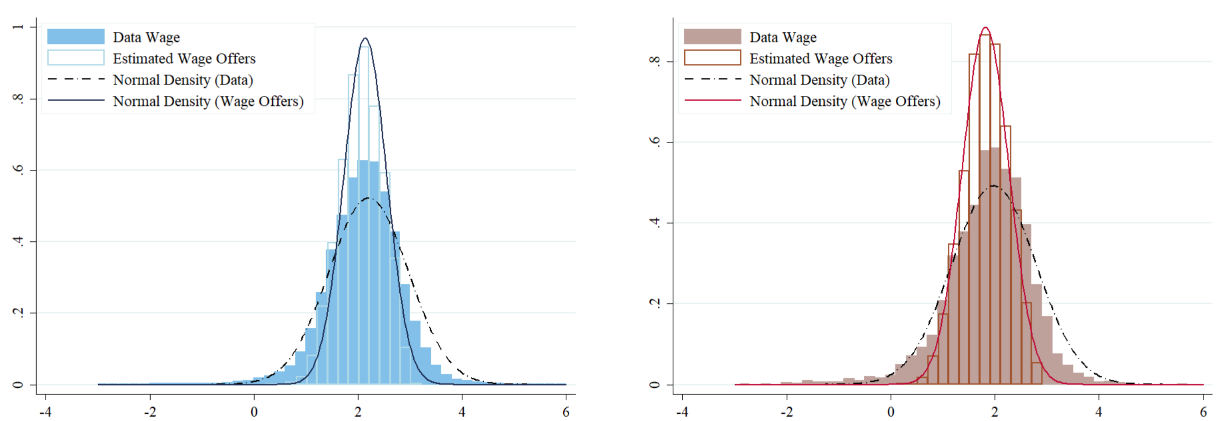
offers and actual wages, suggesting severe selection bias.

Figure 1 illustrates this selection bias graphically by comparing the log-wage distribution by health status. Once we control for the labor supply decision, the mean of the approximated distribution decreases compared to that of observed wages. Using the non-disabled workers' hourly wage rate as the baseline, individuals with work limitations exhibit the decline in their hourly wage by $\$1.64\ (= \exp(2.1) - \exp(1.8))$. Monotonicity in negative impact of health on wage is robust after controlling for gender and education status.

FIGURE 1: SELECTION BIAS: OBSERVED WAGE VS. ESTIMATED WAGE OFFERS

(A): NON-DISABLED

(B): WORK LIMITATIONS



Note: Figure 1 compares the log wage offer distribution with the data by health status, using the hourly wage of working-age individuals in age 25 to 65 in PSID from years 1984 to 2011. The hourly wage rate is defined as the total labor income divided by the annual working hours. The nominal values are converted into 2011 dollars using CPI. Dashed and solid lines represent the normal density approximation of the data and estimated wage offer distributions, respectively.

2.2 Estimation Results: The Role of Health on Labor and Experience

We now estimate the nonlinear wage equation presented in (1), using wage processes after controlling for selection bias as described in 2.1.

Table 4 reports the impact of disability on the efficiency of labor and experience. We measure these effects in relative terms, using the non-disabled workers in the same education group as a benchmark. We find that for college-educated workers, the significant decline in wage is driven by the loss in labor efficiency. In contrast, the decline of efficiency in their experience is relatively small. For workers with high school education, however, the decline in both the effective labor and experience inputs were sizable. Overall, however, regardless of education, disability affects the efficiency of experience less than it does the efficiency of labor.

TABLE 4: ESTIMATED LABOR AND EXPERIENCE EFFICIENCIES RELATIVE TO THE NON-DISABLED

Individual characteristics		Relative Efficiency (Non-Disabled $\equiv 1$)	
Disability Status	Education	Labor	Experience
Disabled	high school	0.709^{***}	0.837^{***}
		[0.57, 0.87]	[0.58, 1.09]
	college	0.575^{***}	0.917^{***}
		[0.48, 0.68]	[0.70, 1.13]
Number of obs.		56,840	
R^2		0.238	

Note: Table 4 reports the coefficient estimation results of the nonlinear wage equation (1). The right hand variables also include quadratic function in age, cubic function in experience, dummy variables for gender, race, marital status and year. The complete list of estimated coefficients are reported in Appendix B. We uses individual-level survey weights for our analysis. Standard errors clustered at individual level. 95% confidence intervals are reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

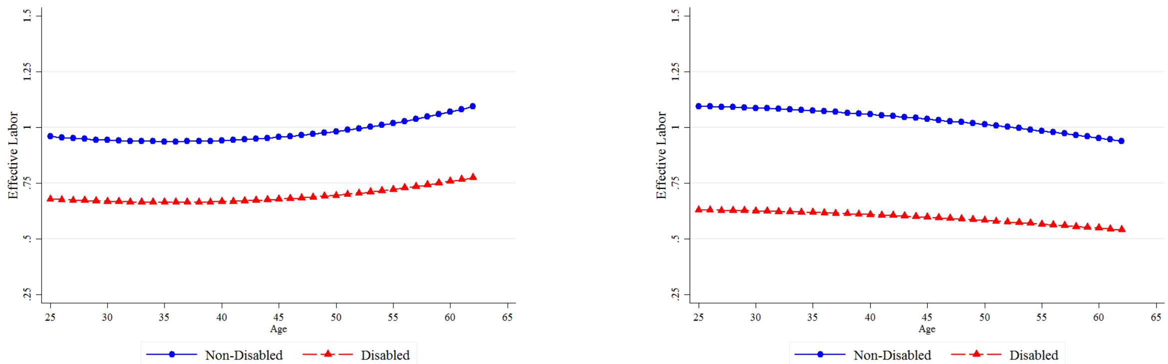
In Figures 2 and 3, we plot the efficiency profiles of labor and experience over the life cycle ($\lambda_L(j,h)$ and $\lambda_E(j,h)$). For both education groups, the efficiency of experience declines as workers age while labor endowment exhibit flatter profiles over the life cycle. This estimation result, however, does not imply that the worker's endowment in experience diminishes as they get older. Though the efficiency of using experience itself might decrease, they accumulate experience through working. The latter, represented by $g(e)$ in the wage equation is presented in Figure 4. Therefore, the total value in experience depends on the relative change in these two factors.

2.3 Estimation Results: The Elasticity of Substitution between Experience and Labor

In this section, we use the time-series variation in total endowments and price estimates, to identify the aggregate production function parameter. In the aggregate economy, homogeneous firms have access to

FIGURE 2: THE EFFICIENCY OF EFFECTIVE LABOR OVER THE LIFE CYCLE ($\lambda_L(j,h)$)

(A): HIGH SCHOOL (B): COLLEGE



Note: Figures 3 illustrate the estimated efficiency of labor by education group. The blue solid line with circles reflect the estimation results of workers without health limitations. The red dashed line with triangles are for workers with disabilities.

a production technology specified as $Y_t = A_t F(L_t, E_t) = A_t (L_t^\rho + \theta E_t^\rho)^{\frac{1}{\rho}}$. This function features constant elasticity of substitution (CES) between two inputs, labor L and experience E, with elasticity of substitution $(1-\rho)^{-1}$, and A_t represents the productivity of the economy at time t. The production function is increasing ($F_x > 0$) and concave ($F_{xx} < 0$) in L and E. Under the assumption of competitive labor markets, the price of each factor is equivalent to its marginal productivity in that period:

$$R_{E,t} = A_t F_{E,t} \quad \text{and} \quad R_{L,t} = A_t F_{L,t}. \tag{2}$$

Therefore, the relative price of experience is given as the ratio of marginal productivities: $\Pi_E \equiv F_E/F_L$. We can also write the aggregate production function $A_t F(L_t, E_t) = A_t (F_{L,t} L_t + F_{E,t} E_t)$ using the Euler theorem.

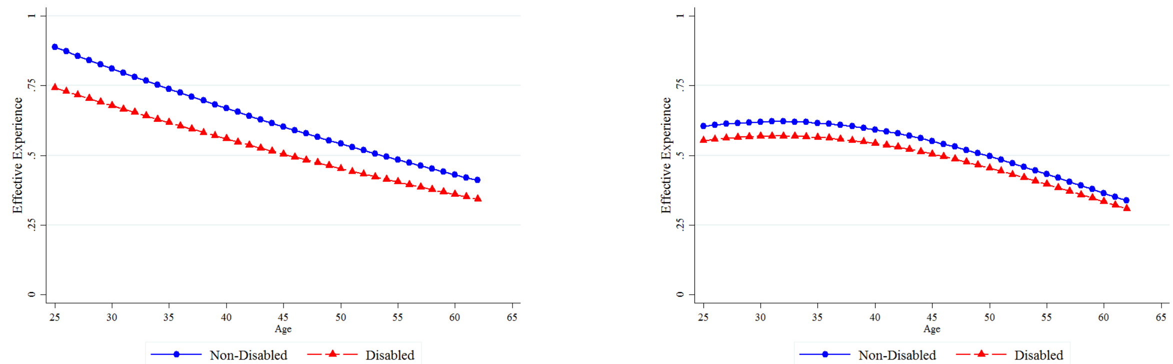
We construct the aggregate amount of labor and experience based on the wage estimation results, along with the estimated relative price of experience. Figure 5 illustrate the evolution of these two time-series variables. Using these data, we estimate the two production technology parameters. From the definition of the relative price,

$$\Pi_{E,t} \equiv \frac{R_{E,t}}{R_{L,t}} = \frac{\theta A_t (L_t^\rho + \theta E_t^\rho)^{\frac{1-\rho}{\rho}} E_t^{\rho-1}}{A_t (L_t^\rho + \theta E_t^\rho)^{\frac{1-\rho}{\rho}} L_t^{\rho-1}} = \theta \left(\frac{E_t}{L_t}\right)^{\rho-1},$$

and $\ln \Pi_{E,t} = \ln\theta + (\rho-1)\ln(E_t/L_t)$. Therefore, a linear regression using the aggregate time-series data of relative price and quantity delivers the values for θ and ρ.

FIGURE 3: THE EFFICIENCY OF EFFECTIVE EXPERIENCE OVER THE LIFE CYCLE ($\lambda_E(j,h)$)

(A): HIGH SCHOOL (B): COLLEGE



Note: Figures illustrate the estimated efficiency of experience by education group. The blue solid line with circles reflect the estimation results of workers without health limitations. The red dashed line with triangles are for workers with disabilities.

FIGURE 5: TRENDS OF LABOR AND EXPERIENCE IN THE U.S.

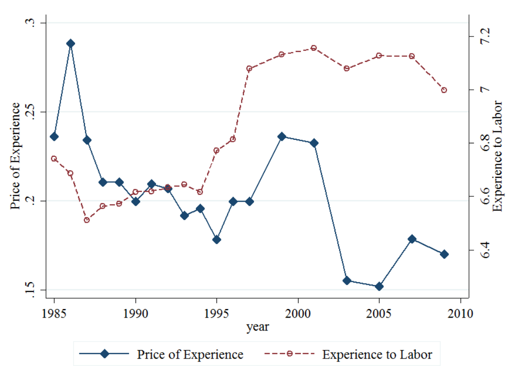


FIGURE 6: PRODUCTION TECHNOLOGY PARAMETERS

Parameters	Coefficient
ρ	-2.18^{***}
	(0.015)
$\ln\theta$	2.60^{***}
	(0.03)
Time periods	1985 to 2009
Adjusted R^2	0.249

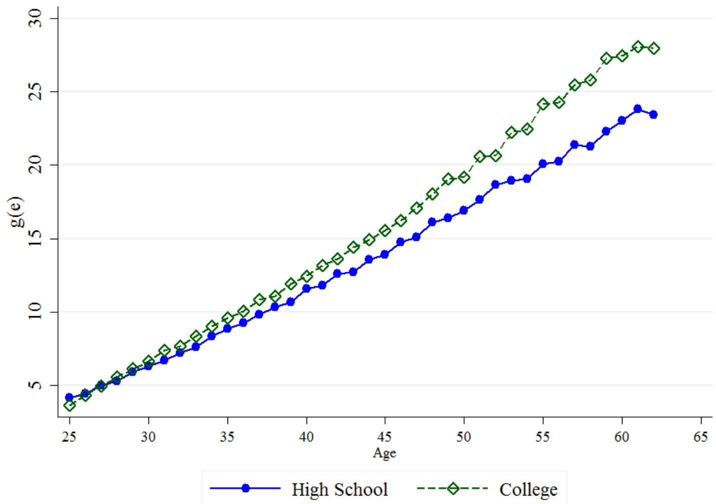
Note: Figure 5 illustrates the trends of aggregate-level experience-to-labor ratio and the price of experience inferred from the estimated wage equation. Table 6 reports the estimated CES production parameters based on the two time series; the experience-to-labor ratio and the price of experience using the reduced form approach.

On the left panel of Figure 5, we plot the relative supply of experience to labor ($R_{E,t}/R_{L,t}$) with the relative price of experience ($\Pi_{E,t}$) from 1985 to 2009, constructed from our estimation in the previous section. We find that the estimated $\rho = -2.18$, suggesting that labor and experience are complementary in production. Detailed estimation results are reported in Table 7.

3 The Model

In this section, we construct a stochastic life-cycle model of labor supply and savings subject to health and earning risks. Our model extends the framework of Imrohoroglu and Kitao (2012) by incorporating

FIGURE 4: THE ROLE OF ACCUMULATED EXPERIENCE $g(e)$



Note: Figure 4 illustrate the mean value of accumulated experience by age group, using the estimated coefficients from the nonlinear wage equation (1). Individuals are categorized as college graduates if they received more than 12 years of schooling. We uses individual-level survey weights to construct the group mean.

interaction of heterogeneous labor inputs in the labor market described in Section 2.

3.1 Environment

Demographics and Health Process The economy consists of overlapping generations of individuals who live for at most J periods. During their lifetime, individuals face exogenous risks on their health status and survival. Health status, h, is discrete and follows an age-specific Markov chain.

$$\pi_j^{ab} = \mathbf{Pr}\left(h_{t+1} = b \middle| h_t = a\right), \quad a, b \in \{1, 2, 3\}.$$

An individual of age j faces an age- and health-specific mortality rate $\delta_j^h \in (0, 1)$. The probability of death in his maximum age J into $J+1$ is 1. Therefore, in any point in time, there exist J different age groups in the economy. The law of motion of population for each age-group in time t is given as

$$n_{j+1} = \sum_h \left(1 - \delta_j^h\right) n_j^h, \quad j = 1, 2, \cdots J - 1.$$

The asset of the deceased will be distributed equally to all surviving members of the economy in a form of bequest transfer, b_t.

Preference We focus on the adulthood of the life-cycle and assume an individual enters the labor market at age 1 and must exit the labor market after reaching the mandatory retirement age $j^R < J$. Individuals have time-separable utility function that depends on the amount of consumption c, the value from leisure l, and their health status h. They maximize their life-time utility, discounting the future expected utility with time discount factor β:

$$\mathbb{E}\left[\sum_{j=1}^{J}\beta^{j-1}u\left(c_j,l_j;h_j\right)\right].$$

Individuals take expectation over the distribution of known survival probabilities and labor market opportunities to optimize their consumption, saving, and labor supply decision over the life-cycle.

3.2 The DI Program

The government uses its revenues to finance two social welfare programs, the social security benefits s_j for the retirees and the disability insurance programs (DI) for the disabled working-age population. To be eligible for DI, working-age individuals must meet the health status criteria and opt out of the labor force.[8] The DI benefits consist of two components: first, the government replaces the recipient's foregone labor income proportional to his previous earning, $d_j\left(\omega\right)$.[9] The government also provides medical insurance coverage for health care for workers with disabilities.

The government budget constraint in each period is given as

$$\sum_{j=j^R}^{J}s_j\mu_j\left(x\right)+\sum_{j=1}^{J^R}\left\{d_j\left(\omega\right)+m_j^h\right\}\mu_j\left(x\right) = \tau^w\sum_{j=1}^{j^R}y_j\mu_j\left(x\right)+\sum_{j=1}^{J}\tau^c a_j\mu_j\left(x\right) \tag{3}$$

where $\mu_j\left(x\right)$ denotes the measure of individuals of age with characteristics $x=\left(a,h,e,\omega\right)$ where a denotes assets holdings, h health status, e the years of work experience, and ω the average of previous earnings. In equilibrium, consumption and labor tax rates are determined such that the budget constraint (3) holds in a steady state.

[8]Social Security pays disability benefits if the worker does not engage in the Substantial Gainful Activity (SGA), which in 2015 is a monthly earning of $1,090 for non-blind individuals and $1,820 for blind individuals. If the worker earns more than the SGA, his benefit terminates.

[9]More specifically, under the current DI system in the United States, disability insurance payments are determined by the worker's Average Indexed Monthly Earnings (AIME), which is the average of the worker's 35 highest annual earnings.

3.3 The Individual's Problem

We characterize the problem of an individual in recursive form. An individual makes consumption (c), saving (a'), and labor supply (l) decisions. The optimization problem of working-age individual is given as

$$\begin{aligned} & V_j\left(a,h,e,\omega\right) & (4) \\ = \quad & \max_{c,l,a'} u\left(c,l;h\right)+\beta\left(1-\delta_j^h\right)\mathbb{E}\left[\left(1-l\right)V_{j+1}\left(a',h',e,\omega\right)+lV_{j+1}\left(a',h',e+1,\omega'\right)\right] & \\ s.t. \quad & c+a'+m_j^h\left(1-\left(1-l\right)\mathbb{I}_{\{h<h^{DI}\}}\right) & \\ & =(1-\tau^w)w_j\left(h,e\right)l+d_j\left(\omega\right)\left(1-l\right)\mathbb{I}_{\{h<h^{DI}\}}+\left\{1+\left(1-\tau_t^a\right)r\right\}a+b & \\ & w_j\left(h,e\right)=R_L\lambda_L\left(j,h\right)+R_E\lambda_E\left(j,h\right)g\left(e\right) & \\ & \omega'=f\left(w_j\left(h,e\right),\omega\right),\ h'\sim\pi_j^{h'h} & \\ & c\geq 0,\ l\in\{0,1\},\ a'\geq A. & \end{aligned}$$

Those workers whose health status falls below the threshold h^{DI} can choose to exit the labor force and receive disability payments. Note that if the worker's health improves, he can always come back to labor market.

Once retire, individuals receive social security benefits based on their accumulated earning history and choose optimal consumption and saving decisions:

$$\begin{aligned} V_j\left(a,h,e,\omega\right) \quad = \quad & \max_{c,a'} u\left(c,l;h\right)+\beta\left(1-\delta_j^h\right)\mathbb{E}\left[\left(1-l\right)V_{j+1}\left(a',h',e,\omega\right)\right] & (5) \\ s.t. \quad & \left(1+\tau_t^c\right)c+a'+m_j^h=ss\left(\omega\right)+\left\{1+\left(1-\tau_t^a\right)r\right\}a+b & \\ & h'\sim\pi_j^{h'h} & \\ & l\in\{0,1\},\ c\geq 0,\ a'\geq A. & \end{aligned}$$

Each period, individuals can trade risk-free bonds and face borrowing constraints with a borrowing limit of A. The interest rate on risk-free bonds is given (r) and constant over time. Individuals can also save at the same rate of return, r.

3.4 Competitive Equilibrium

Definition. Given a set of exogenous demographic parameters $\left\{n_j,\left(\delta_j^h,\delta_j^d\right)\right\}_{j=1}^J$, the stochastic process of health status and medical expenditure $\left\{\pi_j^{h'h},m_j^h\right\}_{j=1}^J$, and government policy variables $\left\{\tau^a,\tau^w,ss\left(\omega\right),d\left(\omega\right),h^{DI}\right\}$, a competitive equilibrium consists of individuals' policy functions $\{c_j(x), a_j(x), l_j(x)\}$ for each state variables $x = (a, h, e, \omega)$, factor price of labor and experience $\{R_E, R_L\}$, tax rates $\{\tau^a, \tau^w\}$, size of bequest transfers b, and the distribution of individuals over the state space $\{\mu_j(x)\}$ such that

1. Individuals solve the optimization problem defined as (4) and (5).

2. Factor prices R_L and R_E are determined competitively:

$$\begin{aligned} R_L &= A\left(L^\rho + \theta E^\rho\right)^{\frac{1-\rho}{\rho}} L^{\rho-1} \\ R_E &= \theta A\left(L^\rho + \theta E^\rho\right)^{\frac{1-\rho}{\rho}} E^{\rho-1}. \end{aligned}$$

3. Factor markets clear:

$$\begin{aligned} L &= \sum_{j \le j^R} \sum_x \mathbb{I}_{\{l_j(x)>0\}} \lambda_L(j,h)\, \mu_j(x)\, dx \\ E &= \sum_{j \le j^R} \sum_x \mathbb{I}_{\{l_j(x)>0\}} \lambda_E(j,h)\, g(e)\, \mu_j(x) dx. \end{aligned}$$

4. The lump-sum bequest transfer equals the amount of assets left by the deceased:

$$b = \sum_j \sum_x a_j(x)\, \delta_j^h \mu_{,j}(x).$$

5. The tax rates satisfy the government's budget constraint (3).

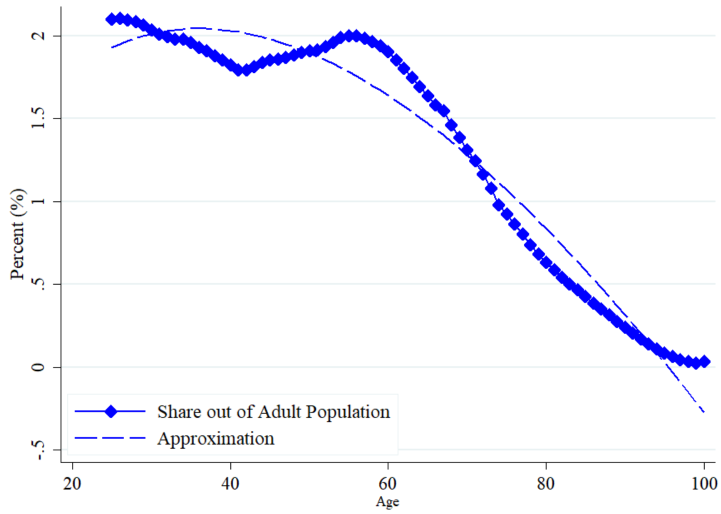
4 Quantitative Analysis

In this section, we quantitatively evaluate the effects by mapping the model in Section 3 into the U.S. economy. We start this section by explaining our choice of demographic parameters.

4.1 Calibration

Demographics The unit of time in our analysis is a year. We simulate at most 75 periods per individual, assuming that a person enters the economy at his age 25 and lives at most up to the age of 100. We target the demographic composition of the economy to be consistent with the US population composition of 2015. Information on demographics is taken from the National Population Projections by the U.S. Census. More detailed information is available in Appendix A.1.

FIGURE 7: THE COMPOSITION CHANGES IN THE U.S. DEMOGRAPHICS



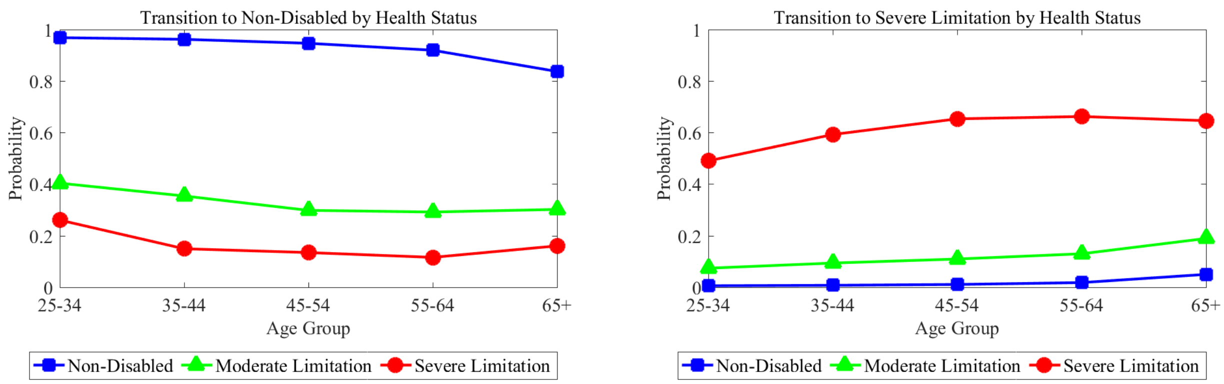
Note: Figure 7 illustrates the 2015 demographic composition of the U.S. based on the Census Population Projections Program. The dotted lines illustrate the percentage share of each age group out of total adult population between 25 and 100, and the dashed lines are quadratic approximation of the data.

Health and Medical Expenditures Our main source of micro-level data is the Panel Study of Income Dynamics (PSID). We use three categories of health statuses: non-disabled, moderately disabled, and severely disabled. Among those who are disabled, one is categorized as moderately disabled, if the disability limits the amount of work "Somewhat," or "Just a little" (rather than "A lot" or "Not at all"). Health status in the model has four roles: it affects the worker's (i) labor productivity, (ii) evolution of health status, (iii) medical expenditures, and (iv) survival probability. One of the key parameters for our model is the impact of health on workers' productivities over the life-cycle. We discuss our parametrization and identification strategy for labor productivity impact of health in Section 2. Here, we discuss how we parameterize the impact of health on its dynamics, medical expenditure risks, and survival probabilities

Health status in the model evolves stochastically according to $\pi_j^{h'h}$, which depends on the worker's age and initial health status. We use panel dimension of the PSID to find the transition probabilities by five age groups (25-34; 35-44; 45-54; 55-64; and 65 and older). Transition probabilities to non-disabled and severe limitation are plotted in Figure 8. As is clear from the plots, health statuses are persistent, and older workers are more likely to transition to having severe limitations.

Medical expenditure risks also differ by age group and health statuses as presented in Table 5. We use adult-equivalent medical expenditures from the PSID. Following Attanasio, Kitao and Violante (2011), we use three medical expenditure bins representing the averages in 1st-60th percentile, 61st-95th percentile,

FIGURE 8: HEALTH STATUS TRANSITION PROBABILITIES



Note: Figures illustrate the yearly transition probabilities of becoming non-disabled (left panel) and having severe limitations (right panel), conditional on the individual's initial health status and age group. The probabilities are computed based on the PSID data between 1984 and 2011, and are weighted by individual weights. In order to ensure sample sizes, we combine individuals into age groups. As the PSID changed its surveys from annual to biannual in 1997, we assume constant yearly transitions to construct annual transition probabilities from two-year transitions.

and 96th-100th percentile. They are chosen to capture the long tail in the medical expenditure distributions from catastrophic events. We assume that all employed workers have health insurance with coverage rate 60% (Kitao, 2014).

Lastly, we estimate the impact of health on conditional survival probabilities, using the life table from the Social Security and micro-level data from the PSID. Following the strategy of Attanasio, Kitao and Violante (2011), we obtain age-dependent survival probabilities ($\bar{s_j}$) from the life table, and the empirical health distribution by age (ϕ_j^h) and survival rate by health status and age s_j^h from PSID. Then, we use the following equations to obtain health-dependent conditional survival probabilities that are consistent with the life tables:

$$\bar{s}_j = \sum_{h \in 0,1,2} \phi_j^h s_j^h$$

$$prem_j^h = s_j^0 - s_j^h, \quad h \in \{1,2\}\,.$$

The second equation represents the survival premium of being non-disabled, relative to having moderate ($prem_j^1$) or severe limitations ($prem_j^2$). Given small samples in the PSID, we smooth out survival premia ($prem_j^h$) by fitting polynomials in age, and extrapolate them for individuals older than 90.[10] The estimated

[10]AKV uses Health and Retirement Survey (HRS) to calculate health-dependent survival probabilities. While HRS, which focuses more on older workers, might be a better dataset for constructing survival probabilities, it does not include disability measures consistent with PSID. When we compare our survival premia to their, the magnitudes seem similar.

TABLE 5: MEDICAL EXPENDITURES BY AGE AND HEALTH STATUS

		Mean within Percentiles			Average
Age Group	Health Status	1-60	61-95	96-100	
25-34	Non-Disabled	538	4,123	43,165	3,915
	Moderate Limitation	707	8,567	57,949	5,772
	Severe Limitation	649	10,124	46,067	5,906
35-44	Non-Disabled	791	4,944	40,847	4,241
	Moderate Limitation	1,179	9,343	98,217	8,733
	Severe Limitation	1,244	15,970	98,823	11,103
45-54	Non-Disabled	976	6,200	62,988	5,859
	Moderate Limitation	1,470	11,835	78,583	8,824
	Severe Limitation	1,415	15,314	120,310	11,747
55-64	Non-Disabled	1,240	7,822	68,988	6,860
	Moderate Limitation	1,761	15,742	101,535	11,170
	Severe Limitation	2,804	22,946	128,267	15,614
65+	Non-Disabled	1,457	9,890	76,955	8,144
	Moderate Limitation	2,197	15,380	116,132	12,457
	Severe Limitation	2,419	18,155	131,143	14,057

Note: Medical expenditures are calculated from the PSID, years from1999 (the earliest year in which medical expenditure data is available) through 2011. In order to ensure enough sample sizes, we combine ages 65 and over.

health-dependent conditional probabilities are plotted in 9.

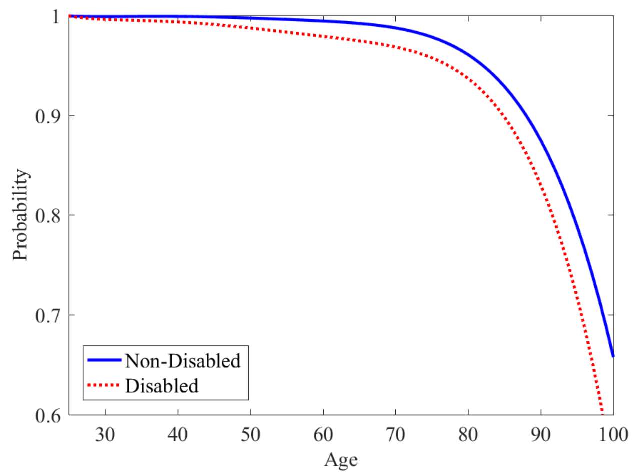
Government Policies Disability Insurance, Social Security, and Medicare are the essential policies of interest. We also capture welfare programs using a consumption floor and allow the government to collect revenues using labor and capital income taxes.[11]

Disability Insurance Both Disability Insurance and Social Security (which we discuss below) payments are determined by the worker's Average Indexed Monthly Earnings (AIME), which is the average of the worker's 35 highest annual earnings. Since it is difficult to keep track of earnings history, we use average prior earnings (ω in our model) to calculate the Primary Insurance Amount (PIA), whose formula is given as follows in 2011 dollars:

$$PIA = \begin{cases} 0.90 \times \omega & \text{if } \omega < \$8,988 \\ \$8,089 + 0.32 \times (\omega - \$8,988) & \text{if } \$8,988 \leq \omega < \$54,204 \\ \$22,559 + 0.15 \times (\omega - \$54,204) & \text{if } \omega \geq \$54,204. \end{cases}$$

[11] Our policy parametrizations closely follow Kitao (2014) and Imrohoroglu and Kitao (2012).

FIGURE 9: CONDITIONAL SURVIVAL PROBABILITY BY HEALTH STATUS



Note: Figure illustrates the conditional survival probability by health status.

As there is a cap on the AIME for benefit calculation, we impose a cap on DI payments, thus $d(\omega) = \min\{PIA, \$30,448\}$.

Social Security and Medicare Social Security taxes are set at $\tau_{ss} = 0.104$, which is levied on labor earnings of workers, with maximum taxable earning of $y_{ss} = \$106,800$. Social Security payments are determined by the PIA (with the maximum amount being $\$30,448$ as in DI) and the retirement age of workers. For cohorts born between 1943 and 1954, the normal retirement age (NRA) is 66; early retirement age (ERA), 62; and delayed retirement age (DRA), 70. For early retirees, their PIA is reduced by the actuarial reduction factor (ARF), which is $\{0.25, 0.20, 0.133, 0.067\}$ respectively for ages between 62 and 65. On the other hand, the delayed retirement credit (DRC) is 0.08 for each year of delay in retirement, up to age 70. Medicare benefits are provided to all individuals of 65 or older and those receiving DI[12]. Beneficiaries pay premium of $p_{med} = \$1,157$, and its coverage rate is $q_{med} = 0.75$. Medicare tax rate is $\tau_{med} = 0.029$, levied on labor earnings.

Other Taxes and Welfare Programs Labor income is taxed at rate $\tau(\omega)$, determined by

$$\tau(\omega) = \tau_0 \left\{ \omega - \left(\omega^{-\tau_1} + \tau_2 \right)^{-\frac{1}{\tau_1}} \right\},$$

where parameters $\{\tau_0, \tau_1, \tau_2\}$ jointly determine the level, progressiveness, and overall scale of tax revenue. We follow the estimation strategy of Gouveia and Strauss (1994) and calibrate τ_0 and τ_1 outside the model to be consistent with current U.S. tax codes. The scale parameter τ_2 is calibrated within the model such that

[12] Technically, only two years later qualify for DI, but for simplicity, we assume everyone is qualified.

the relative size of spending on welfare programs remains constant with respect to the aggregate output of the economy. Capital income tax rate is $\tau_k = 0.30$. We use consumption floor $\underline{c}_f = \$3,150$ per month to capture other welfare programs provided by the government (such as Supplemental Nutrition Assistance Program).

Table 6 summarizes the list of parameters taken from outside the model for describing demographic structure, health and medical expenditure processes, and current government policies.

TABLE 6: PARAMETERS CALIBRATED OUTSIDE THE MODEL

Parameters	Description	Values/Sources
Demographics		
$\{n_j\}_{j=1}^{75}$	2015 US population share	Reference the main text 4.1.
$\{\delta_j^h\}_{j=1}^{75}$	health- and age-specific mortality rates	Reference the main text 4.1.
Preference		
γ	risk aversion parameter	2
Health and Medical Expenditure		
$\left\{\pi_j^{h'h}\right\}$	health- and age-specific health transition probabilities	Reference the main text 4.1.
$\left\{m_j^h, \sigma_{m,j}^h\right\}$	health- and age-specific medical expenditure shocks	Reference the main text 4.1.
Current Government Policy Parameters		
$\{d(\omega)\}$	the disability insurance payment schedule	Reference the main text 4.1.
$\{ss(\omega)\}$	the social security program payment schedule	Reference the main text 4.1.
τ_{ss}	Social Security tax rate	10.4%
y_{ss}	maximum taxable earnings for Social Security	\$106,800
p_{med}	Medicare premium	\$1,158
q_{med}	Medicare coverage rate	0.75
τ_{med}	Medicare tax rate	2.9%
$\{\tau_0, \tau_1, \tau_2\}$	labor income tax function	$\{0.26, 0.73, 0.01\}$
τ_k	capital income tax rate	30%
c_f	other welfare program guaranteeing minimum consumption	\$3,150 per month

Note: Table 6 summarizes the calibration values for the parameters in demography, health-related risks, and the government policies. Monetary values are written in 2011 US dollars, adjusted by CPI.

TABLE 7: PARAMETERS CALIBRATED OUTSIDE THE MODEL: WAGE PROCESS PARAMETERS

Parameters	Description	Values/source
$\{\lambda_L, \lambda_E\}$	productivity process of labor and experience	Reference Appendix B.2.
$\{\phi_L^s, \phi_E^s\}$	the change in productivities on labor and experience due to disabilities	$\{0.71, 0.84\}$ for high school graduates
		$\{0.58, 0.92\}$ for college graduates
$g(e)$	the accumulated amount of experience in effective unit	Reference Appendix B.2.
σ_ε	idiosyncratic component of log hourly wage	0.67 (RMSE of estimation results)

Note: Table 7 summarizes the calibrated parameter values associated with labor market. Please reference the details of these calibration process in maintext 2. The complete list of parameter values can be found in Appendix B.2.

4.2 Calibration Results

Now we turn to our quantitative analysis based on empirical observations and conduct the following counterfactual analyses to measure the impact of DI. We start by reporting the statistics used as calibration targets in our the simulation of method of moments.

Parameters Calibrated with the Simulated Method of Moments We choose the following functional form as the periodic utility function and estimate three parameters to be calibrated within the model: preference parameters on time β, disutility from work θ, and interaction of consumption and health status η:

$$u(c, l) = \frac{\left\{c \exp(\theta l) \times \exp\left(\eta \mathbb{I}_{\{\text{disability}=1\}}\right)\right\}^{1-\gamma}}{1-\gamma}.$$

For our estimation, we directly target the average wealth, employment rates of the economy by health status as our targets. Results are summarized in Tables 8 and 9.

TABLE 8: TARGETS: DATA VS. SIMULATED METHOD OF MOMENTS

Moments	Data	Simulation
Employment rates of the non-disabled	82.1%	79.8%
Employment rates of the disabled	55.3%	47.4%
Consumption	$\$24,682$	$\$26,235$

Note: Table 8 summarizes the moments used in estimation and their simulation counterparts. Statistics are computed from the PSID using individual survey weights.

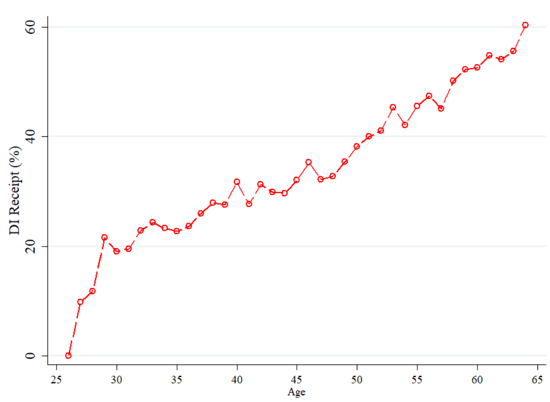
TABLE 9: PARAMETERS CALIBRATED WITHIN THE MODEL

Parameters	Description	Values
β	time discount factor	0.97
θ	interaction of leisure and consumption	-0.23
η	interaction of consumption and health status	-0.31

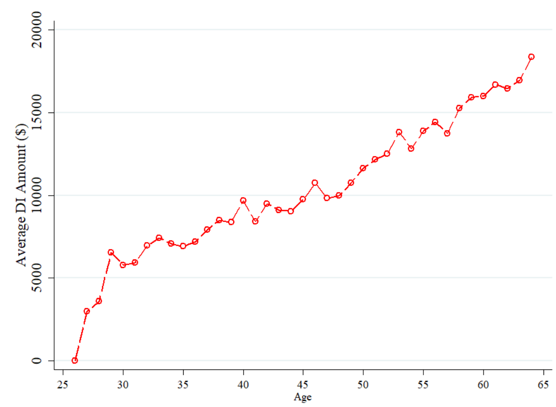
Note: Table 9 reports the value of three preference parameters estimated within the model using the simulated method of moments. The detailed discussion on estimation procesures can be found in appendix C.

FIGURE 10: SIMULATION: DI PROGRAM UTILIZATION

(A): THE SHARE OF DI RECIPIENTS

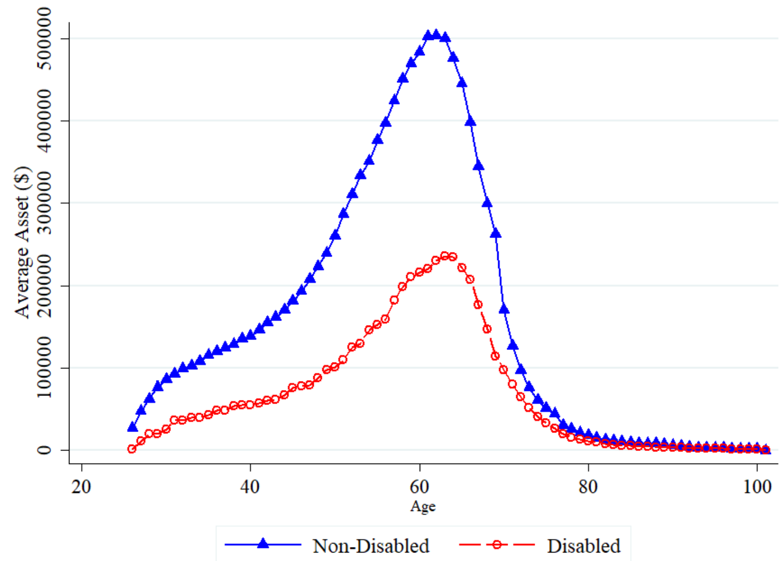


(B): THE DI PAYMENT



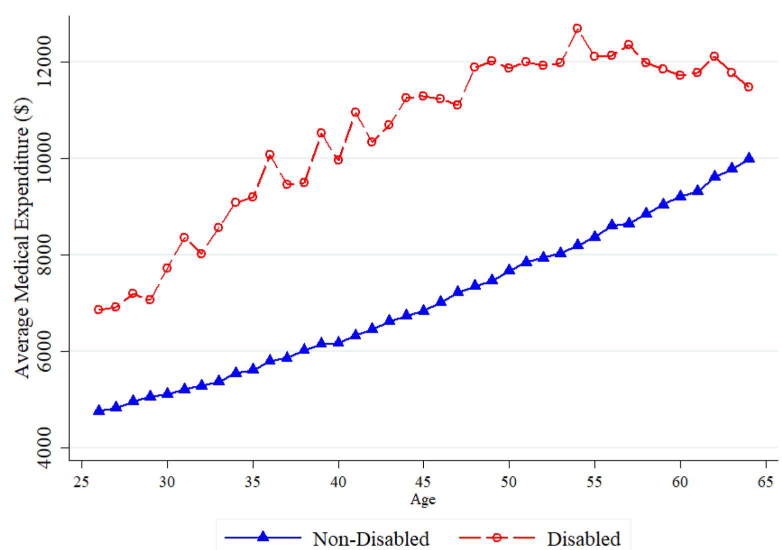
Note: Figures show the pattern of DI program utilization in simulated economy. Figure 10(a) illustrates the share of DI recipients among the disabled over the life-cycle. Figure 10(b) computes the average monthly payment of DI for those who collect the DI benefits.

FIGURE 11: SIMULATION: ASSET ACCUMULATION OVER THE LIFE CYCLE



Note: Figure 11 illustrates the trends of median wealth in simulated economy by health and age. The blue solid line with triangles indicate workers without work limitation. The red dashed line with circles is the same statistics of the non-disabled.

FIGURE 12: SIMULATION: MEDICAL EXPENDITURE OVER THE LIFE-CYCLE



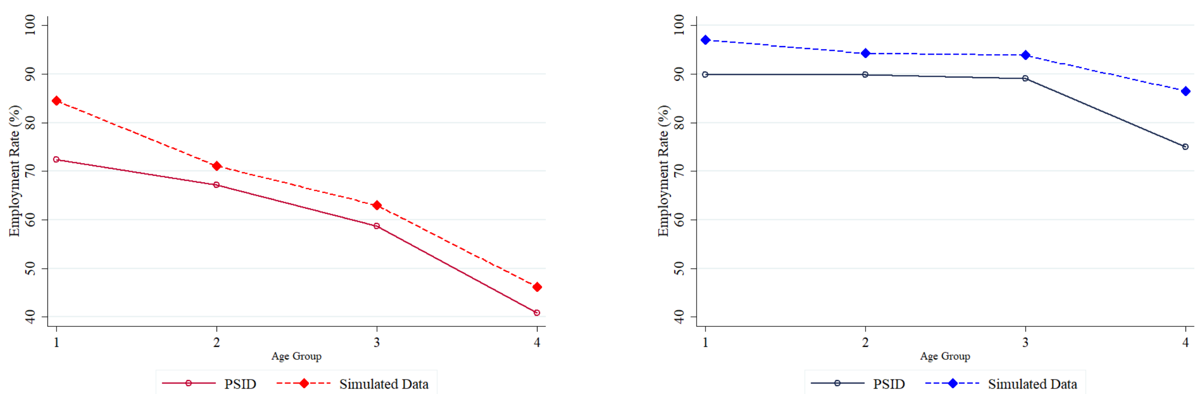
Note: Figure 12 illustrates the trends of medical expenditure by health and age. The blue solid line with triangles indicate workers without work limitation. The red dashed line with circles is the same statistics of the non-disabled.

The Model's Fit on Nontargeted Moments Even though we do not directly target the employment rate by age group, the simulated data replicates well the life-cycle pattern of employment rates and exhibit the rapid outflow from the labor force starting from the ages in late 50s. It also reflects the more faster rate of early retirement for workers with disabilities. These patterns are documented in Figure 13.

FIGURE 13: SIMULATION: EMPLOYMENT RATE OVER THE LIFE CYCLE

(A): DISABLED

(B): NON-DISABLED



Note: Figure 13(a) contrasts the employment rates of working-age population with work-limitation by age group between the PSID and simulated economy. Figure 13(b) is the same statistics for workers without work-limitations.

Another non-targeted moments include the labor market wages compared to the potential offers by the health status. As Table 10 shows, the model generates the selection in labor market participation decision in a similar magnitude as we observed in Section 2.

TABLE 10: ACCEPTED WAGE VS. WAGE OFFER DISTRIBUTIONS

Disability Status	Ratio between offer-to-accepted wages (%)	
	Data	Simulation
Non-Disabled	99.0	97.6
Disabled	91.9	90.7

Note: Table 10 reports the ratio of mean potential wage offer compared to the mean of accepted wages computed from the PSID and compare the counterparts in simulated data.

4.3 The Impacts of the Disability Insurance Program

To evaluate the impact of current DI program on labor market participation decisions, we conduct the following counterfactual analysis. Holding all the calibrated parameters constant, we simulate an economy without the current DI program. As the agents learn this information when they optimize their utilities, this change alters the equilibrium prices of labor and experience as well. While the tax system is remained the same, collected revenue is returned in form of lump-sum transfer. By doing so, we preserve the progresiveness of the current tax policy and maintain budget neutrality simultaneously. Table 11 summarizes the differences in these two economies. Without disability insurance, individuals participate the labor markets. As the older population attach to the market, the relative supply of experience compared to the labor increases.

TABLE 11: THE IMPACT OF DI PROGRAM: SUMMARY STATISTICS

Variables	Current Economy with the DI program	Counterfactual Economy without the DI program	Difference
DI recipients	22.6%	-	-
Relative supply of experience	6.7	6.1	8.9%
Aggregate employment rate	80.1%	82.5%	2.4 ppt
Non-disabled	84.8%	87.5%	2.7 ppt
Disabled	57.4%	63.7%	6.5 ppt

Note: Table 11 reports the summary statistics of the counterfactual economy compared to the empirical counterparts computed from the PSID and CFS.

Decomposition Exercise The difference between the two economies is a result of two changes: the direct impact of changes in insurance against health-related risks and the changes in the value of labor and experience. The latter impact comes from the fact that the price of each production input is interdependent to the composition of the labor force. To measure the relative importance between these two mechanisms, we conduct additional counterfactual analysis.

In this quantitative exercise, we assume the experience and labor prices in the labor market remains the same as the stationary equilibrium of the U.S. with the current form of DI program, and remove the DI policy. Therefore, the changes in individuals' behavior is effectively generated by the absence of social safety net against health risks.

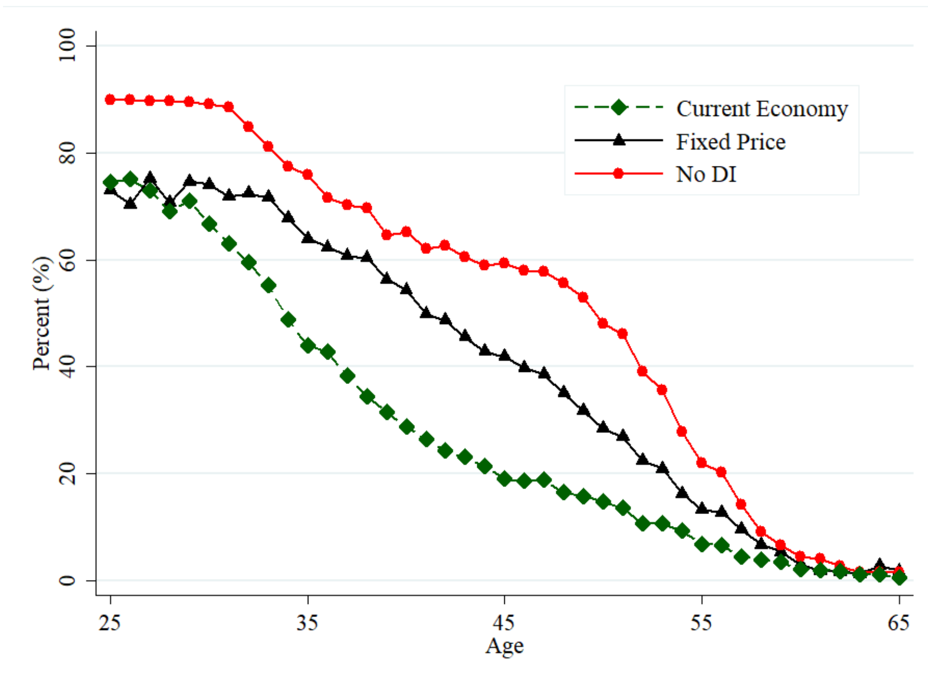
Table 12 reports the summary statistics of three economies. The first column reproduces the summary statistics of the economy with benchmark calibration. The second column of Table 12 computes the changes in an economy without DI program, under the fixed market prices of labor and experience. Removal of the DI program induces higher labor market participation of the disabled. However, quantitative importance of this direct impact is moderate; the employment increases by 0.4 percentage points. The total impact accompanied with the indirect impact, however, is approximately twice larger. This is because the change in factor prices affect both the disabled and non-disabled. Removal of social safety net increases the employment rates of the older population, causing the relative supply of employment improve. This change improves the overall earnings, increasing the employment rates of all.

TABLE 12: DECOMPOSITION EXERCISE: SUMMARY STATISTICS

Variables	(1) Current Economy	(2) Changes in Price & removal of DI (%)	(3) Fixed Price & removal of DI (%)
DI recipients	22.6%	-	-
Relative supply of experience	6.7	6.1	6.3
Aggregate employment rate	87.1%	88.1%	87.5%
Non-disabled	93.1%	97.5%	93.8%
Disabled	48.2%	53.7%	50.3%

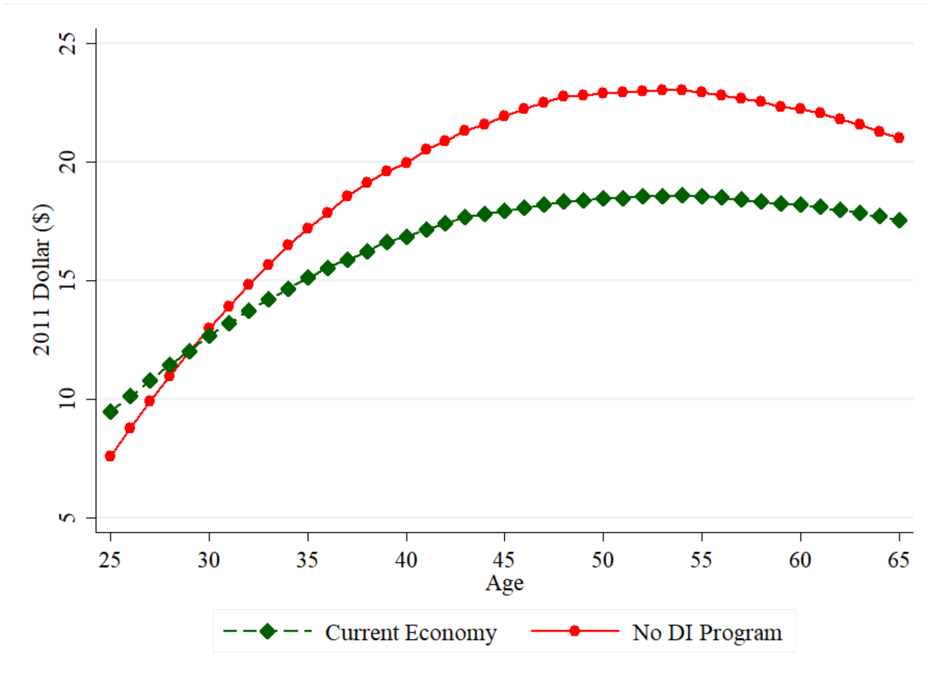
Note: Table 12 reports the summary statistics of the two counterfactual economies compared to the empirical counterparts computed from the PSID and CFS.

FIGURE 14: COUNTERFACTUAL ANALYSIS: IMPACTS ON THE LABOR FORCE PARTICIPATION RATES



Note: Figure illustrates the employment rates.

FIGURE 15: COUNTERFACTUAL ANALYSIS: IMPACTS ON THE LABOR FORCE PARTICIPATION RATES



Note: Figure illustrates the employment rates.

5 Conclusion

Disability insurance serves as an important source of insurance for workers with disabilities. However, it is well-known that it creates sizable disincentives on labor supply of workers. Our goal in this paper is understanding the aggregate implications of DI, especially its impact on labor supply, output, and aggregate productivity. Towards that goal, we estimated the role of disability on a worker's two kinds of human capital–(pure) labor and experience. We find that disability has a relatively small impact on a worker's effective experience, most of which is owned by old individuals who are more likely to be disabled and receive DI. Thus, the foregone amount of effective experience from withdrawal of older disabled workers is large. Moreover, as labor and experience are complementary inputs in the aggregate production, lower supply of experience has a pecuniary externality, changing the relative price of experience. We find that, in our quantitative model, the current DI program reduces the output by 3% as the aggregate labor and experience decline by 0.7 and 2.5%, respectively.

Our findings in this paper are not limited to the context of DI, but have far-reaching implications especially in aging economies, like the U.S. Demographic change naturally affects the relative supply of labor and experience. Thus, it would be interesting to understand the aggregate efficiency implications within our model framework and study the role of policies that affect labor supply decisions of workers (e.g., increase in mandatory retirement age, or changing the Social Security payment schedule). Moreover, we can also analyze the interaction of DI and Social Security. We leave these important questions to future research.

References

Attanasio, Orazio, Sagiri Kitao, and Giovanni L. Violante. 2011. "Demography and the Economy." , ed. John B. Shoven, Chapter Financing Medicare: A General Equilibrium Analysis, 333–366. University of Chicago Press.

Bound, John. 1989. "The Health and Earnings of Rejected Disability Insurance Applicants." *American Economic Review*, 79: 482–503.

Braun, R. Anton, Karen Kopecky, and Tatyana Koreshkova. 2017. "Old, Sick, Alone and Poor: A Welfare Analysis of Old-Age Social Insurance Programs." *Review of Economic Studies*, 84: 580–612.

Card, David, and Thomas Lemieux. 2001. "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." *Quarterly Journal of Economics*, 116(2): 705–746.

Congressional Budget Office. 2016. "Social Security Disability Insurance: Participation and Spending."

Currie, Janet, and Gruber Jonathan. 1996. "Health Insurance Eligibility, Utilization of Medical Care, and Child Health." *Quarterly Journal of Economics*, 111(2): 431–466.

De Nardi, Mariacristina, Eric French, and John B. Jones. 2016. "Medicaid Insurance in Old Age." *American Economic Review*, 106(1): 3480–3520.

French, Eric, and Jae Song. 2014. "The Effect of Disability Insurance Receipt on Labor Supply." *American Economic Journal: Policy*, forthcoming.

Gouveia, Miguel, and Robert P. Strauss. 1994. "Effective Federal Individual Tax Functions: An Exploratory Empirical Analysis." *National Tax Journal*, 47(2): 317–339.

Gruber, Jonathan, and Kevin Milligan. 2010. "Social Security Programs and Retirement around the World: The Relationship to Youth Employment." , ed. Jonathan Gruber and David A. Wise, Chapter Do Elderly Workers Substitute for Younger Workers in the United States?, 345–360. University of Chicago Press.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1): 153–161.

Hubbard, R.G., J Skinner, and S. P. Zeldes. 1995. "Precautionary Savings and Social Insurance." *Journal of Political Economy*, 103: 360–399.

Imrohoroglu, Selahattin, and Sagiri Kitao. 2012. "Social Security Reforms: Benefit Claiming, Labor Force Participation, and Long-run Sustainability." *American Economic Journal: Macroeconomics*, 4(3): 96–127.

Jeong, Hyeok, Young Kim, and Iourii Manovskii. 2015. "The Price of Experience." *American Economic Review*, 105(2): 784–815.

Karabarbounis, Loukas, and Brent Neiman. 2014. "The Global Decline of the Labor Share." *Quarterly Journal of Economics*, 129(1): 61–103.

Kitao, Sagiri. 2014. "A Life-Cycle Model of Unemployment and Disability Insurance." *Journal of Monetary Economics*, 68: 1–18.

Krusell, Per, Lee E. Ohanian, Jose-Victor Rios-Rull, and G. L. Violante. 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica*, 68: 1029–1054.

Low, Hamish, and Luigi Pistaferri. 2015. "Disability Insurance and the Dynamics of the Incentive-Insurance Tradeoff." *American Economic Review*.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand. 2013. "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt." *American Economic Review*, 103(5): 1797–1829.

Munnell, Alicia H., and April Yanyuan Wu. 2012. "Will Delayed Retirement by the Baby Boomers Lead to Higher Unemployment Among Younger Workers?" *Working Paper, Center for Retirement Research at Boston College*.

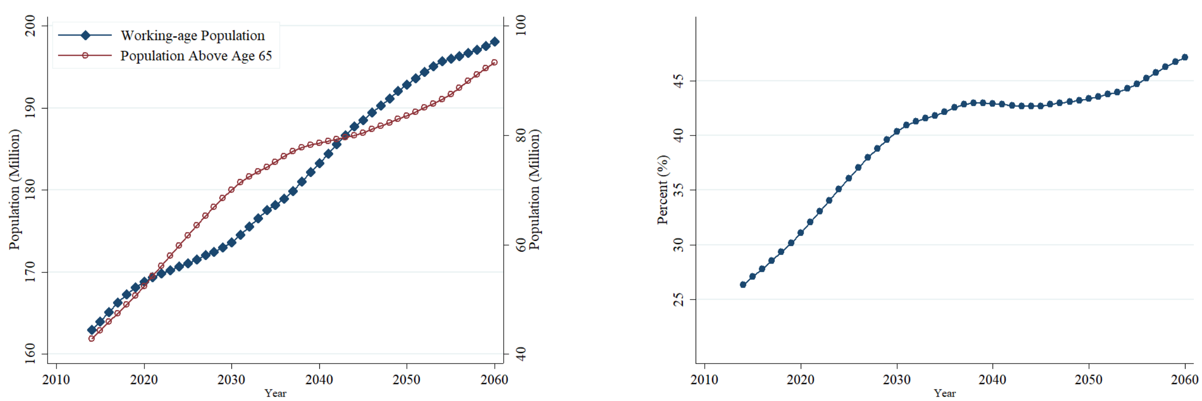
A Data Appendix

A.1 Census Population Estimates

We constructed demographic variables using 2014 version of the Population Projections Program, obtained from the U.S. Census Bureau. The Population Projections Program provides projected estimates of demographic compositions by age, sex, race, and ethnicity using the most recent decennial Census. The 2014 Population Projection is based on the 2010 Census, and the analysis was conducted in 2013 based on the cohort method under the assumptions on future fertility, mortality, and migration rates.

FIGURE 16: POPULATION PROJECTION OF THE UNITED STATES

(A): POPULATION SIZE BY AGE (B): THE SHARE OF OLDER POPULATION



Note: Figure 16(a) illustrates the projected total population of working-age population (age between 25 and 65) and older population (age between 65 and 100) of the U.S. based on the 2014 Census Population Projections Program. Figure 16(b) is the ratio of the older population out of the total working-age population in percentage term.

A.2 Welfare Programs, Tax Credits and Potential Benefits

We only observe wages of workers who joined the labor force and employed. Therefore, we may underestimate the effects of health risks on hourly wage if we do not take into account individuals who opted out of the labor force. To address this selection bias problem in our wage estimation, we adopt Heckman's Two-Stage Estimation (Heckman, 1979). In the first stage estimation, we run the following probit regression using potential government transfers as our exclusion restriction:

$$L_{it} = Z_i'\delta + \alpha D_{it} + \phi T_{it} + \nu_i \tag{6}$$

where $L_{it} \in \{0, 1\}$ is an indicator variable of employment; T_{it} represents the vector of exclusion restrictions; and D_{it} is the dummy for individuals with work limitations. Using the estimates, we construct the inverse Mills ratio $\hat{\lambda} = \Phi\left(\hat{L}\right)$ and include it in our second stage wage equation estimation along with other regressors presented in (1). As we report in Table 14, the estimated coefficient of the inverse Mills ratio in the second stage is positive and significant, indicating the existence of selection bias: the observed wages of the disabled are from the right tail of the underlying wage offer distribution.

Exclusion restrictions are variables that influence labor market participation decision (L_{it}) but not affect the wages individuals actually receive once we control for their participation decision ($L_{it} = 1$) as well as other observable characteristics. Similar to Currie and Jonathan (1996) and Low and Pistaferri (2015), we take the "potential benefits" from the government as well as its interaction with health status as our exclusion restrictions. We compute the potential benefits for representative household to each federal or state-level welfare programs in a given year. Unlike the actual transfer amounts, which are endogenous, these potential benefits are exogenous by default.

Following Low and Pistaferri (2015), we construct a part of variables in T_{it} based on the following welfare programs: Earned Income Tax Credit (EITC), Unemployment Insurance (UI), Food Stamps, Aid to Families with Dependent Children (AFDC) and Temporary Assistance for Needy Families (TANF). We also include potential tax credits in our exclusion restriction: American Opportunity Credit, Student Loan Interest Deduction, and Home Mortgage Interest Deduction. We find that having a wide range of potential benefits as our exclusion restrictions serve a better purpose as our research is studying both high school and college graduates.[13]

B Estimation of Wage Equation

B.1 Sample Selection and Variable Construction

We use PSID as our main source of data for wage process estimation. Our sample consists of individuals in working age between 25 to 65, both male and female in all education level. We exclude observations missing key information on disability status. We take the self-reported measure of disability questionnaire in PSID as our main indicator of health status. We also drop observations missing key information such as

[13] Low and Pistaferri (2015) focused on samples with high school education to study tradeoffs between welfare benefit from disability insurance and its costs from limiting work-incentives.

age, schooling, years of experience if we couldn't fill the gaps even after exploring the past observations in panel data.

One of the key variables in our empirical analysis is the years of work experience. Following Jeong, Kim and Manovskii (2015), we take the basis years which PSID directly asked the years of prior work and construct the experience variable by adding experience once an individual worked at least 750 hours per year.[14] Similarly, when we observe an individual with his first experience variable larger than one when he is older than 18, we construct the experience variable backward in time for his younger working life. Starting from 1999, PSID surveyed samples from annual to biennial frequency. Accordingly, we adjusted the gap and add two years of experience when he worked full time in the past year.

B.2 Selection Bias in Wage Estimation

As we describe in A.2, we address the selection bias in our wage equation estimation by taking the approach of Heckman's two-stage estimation.

[14]PSID asked the years of experiences in 1974, 1975, 1976, and 1985 for every heads and wives of households. In subsequent sample years, PSID collects this information for new heads and wives.

TABLE 13: FIRST-STAGE ESTIMATION RESULTS

		(1)		(2)
Disability Status	Disabled	-1.362^{***}	Moderate Limitation	-0.614^{***}
		(0.100)		(0.102)
			Severe Limitation	-1.526^{***}
				(0.158)
Demography	Years of schooling	0.047^{***}	Years of Schooling	0.047^{***}
		(0.006)		(0.005)
	Male	0.060^{**}	Male	0.064^{**}
		(0.027)		(0.027)
	Black	-0.051^{*}	Black	-0.051^{*}
		(0.030)		(0.030)
	Married	-0.004	Married	-0.003
		(0.023)		(0.023)
	Age	-0.035^{***}	Age	-0.034^{***}
		(0.005)		(0.005)
	Age^2	-0.0002^{**}	Age^2	-0.0003^{**}
		(0.0001)		(0.0001)
	Experience	0.258^{***}	Experience	0.258^{***}
		(0.008)		(0.008)
	Experience^2	-0.009^{***}	Experience^2	-0.009^{***}
		(0.0005)		(.0005)
	Experience^3	0.0001^{***}	Experience^3	0.0001^{***}
		$(7.79e-06)$		$(7.74e-06)$
Exclusion Restrictions	Potential benefits	-0.069^{***}	Potential benefits	-0.071^{***}
		(0.026)		(0.026)
	Potential benefit	-0.290^{***}	Potential benefit	-0.062
	× Moderate Limitation	(0.081)	× Moderate Limitation	(0.072)
	Potential benefit	0.244^{***}	Potential benefit	-0.020
	× Severe Limitation	(0.110)	× Severe Limitation	(0.111)
	Potential benefits 2 (housing)	-0.025^{***}	Potential benefits 2 (housing)	-0.019^{**}
		(0.008)		(0.008)
	Potential benefit 2× Disability	0.075^{***}	Potential benefit 2	0.033^{**}
		(0.014)	× Moderate Limitation	(0.014)
			Potential benefit 2	0.044^{*}
			× Severe Limitation	(0.023)
	Year Dummy		Year Dummy	Yes
	State Dummy		State Dummy	Yes
R^2		0.272		0.274
Number of observations		68,332		68,332

Note: Table 13 reports the estimated coefficients from the first stage of Heckman two-stage estimation. Other regressors include individual characteristics such as age, experience, years of schooling, marital status, states as well as time-varying year dummies, male dummies, and race dummies. Standard errors clustered at individual level and reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

TABLE 14: SECOND-STAGE WAGE ESTIMATION RESULTS

		(1) No Control	(2) Control		(2) No Control	(2) Control
Disability Status	Disabled	-0.151^{***}	-0.258^{***}	Moderate Limitation	-0.146^{***}	-0.209^{***}
		(0.018)	(0.033)		(0.019)	(0.026)
				Severe Limitation	-0.174^{***}	-0.411^{***}
					(0.039)	(0.072)
Demography	Years of schooling	0.118^{***}	0.124^{***}		0.118^{***}	0.125^{***}
		(0.003)	(0.003)		(0.003)	(0.004)
	Male	0.163^{***}	0.162^{***}		0.163^{***}	0.162^{***}
		(0.015)	(0.015)		(0.015)	(0.015)
	Black	-0.179^{***}	-0.185^{***}		-0.179^{***}	-0.185^{***}
		(0.030)	(0.015)		(0.015)	(0.015)
	Married	0.034^{***}	0.033^{***}		0.034^{***}	-0.033
		(0.011)	(0.011)		(0.011)	(0.011)
	Age	0.005^{*}	-0.001		0.005^{*}	-0.002
		(0.005)	(0.003)		(0.003)	(0.004)
	Age^2	-0.0002^{***}	-0.0001^{**}		-0.0001^{**}	-0.0002^{*}
		(0.0001)	(0.00006)		(0.00006)	(0.00006)
	Experience	0.061^{***}	0.099^{***}		0.061^{***}	0.102^{***}
		(0.004)	(0.009)		(0.004)	(0.011)
	Experience^2	-0.002^{***}	-0.003		-0.002^{***}	-0.003^{***}
		(0.0002)	(.0004)		(0.0002)	(0.0004)
	Experience^3	0.00002^{***}	0.00004^{***}		0.00002^{***}	0.00004^{***}
		$(4.00e-06)$	$(5.24e-06)$		$(4.00e-06)$	$(5.82e-06)$
	Inverse Mills Ratio		0.297^{***}			0.325^{***}
			(0.073)			(0.087)
	Year Dummy	Yes	Yes		Yes	Yes
	State Dummy	Yes	Yes		Yes	Yes
R^2		0.266	0.266		0.266	0.266
Number of observations		56,840	56,840		56,840	56,840

Note: Table 14 reports the estimated coefficients from the first stage of Heckman two-stage estimation. Other regressors include individual characteristics such as age, experience, years of schooling, marital status, states as well as time-varying year dummies, male dummies, and race dummies. Standard errors clustered at individual level and reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

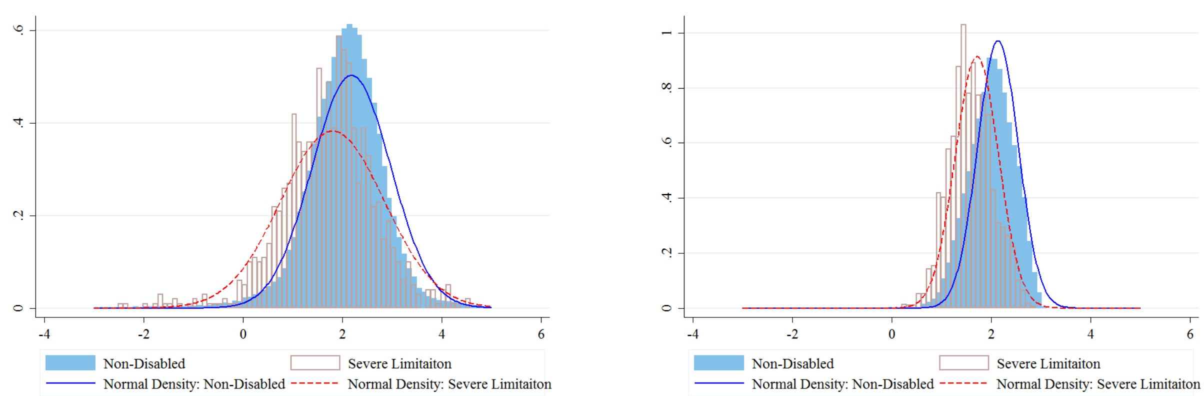
TABLE 15: ACCEPTED WAGE VS. WAGE OFFER DISTRIBUTIONS

	High School			College		
	Non-Disabled	Moderate	Severe	Non-Disabled	Moderate	Severe
Average wage of the employed	1.864	1.696	1.588	2.282	2.099	2.012
	(0.698)	(0.834)	(0.993)	(0.772)	(0.790)	(0.841)
Estimated mean wage	1.835	1.570	1.486	2.286	2.096	1.992
	(0.334)	(0.352)	(0.321)	(0.309)	(0.302)	(0.276)
Ratio between offer-to-accepted wages (%)	97.1	88.2	90.3	100	99.7	98.0

	Male			Female		
	Non-Disabled	Moderate	Severe	Non-Disabled	Moderate	Severe
Average wage of the employed	2.286	2.082	1.918	1.943	1.781	1.634
	(0.762)	(0.850)	(0.866)	(0.742)	(0.787)	(0.860)
Estimated mean wage	2.306	2.067	1.837	1.934	1.681	1.549
	(0.321)	(0.357)	(0.362)	(0.356)	(0.386)	(0.366)
Ratio between offer-to-accepted wages (%)	102.0	98.5	92.2	99.1	90.5	91.9

Note: Table 15 reports the mean and standard variation of log hourly wage of the employed workers. Based on the PSID, the hourly wage variable is computed by dividing the total labor income variable with the total annual working hours. We use sample periods from 1984 to 2011, and converted wage variables into 2011 US dollars using the CPI. Observations include both man and woman with any education level in working-age between 25 and 65, and considered being employed if their reported more than 500 annual working hours. College graduates are those who received more than 12 years of education. Estimated wage offer statistics are constructed based on the coefficients from the wage equation with Heckman's two-stage linear regression. Both mean and standard deviations are weighted using the individual weights.

FIGURE 17: DATA VS. SELECTION-CORRECTED WAGE DISTRIBUTION



Note: Figure compares the wage distribution of individuals to the estimated wage offer distribution using the Heckman's two-stage correction.

B.3 Nonlinear Wage Equation Estimation

The log-wage equation of our benchmark analysis is given by

$$\ln w_{ijt} = \ln R_{L_t} + \ln \lambda_L (j, x_{it}, h_j) + \ln \left(1 + \Pi_{E_t} \frac{\lambda_E (j, x_{it}, h_j)}{\lambda_L (j, x_{it}, h_j)} g(e_{it})\right) + \ln z_{it} + \epsilon_{ijt}$$

where ϵ_{ijt} is classic measurement error.

A Two-Period Example: Illustration of Our Identification Mechanism We simplify our estimation equation and present our identification mechanism using examples. We assume

$$\phi_L^U (j) = \exp\left(\alpha_0^U + \alpha_1^U j\right) \quad \text{and} \quad \phi_E^U (j) = \exp\left(\beta_0^U + \beta_1^U j\right).$$

If we model the endowment profile $\lambda_s (j) = \lambda_s + \lambda_s j$ and $g(0) \equiv 0$, then the nonlinear wage equation can be further simplified to

$$\begin{aligned} \ln w_j^h (e) &= \ln R_L + \left\{\alpha_0^h + \alpha_1^h j\right\} + \{\lambda_{L,0} + \lambda_{L,1} j\} \\ &+ \ln \left[1 + \Pi_E \exp\left(\sum_j \left(\underbrace{\beta_0^h - \alpha_0^h}_{\equiv \delta^h}\right) j + \sum_j \lambda_{E/L} (j)\right) g(e)\right]. \end{aligned} \tag{7}$$

Table 16 shows that by comparing the difference in wages of workers by health status, we can identify the coefficients for the impact of health on the labor, $\left\{\alpha_i^U\right\}$. We can further proceed our example by solving the wages of workers with one year of experience at $j = 1$ to illustrate identification of $\left\{\beta_i^U\right\}$.

TABLE 16: IDENTIFICATION OF THE LABOR PROFILES: AN EXAMPLE

Disability	Characteristics (age, experience)	Log-Wage
Healthy	$(0,0)$	$\ln R_L + \lambda_{L0}$
	$(1,0)$	$\ln R_L + \lambda_{L0} + \lambda_{L1}$
Unhealthy	$(0,0)$	$\ln R_L + \alpha_0^U + \lambda_{L0}$
	$(1,0)$	$\ln R_L + \alpha_0^U + \alpha_1^U + \lambda_{L0} + \lambda_{L1}$

Impact of disability on wage $\left(\ln w^U - \ln w^H\right)$			
(age, experience)		Healthy	
		$(0,0)$	$(1,0)$
Unhealthy	$(0,0)$	α_0^U	$\alpha_0^U - \lambda_{L1}$
	$(1,0)$	$\alpha_0^U + \alpha_1^U + \lambda_{L1}$	$\alpha_0^U + \alpha_1^U$

TABLE 17: IDENTIFICATION OF THE EXPERIENCE PROFILES: AN EXAMPLE

Disability	Characteristics (age, experience)	Log-Wage	Value of experience on wage $\left(\ln w^h_{(1,1)} - \ln w^h_{(1,0)}\right)$
Healthy	$(1,0)$	$\ln R_L + \lambda_{L0} + \lambda_{L1}$	$\ln\left(1 + \Pi_E\left(\lambda_{E/L,0} + \lambda_{E/L,1}\right) g(1)\right)$
	$(1,1)$	$\ln R_L + \lambda_{L0} + \lambda_{L1}$ $+ \ln\left(1 + \Pi_E\left(\lambda_{E/L,0} + \lambda_{E/L,1}\right) g(1)\right)$	
Unhealthy	$(1,0)$	$\ln R_L + \alpha^U_0 + \alpha^U_1 + \lambda_{L0} + \lambda_{L1}$	$\ln\left(1 + \Pi_E\left(\delta_0 + \delta_1 + \lambda_{E/L,0} + \lambda_{E/L,1}\right) g(1)\right)$
	$(1,1)$	$\ln R_L + \alpha^U_0 + \alpha^U_1 + \lambda_{L0} + \lambda_{L1}$ $+ \ln\left(1 + \Pi_E\left(\delta_0 + \delta_1 + \lambda_{E/L,0} + \lambda_{E/L,1}\right) g(1)\right)$	

In general, the health premium within the same group of characteristics comes from two components:

$$\begin{aligned}
& \ln w^H_{1t} - \ln w^U_{2t} \\
= \; & \ln\left(\lambda_L(j) + \Pi_t \lambda_E(j)\, g(e)\right) - \ln\left(\phi^U_L \lambda_L(j) + \Pi_t \lambda_E(j)\, \phi^U_E g(e)\right) \\
= \; & \ln\left(\lambda_L(j) + \Pi_t \lambda_E(j)\, g(e)\right) - \ln\left(\phi^U_L \left\{\lambda_L(j) + \Pi_t \lambda_E(j)\left(\frac{\phi^U_E}{\phi^U_L}\right) g(e)\right\}\right) \\
= \; & \ln\left(\lambda_L(j) + \Pi_t \lambda_E(j)\, g(e)\right) - \ln \phi^U_L - \ln\left(\lambda_L(j) + \Pi_t \lambda_E(j)\left(\frac{\phi^U_E}{\phi^U_L}\right) g(e)\right)
\end{aligned}$$

The constant term in the wage differential equation will identify the effect on the labor efficiency schedule. The relative impact of health on experience compared to labor can be estimated using nonlinear MLE.

Estimation Results Table 18 reports the effects of health on labor and experience efficiency profiles based on (1). The remaining parameter estimation results are reported in Table 19.

TABLE 18: EFFECTS OF HEALTH ON LABOR AND EXPERIENCE: ROBUSTNESS ANALYSIS

Individual characteristics		Relative Efficiency (Non-Disabled $\equiv$ 1)	
Disability Status	Education	Labor	Experience
Disabled	high school	0.709	0.837
		[0.57, 0.87]	[0.58, 1.09]
	college	0.575	0.917
		[0.48, 0.68]	[0.70, 1.13]
Number of obs.		56,840	
R^2		0.238	

Note: Table 18 reports the estimated $\hat{\phi^{xy}}$, which represents the relative shifts of λ_E and λ_L by health realizations compared to those who remain healthy. Other regressors include individual characteristics such as age, experience, marital status, as well as time-varying year dummies, male dummies, race dummies, and annual dummies for return from school. Standard errors clustered at individual level and reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

TABLE 19: COEFFICIENT ESTIMATES OF THE NONLINEAR WAGE EQUATION

Parameter description			Coefficient
labor endowment profile	High school	$\lambda^H_{L,2}$	-0.008
($\ln \lambda^s_L = \sum_{n=1}^{3} \lambda_{L,n} j^{n-1}$			(0.013)
with $s = H, C$)		$\lambda^H_{L,3}$	0.0002
			(0.0002)
	College	$\lambda^C_{L,1}$	0.092
			(0.145)
		$\lambda^C_{L,2}$	0.0003
			(0.009)
		$\lambda^C_{L,3}$	-0.00009
			(0.0002)
experience endowment profile	High school	$\lambda^H_{L,2}$	-0.016
($\ln \lambda^s_E = \sum_{n=1}^{3} \lambda_{E,n} j^{n-1}$			(0.012)
with $s = H, C$)		$\lambda^H_{L,3}$	-0.00009
			(0.0002)
	College	$\lambda^C_{L,1}$	-0.597***
			(0.150)
		$\lambda^C_{L,2}$	0.018*
			(0.010)
		$\lambda^C_{L,3}$	-0.0007***
			(0.0002)
accumulation of experience		θ_2	-0.016*
($g(e) = e + \sum_{n=2}^{4} \theta_n e^n$)			(0.008)
		θ_3	0.00002
			(0.0004)
		θ_4	3.26e-06
			(5.31e-06)
Number of obs.		56,840	
R^2		0.238	

Note: Table 19 reports the estimated endowment profiles, λ_E and λ_L, by education group. Other regressors include individual characteristics, such as regional dummies and marital status, as well as time-varying year dummies, male dummies, race dummies, and annual dummies for return from school. Standard errors clustered at individual level and reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Finally, Figures summarize the changes in time-varying dummies for gender, race, and college education.

FIGURE 18: ESTIMATION RESULTS: TIME-VARYING DUMMIES

(A): GENDER DUMMIES

(B): RACE DUMMIES

(C): EDUCATION DUMMIES

Note:

B.4 A Model Comparison with Other Specifications

A Model Comparison with Card and Lemieux (2001) In Card and Lemieux (2001), the aggregate production function takes separate factors by education group as its inputs:

$$Y = (H^\rho + \theta C^\rho)^{\frac{1}{\rho}}$$

where H and C represent the total amount of labor supplied by the high school and college graduates, respectively. They assumed the total amount of labor H and C would be characterized by its own education-specific CES production function where each age-group functions its designated role:

$$H = \left(\sum_j \alpha_j H_j^\rho\right)^{\frac{1}{\rho}} \quad \text{and} \quad C = \left(\sum_j \alpha_j C_j^\rho\right)^{\frac{1}{\rho}}.$$

Instead of assuming that an individual in age j exclusively supplies age-specific factor, we follow the approach of Jeong, Kim and Manovskii (2015) and consider the case where individuals take his labor (l) and experience (e). Therefore, the aggregate amount of H and C through the lens of Card and Lemieux (2001) would be equivalent to the aggregation of these individual factor supplies:

$$\begin{aligned} H &= F^H\left(\sum_i l_i, \sum_i e_i\right) \\ C &= F^C\left(\sum_i l_i, \sum_i e_i\right), \end{aligned}$$

where F^X is a education-group-level production function which translates labor and experience of the economy with same education to factor X. Suppose the production function of education production function satisfies the Euler theorem with homogeneous of degree 1. In this case,

$$\begin{aligned}
X &= F^X\left(\sum_i l_i, \sum_i e_i\right) \\
&= F^X_L\left(\sum_i l_i, \sum_i e_i\right) \times \sum_i l_i + F^X_E\left(\sum_i l_i, \sum_i e_i\right) \times \sum_i e_i \\
&\equiv R^X_L \times \sum_i l_i + R^X_E \times \sum_i e_i
\end{aligned}$$

where R^X_L and R^X_E represent the first-order derivative with respect to labor and experience for education group x. Using this expression, the aggregate production function can be written as

$$Y = \left[\left\{R^H_L L^H + R^H_E E^H\right\}^\rho + \theta\left\{R^C_L L^C + R^C_E E^C\right\}^\rho\right]^{\frac{1}{\rho}}$$

where $L^X \equiv \sum_i l_i$ and $E^X \equiv \sum_i e_i$ in individuals in education group. In the labor market, individual-level labor market income will be determined by the marginal productivity of his factor supplies:

$$\begin{aligned}
\frac{\partial Y}{\partial H} \times \frac{\partial H}{\partial L} &= \Psi H^{\rho-1} \times R^H_L = P_H \times R^H_L \\
\frac{\partial Y}{\partial H} \times \frac{\partial H}{\partial E} &= \Psi H^{\rho-1} \times R^H_E = P_H \times R^H_E \\
\frac{\partial Y}{\partial C} \times \frac{\partial C}{\partial L} &= \theta\Psi C^{\rho-1} \times R^C_L = P_C \times R^C_L \\
\frac{\partial Y}{\partial C} \times \frac{\partial C}{\partial E} &= \theta\Psi C^{\rho-1} \times R^C_E = P_C \times R^C_E
\end{aligned}$$

where $\Psi \equiv (H^\rho + \theta C^\rho)^{1/\rho-1}$, $P_H \equiv \Psi H^{\rho-1}$, and $P_C \equiv \theta\Psi C^{\rho-1}$. Thus, the total labor market earning of a person i in education group X is

$$w^X_i = P_X\left\{R^X_L l_i + R^X_E e_i\right\} = P_X R^X_L\left\{l_i + \Pi^X e_i\right\}$$

where Π^X represents the relative return of experience compared to labor as described in the main text. If the production function of each education is also CES as in Card and Lemieux (2001),

$$H = \left(L^{\mu_H} + \theta^H E^{\mu_H}\right) \quad \text{and} \quad C = \left(L^{\mu_C} + \theta^C E^{\mu_C}\right)^{\frac{1}{\mu_C}}.$$

Under this assumption, the return of labor and experience can be written as

$$\begin{aligned} R_L^H &= \left(L^{\mu_H} + \theta^H E^{\mu_H}\right)^{\frac{1}{\mu_H}-1} L^{\mu_H-1} \\ R_E^H &= \theta^H \left(L^{\mu_H} + \theta^H E^{\mu_H}\right)^{\frac{1}{\mu_H}-1} E^{\mu_H-1} \\ R_L^C &= \left(L^{\mu_C} + \theta^C E^{\mu_C}\right)^{\frac{1}{\mu_C}-1} L^{\mu_C-1} \\ R_E^C &= \theta^C \left(L^{\mu_C} + \theta^C E^{\mu_C}\right)^{\frac{1}{\mu_C}-1} E^{\mu_C-1}, \end{aligned}$$

and $\Pi^X = \theta^X \left(\frac{E^X}{L^X}\right)^{\mu_X-1}$. If the average endowment of labor and experience of age j and education x as $l_X(j)$ and $e_X(j)$, then the college premium equation can be written as

$$\begin{aligned} \ln\left(\frac{w_j^H}{w_j^C}\right) &= \ln\left(\frac{P_H\left\{R_L^H l_H(j) + R_E^H e_H(j)\right\}}{P_C\left\{R_L^C l_C(j) + R_E^C e_C(j)\right\}}\right) \\ &= -\ln\theta + \left(\frac{\rho}{\mu_H}-1\right)\ln\left(L_H^{\mu_H} + \theta^H E_H^{\mu_H}\right) - \left(\frac{\rho}{\mu_C}-1\right)\ln\left(L_C^{\mu_C} + \theta^C E_C^{\mu_C}\right) \\ &\quad + (\mu_H-1)\ln L_H - (\mu_C-1)\ln L_C + \ln\left(l_H(j) + \theta^H\left(\frac{E_H}{L_H}\right)^{\mu_H-1} e_H(j)\right) \\ &\quad - \ln\left(l_C(j) + \theta^C\left(\frac{E_C}{L_C}\right)^{\mu_C-1} e_C(j)\right) \end{aligned}$$

Therefore, the analysis of Card and Lemieux (2001) can be considered as a special case where each group provides age-specific factor. For a two-period example with a young and old workers, the above equation becomes

$$\begin{aligned} \ln\left(\frac{w_Y^H}{w_Y^C}\right) &= -\ln\theta + \left(\frac{\rho}{\mu_H}-1\right)\ln\left(L_H^{\mu_H} + \theta^H E_H^{\mu_H}\right) - \left(\frac{\rho}{\mu_C}-1\right)\ln\left(L_C^{\mu_C} + \theta^C E_C^{\mu_C}\right) \\ &\quad + (\mu_H-1)\ln L_H - (\mu_C-1)\ln L_C + \ln(l_H) - \ln(l_C) \\ \ln\left(\frac{w_O^H}{w_O^C}\right) &= -\ln\theta + \left(\frac{\rho}{\mu_H}-1\right)\ln\left(L_H^{\mu_H} + \theta^H E_H^{\mu_H}\right) - \left(\frac{\rho}{\mu_C}-1\right)\ln\left(L_C^{\mu_C} + \theta^C E_C^{\mu_C}\right) \\ &\quad + (\mu_H-1)\ln L_H - (\mu_C-1)\ln L_C \\ &\quad + \ln\left(\theta^H\left(\frac{E_H}{L_H}\right)^{\mu_H-1} e_H\right) - \ln\left(\theta^C\left(\frac{E_C}{L_C}\right)^{\mu_C-1} e_C\right) \end{aligned}$$

When a worker experiences health deterioration, we allow their labor and experience endowments can vary by δ_L^X.

A Model Comparison with Low and Pistaferri (2015) Low and Pistaferri (2015) is one of the recent papers evaluating the health risks on labor market income risks. Using the labor income growth rate as their measure, which is defined as the compound of both hours and productivity, they find substantial decline in income growth is associated with the onset of disability among the workers with high school education males. In this section, we provide interpretation of their wage equation estimation in our model framework and compare their findings with our estimation results. The log-wage equation of Low and Pistaferri (2015) is given as

$$\ln w_{it} \quad = \quad X_{it}'\beta + \varphi L_{it}^{j} + f_i + \epsilon_{it} + \omega_{it}.$$

The coefficient φ captures the change in income when a worker' health status changes to j: $L_{it}^{j} = 1$. Along with the effect φ, there is another mechanism that health influence worker's labor income in their analysis: individual's fixed-type (f_i) affects the probability that he experiences certain health status, $\Pr\left(L_{it} = j \middle| L_{it-1} = k, f_i\right)$. They introduce productivity shock which is independent from health status $\epsilon_{it} = \epsilon_{it-1} + \zeta_{it}$ and standard measurement error ω_{it}. The value of the unobserved heterogeneity $\hat{f}_i$ is pinned once they estimate key coefficients based on the differences in wage equations:

$$\Delta \ln w_{it} \quad = \quad \Delta\left(X_{it}'\beta + \varphi L_{it}^{j} + \epsilon_{it} + \omega_{it}\right),$$

and $\hat{g}_{it} = \Delta\left(\ln w_{it} - X_{it}'\hat{\beta} - \hat{\varphi} L_{it}^{j}\right)$. The series of residuals, $\hat{g}_{it}$ allows them to estimate the variance of productivity shock (σ_{ζ}^{2}) as $g_{it} = \zeta_{it} + \Delta\omega_{it}$.

In our model representation, individual's hourly wage is a function of (i) his idiosyncratic component; (ii) the number of years of experience and his efficiency in translating his experience into human capital; and (iii) the current labor endowment and his efficiency in translating it into human capital:

$$\ln w_{it} \quad = \quad X_{it}'\beta + \ln f\left(h_{it}\right) + \ln g\left(e_{it}, h_{it}\right) + \nu_{it},$$

where ν_{it} is classic measurement error. When a worker experience a chance in health status, the total effects in productivity is

$$\Delta\left(\ln f\left(h_{it}^{j}\right) + \ln g\left(e, h_{it}^{j}\right)\right) + \Delta\nu_{it} \quad = \quad \varphi + \zeta + \Delta\omega.$$

Returning to the wage level-estimation, we can write the labor productivity of an individual with health status L_{it}^{j} in two different ways:

$$\sum_{k=0}^{t-1} \zeta_{i,k} + \varphi_{it}^{j} + f_i \quad = \quad \ln f\left(h_{it}^{j}; \phi_L^{j}\right) + \ln g\left(e_{it}, h_{it}^{j}; \phi_E^{j}\right)$$

On the left hand side, his labor productivity is summarized by the accumulation of health-independent shocks up to $t-1$ periods and his newly realized health shock φ_{it}^{j}. On the right hand side, we have nonlinear equation as a function of two coefficients, ϕ_L^{j} and ϕ_E^{j}. While we do not introduce unobserved ex-ante heterogeneity in our model, we do explicitly consider the years of experience as additional source of information in estimation process. By doing so, our model can generate certain subset of workers exhibiting higher productivity through out their working life compared to their less fortunate counterparts endogenously.

C Computational Algorithms

TBA