Understanding the Aggregate Effects of Disability Insurance*

Soojin Kim† Serena Rhee‡

April 23, 2020

Abstract

We study the aggregate consequences of the Social Security Disability Insurance (DI) program, focusing on the role of complementarity between heterogeneous human capital. First, we develop and estimate a wage process in which individuals' human capital is composed of (pure) labor and experience, and their efficiencies are affected by disability. We find that older workers are more experience-abundant, and that disability causes a smaller loss in the efficiency of experience than it does in the efficiency of labor. Further, the estimated aggregate production technology shows that labor and experience are complementary inputs. Combining these empirical results with a structural general equilibrium model, we analyze the labor market implications of removing the DI program. Removal of the DI program induces an increase in the relative supply of experience, thus affecting the marginal products of inputs and wages of all workers in the economy. While the entry of less productive workers lowers the average productivity of the workforce, its negative effect is limited because of the complementarities between labor and experience.

JEL Codes: J31, J24, E24, I18, I38
Keywords: disability insurance, labor supply, wage risk, skill complementarity, human capital

---

* We thank Rasmus Lentz, José-Víctor Ríos-Rull, Minchul Yum, and seminar and conference participants at numerous institutions for their valuable comments and discussions. We are grateful to Iourii Manovskii for helping us generate the labor market experience variables using the Panel Study of Income Dynamics. This paper was written in part while Rhee was visiting the RAND Corporation and the USC Schaeffer Center. Rhee is grateful to both organizations for their hospitality during her visit. Rhee's research was supported by the National Institutes of Health under award number 2T32AG000244. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the National Institutes of Health. Any errors are our own.

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

‡ School of Economics, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06724, Republic of Korea, srhee@cau.ac.kr.

1 Introduction

In 2018, approximately 10 million people in the United States benefited from the Social Security Disability Insurance (DI) program (Social Security Administration, 2017), and its growth is accelerating with the aging of the population.[1] While DI serves as an important safety net against health risks, recent empirical studies (e.g, Maestas et al., 2013 and French and Song, 2014) have found that it suffers from significant disincentive effects. Given the large scale of the DI program, it is important to understand the aggregate implications of the the labor supply effects of DI. In this paper, we study this question by evaluating the individual-level effects of disability on workers' productivities, and combining these micro-level results with the cross-sectional distribution of workers to measure the aggregate implications of DI.

To assess the DI program, we need to know the productivity of disabled workers, and how the loss of these workers impacts the labor market and aggregate production. Thus, we first estimate the productivity effects of disability on workers. Following the seminal paper by Katz and Murphy (1992) and expanding the work of Jeong et al. (2015), we assume that workers are endowed with two inputs: "(pure) labor" and "experience." Using the detailed micro-level data, we separately quantify how detrimental a disability is on the efficiencies of labor and experience, thereby identifying the sources of the low productivity (wage) of disabled workers. Then, we further exploit the time-series variations in the relative price and quantity of labor and experience to measure the substitutability between the two inputs in the aggregate production. The modeling of these heterogeneous inputs helps measure not only the direct productivity loss but also the potential aggregate efficiency consequences from losing workers due to the DI program. Lastly, we use a general equilibrium life-cycle model of workers to evaluate the aggregate labor market effects of DI within our heterogeneous input model, and to measure the value of the DI program for workers.

The micro-level estimation uses data from the Panel Study of Income Dynamics (PSID), which contains work history (years of experience) and health status information. Using the hourly wage rate as a measure of productivity, we estimate the amounts of efficiency units of labor and experience, and how much the efficiency of these inputs are affected by an individual's disability.[2] We find that having a disability lowers a worker's efficiency units of labor by 27% (37%) and his efficiency units of experience by 13% (2%) for high school graduates (college graduates). These findings, in conjunction with the fact that the experience is the primary source of human capital for older workers, suggest that the amount of experience lost from the

[1] In 2015, the total benefit payments for DI and Medicare for qualified beneficiaries exceeded \$220 billion (5.8% of the federal budget). Under the current system, the Congressional Budget Office (CBO) projects that the trust fund for the DI program will be exhausted in 2022 as the size of beneficiaries is expected to grow by 0.8% a year (Congressional Budget Office, 2016).

[2] Since wage variables are only available for the employed, we control for selection bias following Heckman's two-step procedure using the generosity of welfare programs and tax credits of the local government as instrumental variables.

reduced labor market participation of older workers, the majority of DI recipients, might be significant. Further, if these inputs are imperfectly substitutable in aggregate production, the changes in the relative supply of inputs can cause indirect effects through equilibrium factor prices and labor productivity. To capture the aggregate effect, we estimate the elasticity of substitution between the inputs, assuming a constant elasticity of substitution (CES) production function. Exploiting the time-series variations in the relative supply and price of labor and experience, we find that labor and experience are gross complements with the elasticity of substitution of 0.41.

We incorporate these empirical findings into a general equilibrium model, which we use to quantify the aggregate impacts of the DI program. In the model, finitely-lived households are subject to health (disability and mortality), medical expenditures, and idiosyncratic productivity risks. These workers make the endogenous labor supply and saving decisions, and, if disabled, are allowed to apply to the DI program. Importantly, we model the key features of the DI program including the application processes and the risks (e.g., acceptance, termination) associated with the policy. We use the estimated wage processes and aggregate production technologies, and calibrate the model to match the life-cycle statistics of worker outcomes by health statuses. We further validate that the calibrated model is able to replicate the characteristics of DI recipients and the empirical estimates of labor supply elasticities.

Finally, we evaluate the impact of DI on aggregate outcomes. In the calibrated economy, the removal of the DI program increases the work incentives of all workers, thereby increasing the aggregate supplies of labor and experience, and output. However, as the employment rates of the experience-abundant old workers (who used to be DI recipients) increase disproportionately, the relative supply of experience rises. This in turn heterogeneously impacts the wages of workers over the life cycle. In particular, young workers benefit because of the increased price of labor, while old workers benefit because of their higher amount of experience despite its price being lower than in the benchmark economy. To understand the role of the input complementarity, we conduct the same counterfactual analysis in a recalibrated economy where labor and experience are perfect substitutes. We find that accounting for the complementarity between inputs is important for gauging the productivity effects of removing the DI program. Thanks to the complementarity between young (labor) and old (experience) workers, the entry of old workers, which is induced by the removal of DI, leads to a relatively smaller productivity loss compared to the economy where inputs are perfect substitutes. Lastly, we measure the value of DI to workers, which accounts for both the insurance benefits of the program and its labor market effects through changes in worker productivities (wages). The value of the DI program varies widely across worker characteristics, with higher valuations from the old, poor, and disabled workers.

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 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 young (those between 20 and 24 in age) and prime-aged (between 25 and 54) workers drop by 0.492 and 0.258 points when elderly (between 55 and 64) employment increase by 1 percentage point (*pp*), 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 a "crowd-out" effect of young workers when the employment of old workers increases. Instead, they find that one percent increase employment rate of the old leads to an increased employment rate of the prime-aged workers by 0.34*pp*. These are empirical evidence of complementary between young and old workers, consistent with our finding that labor and experience (implicitly, young and old workers) are complementary in production.

Relatedly, similar to the previous literature (e.g., Card and Lemieux, 2001; Krusell et al., 2000; and Karabarbounis and Neiman, 2014), we estimate the degree of substitutability across heterogeneous inputs in production using empirical data, assuming a Constant Elasticity of Substitution (CES) production function.[3] In terms of methodology, we are most closely related to Jeong et al. (2015), which estimates the amount 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 the labor and experience over the life cycle with consideration of potential selection bias.

Secondly, this paper builds on and expands the studies of the labor supply disincentive effects of DI, which has long been a topic of interest, starting with Bound (1989). Maestas et al. (2013) and French and Song (2014) use random assignments of disability examiners and judges to estimate the disincentive effects of disability insurance on the labor supply of workers. Both papers find a strong disincentive effect of disability insurance.[4] While these papers use an econometric approach to study individual behavior,

[3] 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 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.

[4] Maestas et al. (2013) shows that for marginal applicants, the employment would have been 28 *pp* higher in the absence of the DI program, using the data on behaviors of rejected DI applicants. Similarly, French and Song (2014) also finds that benefit receipt reduced participation by 26 *pp* three years after the decision. While the focus is different, Low and Pistaferri (2019) uses administrative data to explore broader aspects of the institutional features of the DI program, and finds systematically higher false rejection rates against female applicants during the screening process.

Kitao (2014), Low and Pistaferri (2015), and Autor et al. (2019) are among the few that develop life-cycle models to analyze the effects of DI. Kitao (2014) focuses on the interaction between DI and unemployment insurance, while Low and Pistaferri (2015) focuses on the incentive and insurance trade-off that individual workers face. Autor et al. (2019) evaluates welfare effects of DI by explicitly incorporating household structures and finds that spousal labor supply serves as an important insurance against disability. This paper is distinct from theirs in two dimensions. First, while most analyses on disability assume that a worker's human capital is one-dimensional, we explicitly model and estimate the effects on disability on heterogeneous human capital endowments of workers. Thus, our analysis provides an understanding of the sources of the productivity losses disabled workers face, and how these effects might differ over the life cycle of workers. Secondly, we further use these micro-level findings and incorporate the interactions between inputs in aggregate production to measure the aggregate output and productivity effects of the DI program.

Finally, this paper contributes to the broad literature analyzing the effects of social insurance policies, especially with respect to health or medical expense risks (e.g., Hubbard et al., 1995; Attanasio et al., 2011; Pashchenko and Porapakkarm, 2017; De Nardi et al., 2018). Some of the recent papers in the literature include De Nardi et al. (2016) and Braun et al. (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 aggregate implications in a general equilibrium model.

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 discuss the calibration of the model, and use it to conduct counterfactual analyses in Section 5. We conclude in Section 6.

2 Empirical Analysis

In this section, we estimate the role of disability on workers' productivities and use the estimates to measure the parameters of the aggregate production function.

2.1 Wage Equation

We consider workers who provide two distinctive inputs—(pure) labor and experience—and empirically examine the relationship between health and factor productivities. Labor is physical effort or abilities, and experience represents human capital accumulated from participating in labor markets. These concepts correspond to "pure labor" and "pure experience" in the seminal paper by Katz and Murphy (1992), where they model them as separate inputs exclusively supplied by "young" and "old."

Our wage equation incorporates the health effects on wages, extending those of Jeong et al. (2015), who generalized Katz and Murphy (1992) to allow all workers to supply a bundle of labor and experience. Through the lens of Jeong et al. (2015), the hourly wage rate (or productivity) of a worker is determined by endowed human capital–labor $\hat{l}$ and experience $\hat{e}$–and their prices R_L and R_E: $w = R_L\hat{l} + R_E\hat{e}$.

We denote the endowed units of labor ($\hat{l}$) for an individual with age j and health h as $\lambda_L(j,h)$. Unlike this deterministic life-cycle profile of labor, the amount of experience ($\hat{e}$) can vary within the same demographics as workers may have different employment histories over time. Therefore, we consider that the total amount of experience (in efficiency units) is a product of both the deterministic component $\lambda_E(j,h)$ and a function of a worker's endogenously accumulated work experience $g(e)$.[5] Using these notations, the hourly wage rate of a worker with age j and health status h can be rewritten as

$$w(j,h,e) = R_L\lambda_L(j,h) + R_E\lambda_E(j,h)\,g(e), \tag{1}$$

where e is the actual years of work in the labor market.

For the purposes of implementing the estimation of Equation (1), we follow the functional form choices directly from Jeong et al. (2015) and use polynomials, which are shown to be superior to alternative specifications in fitting the wage data. Specifically, an individual's accumulated experience is determined by $g(e) = e + \zeta_1 e^2 + \zeta_2 e^3 + \zeta_3 e^4$, allowing for possible non-linear effects of work experience. The deterministic components of labor and experience are approximated by polynomial functions of age j: $\tilde{\lambda}_X(j) = \exp\left(\lambda_{X,0} + \lambda_{X,1}j + \lambda_{X,2}j^2\right)$ with $X = L$ and E. Further, we incorporate health effects on labor and experience profiles by including a scaling factor $\phi_X(h)$, so that $\lambda_X(j,h) = \phi_X(h)\tilde{\lambda}_X(j)$ for $X = L, E$. Thus, in our implementation, health proportionately affects the factor profiles. Given the functional form assumptions, the relative efficiency of experience compared to labor is given as $\tilde{\lambda}_E(j)/\tilde{\lambda}_L(j) = \exp\left(\lambda_{E/L,0} + \lambda_{E/L,1}j + \lambda_{E/L,2}j^2\right)$, where $\lambda_{E/L,k} \equiv \lambda_{E,k} - \lambda_{L,k}$ for $k \in \{0,1,2\}$.

[5] We can interpret that the life-cycle profile $\lambda_E(j,h)$ represents how effectively an individual uses his experience, which depends on his age and health status. The quantity of accumulated experience is captured by the term $g(e)$, a function of the worker's actual years of work, e. We denote the product of the two components as the *efficiency* units of experience.

We allow the coefficients for deterministic components to depend on education and health status. We categorize workers into two education groups—$s_{it} \in \{HS, Col\}$, high school (HS) graduate or less and some college or more (Col)—and two health groups—$h_{it} \in \{ND, D\}$, those with a work-limiting disability (D) and the rest non-disabled (ND). Thus, the coefficients are denoted as $\lambda_{X,k}(s_{it})$ and $\phi_X(s_{it}, h_{it})$ for $X \in \{L, E\}$ and $k \in \{0, 1, 2\}$. Rewriting Equation (1) as $w(j, h, e) = R_L \lambda_L(j, h)\left[1 + \frac{R_E}{R_L} \cdot \frac{\lambda_E(j,h)}{\lambda_L(j,h)} g(e)\right]$, replacing $\ln R_{L_t}$ with a year-dummy variable d_t, and denoting relative price of experience as $\Pi_{E_t} \equiv R_{E_t}/R_{L_t}$, the log-wage equation for estimation is expressed as

$$\begin{aligned} \ln w_{it} \quad = \quad & d_t + \ln \phi_L(s_{it}, h_{it}) + \left\{\lambda_{L,0}(s_{it}) + \lambda_{L,1}(s_{it}) j_{it} + \lambda_{L,2}(s_{it}) j_{it}^2\right\} \\ & + \ln\left[1 + \Pi_{E_t} \frac{\phi_E(s_{it}, h_{it})}{\phi_L(s_{it}, h_{it})} \exp\left(\lambda_{E/L,0}(s_{it}) + \lambda_{E/L,1}(s_{it}) j_{it} + \lambda_{E/L,2}(s_{it}) j_{it}^2\right)\right. \\ & \left.\times\left(e_{it} + \zeta_1 e_{it}^2 + \zeta_2 e_{it}^3 + \zeta_3 e_{it}^4\right)\right] + \boldsymbol{\beta} \boldsymbol{X}_{it} + \varepsilon_{it}. \end{aligned} \tag{2}$$

We control for individual-level characteristics through the regressor $\boldsymbol{X}_{it}$, which includes region, and time-specific dummies for college degree, gender, and race. The classical measurement error is denoted as ε_{it}. We normalize $\lambda_{L,0}(HS) = \lambda_{E,0}(HS) = 1$ and $\phi_L(s_{it}, ND) = \phi_E(s_{it}, ND) = 1$. Thus, the coefficient $\phi_X(s_{it}, D)$ reflects the relative efficiency of disabled workers compared to non-disabled workers within the same education group.[6,7]

2.2 Selection Bias and Identification Strategy

One challenge in estimating Equation (2) is that we only observe the wages of employed individuals; these workers, especially those who are participating in the labor market despite their disabilities, may systematically differ from the non-employed disabled. Therefore, the estimated effects of disability on labor and experience can be biased if we do not correct for this potential selection bias.

We address this concern by estimating the wage equation using a standard two-stage procedure described in Heckman (1979). We first estimate the underlying participation decision using a probit model with instrument variables, then estimate the wage equation including the inverse Mills' ratio from the first stage. In line with the idea of simulated IV in public economics (as in Currie and Gruber, 1996a,b and Low and Pistaferri, 2015), we exploit the spatial and time variation of public policies as our first-stage instruments.

[6] We conduct various robustness analyses regarding the specification, which are discussed in Section 2.3 and in Appendix B.2.

[7] While we estimate the impact of disability on productivity of labor and experience, we do not estimate the full wage processes that include productivity risks as does Low and Pistaferri (2015). However, when we use the wage processes for the structural model, we take as variance of the iid productivity shock, the data-implied residual variances that depend on the disability status of the worker.

Specifically, we construct the generosity measures of welfare programs and tax systems by simulating potential transfers and taxes that a "representative" earner would receive from his residential state and year. The generosity of public policies vary by state and year, thus generating heterogeneity in the labor force participation incentives of the representative earner.

Note that both the transfers and taxes are computed for a representative earner, not for each individual using their own characteristics. Indeed, having actual benefits would be inappropriate due to their endogenous relation with wages. With simulated potential transfers, we capture the effects of public policies on labor supply decisions, independent from individual characteristics. Still, to be valid, our identification strategy relies on two assumptions. First, we assume that given policy variations are not systematically related to labor market conditions, and that these potential benefits affect individuals' labor market participation decisions but not their wage rates.

For constructing transfers from welfare programs, we use the Earned Income Tax Credit, Unemployment Insurance, the Supplemental Nutrition Assistance Program, and Aid to Families with Dependent Children, which later became Temporary Assistance for Needy Families. For tax credits, we first supplement the PSID with the Survey of Consumer Finances (SCF) data, which provides rich information on individuals' financial status, such as mortgage interest payments. Including this information helps us to generate a more reliable measure of taxable income, as we can better approximate tax liabilities and credits such as mortgage deductions. Using the predicted taxable income of representative earners, we simulate their taxes using the NBER TAXSIM v.27.[8] Further details of the estimation process and transfer variable construction are documented in Appendix B.1.

2.3 Estimation Results

The First-Stage Probit Regression. In the first-stage, we estimate a probit model of labor market participation decision using the entire working-age sample in the data. The independent variables include, along with standard controls for individual characteristics, the two instrumental variables by disability status.[9] Table 1 presents the results from the first-stage probit regression. We observe that disability has a significant impact on employment probability; for a marginal worker, having a disability lowers the employment probability by 23.3 percentage points (pp), and for an average worker, the employment probability decreases by 19.8pp.

[8] See Feenberg and Coutts (1993) and http://www.nber.org/taxsim/ for more information regarding the NBER TAXSIM.

[9] Thus, we use a total of four variables to instrument for the labor supply decision estimation. To avoid computational burden, we conduct an over-identification test with the standard linear wage equation and find that the exclusion restriction holds (J-test is not rejected), as detailed in Appendix B.1.

Table 1: The First-Stage Probit Regression Results

Independent Variables	Coefficients	Effects on Probability of Employment	
		Marginal Effects	Average Effects
Disability	-0.809	-0.233	-0.198
	(0.032)	(0.009)	(0.008)
Number of Obs.		$101,414$	
Pseudo R^2		0.225	

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. Independent variables include individual characteristics (age, experience, years of schooling, male, race, marital status, state, and time-varying year dummies). We use state- and year-specific amounts of potential transfers and taxes as exclusion restrictions. We use individual-level survey weights, and standard errors clustered at the individual level are reported in parentheses. The complete list of estimated coefficients is reported in Appendix B.1.

The Role of Disability on Labor and Experience. We now estimate the nonlinear wage equation, presented in Equation (2), controlling for selection bias. Table 2 reports the estimated coefficients of the wage profile.

Table 2: Estimated Coefficients of Wage Profile

(a) Labor $\lambda_L(s)$		(b) Experience $\lambda_E(s)$		(c) Experience $g(e)$	
$\lambda_{L,1}(HS)$	0.0213 (0.0052)	$\lambda_{E,1}(HS)$	0.0048 (0.0137)	ζ_2	-0.0486 (0.0080)
$\lambda_{L,2}(HS)$	-0.0004 (0.0001)	$\lambda_{E,2}(HS)$	-0.0003 (0.0003)	ζ_3	0.0012 (0.0004)
$\lambda_{L,0}(Col)$	-0.2467 (0.0555)	$\lambda_{E,0}(Col)$	-0.3748 (0.1785)	ζ_4	-0.00001 (0.000)
$\lambda_{L,1}(Col)$	0.0526 (0.0056)	$\lambda_{E,1}(Col)$	0.0088 (0.0184)	Inverse Mills Ratio	
$\lambda_{L,2}(Col)$	-0.0010 (0.0001)	$\lambda_{E,2}(Col)$	-0.0003 (0.0004)	0.2663 (0.0891)	
(d) Health $\phi_X(s,h)$		(e) Implied Effects of Disability			
$\ln\phi_L(HS,D)$	-0.3083 (0.0838)	$\phi_L(HS,D)$	0.7347		
$\ln\phi_L(Col,D)$	-0.4614 (0.0787)	$\phi_L(Col,D)$	0.6305		
$\phi_E(HS,D)/\phi_L(HS,D)$	1.1847 (0.1941)	$\phi_E(HS,D)$	0.8704	N	83,532
$\phi_E(Col,D)/\phi_L(Col,D)$	1.5606 (0.2602)	$\phi_E(Col,D)$	0.9838	R^2	0.998

Note: Table 2 reports the coefficient estimation results of the nonlinear wage equation (2). The control variables include region and year-specific dummy variables for gender, race, and schooling (college). We use individual-level survey weights, and standard errors clustered at the individual level are reported in parentheses. N is number of observations and $R^2$is adjusted R^2. The complete list of estimated coefficients is reported in Appendix B.2.

Based on these estimates, Figure 1 illustrates the age-efficiency profiles of labor ($\lambda_L(s,j,h)$) and experience ($\lambda_E(s,j,h)$) for high school graduates by disability status.[10] As seen in Figure 1(a), the efficiency units of labor are hump-shaped over the life cycle, peaking in the mid 40s. On the other hand, their experience profile, shown in Figure 1(b), is downward-sloping, implying that one unit of experience at an early age is more valuable than it is at a later age.

While the age-efficiency profile of experience decreases over the life cycle, this does not necessarily

[10] We show the corresponding profiles for college graduates in Appendix B.2 for brevity in the main text.

mean that a worker's effective experience declines as he ages. This is because accumulated experience $g(e)$ increases along with years of employment (Figure 2(a)). When we compute the total experience $\lambda_E(s, j, h)\, g(\bar{e}_j)$ for high school graduates using the average experience for each age j from the data, the life-cycle profile of experience in efficiency units reveals an increasing trend, as shown in Figure 2(b).[11]

Figure 1: Efficiency of Human Capital over the Life cycle–High School Graduates

(a): Efficiency of Labor: $\lambda_L(HS, j, h)$ (b): Efficiency of Experience: $\lambda_E(HS, j, h)$

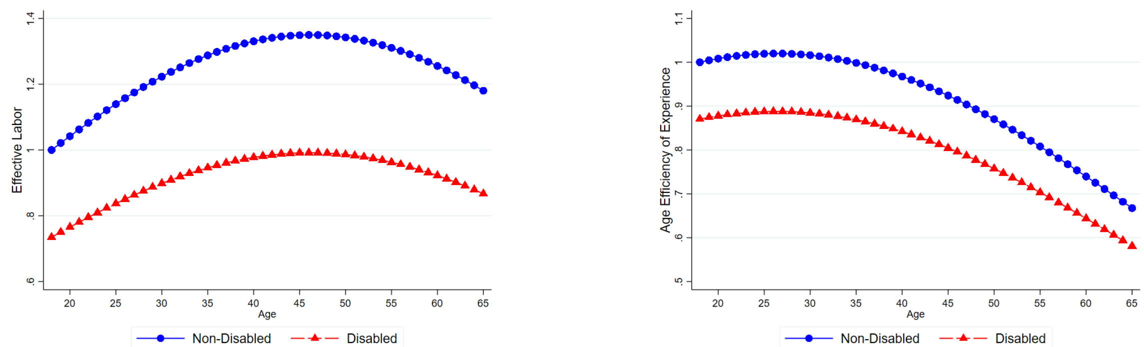
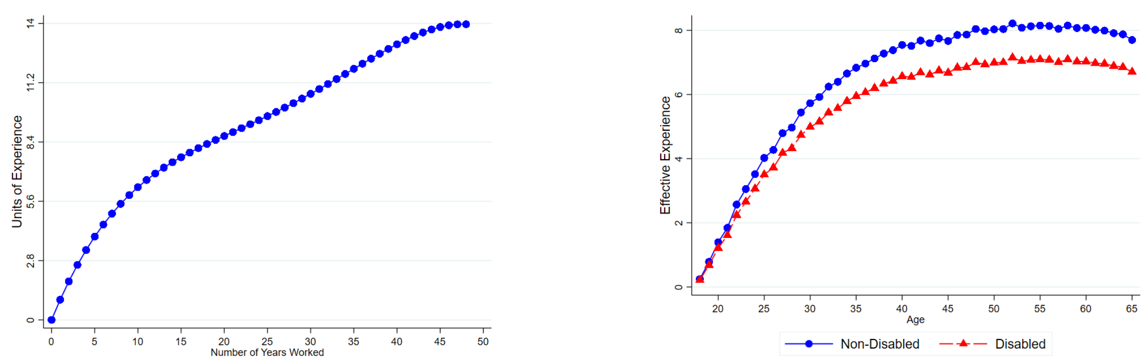


Figure 2: Experience Profiles for Workers with High School Education

(a): Accumulated Experience $g(e)$ (b): Effective Experience: $\lambda_E(HS, j, h)g(\bar{e}_j)$



Note: Figure 1(b) plots the empirical pattern of the total amount of effective experience, using the average years of experience for each age in calculating $\lambda_E(HS, j, h)\, g(\bar{e}_j)$.

As we model heterogeneous human capital, we are able to uncover the sources of productivity losses due to disability. In column (e) in Table 2, we report the implied effects of disability from the estimated coefficients. First, we note that the efficiency units of both labor and experience are lower for disabled workers. Second, we see that for workers of both education groups, disability is relatively less detrimental to efficient units of experience than it is to labor ($\phi_E > \phi_L$). For high school graduates, efficiency of labor

[11] As shown by the R^2 values, the estimated wage profiles fits the data profiles very well. The estimated wage profiles by education and health status along with their data counterparts are presented in Appendix B.2.

is 26% lower for disabled workers compared to their non-disabled counterparts, whereas the efficiency of experience of disabled workers is 13% lower than that of non-disabled workers. For college-educated workers, the significant decline in wage is driven by the loss in labor efficiency of 37%, and the efficiency decline in their experience is relatively small at 2%.

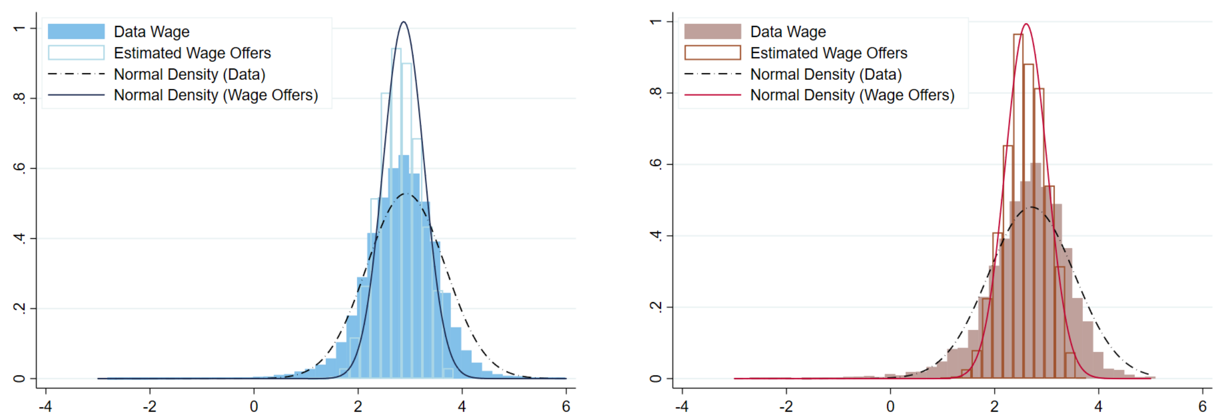
Importance of Selection Control. It is worth mentioning the importance of selection control in estimating the productivity impacts of disability. As shown in Table 2, the coefficient on the inverse Mills' ratio is significant. When we estimate the wage equation without controlling for selection, the estimated productivity impacts of disability are 0.86 (high school) and 0.73 (college) for labor, and 0.92 (high school) and 1.10 (college) for experience.

Using these estimates, we illustrate the selection bias graphically by comparing the log-wage and wage offer distributions by disability status in Figure 3, where we construct the wage offer distribution by applying the estimated coefficients on observable characteristics for working-age individuals. As seen in Figure 3, the two distributions show differences, notably more so for disabled workers. For non-disabled workers, the ratio between average wage and average wage offer is 1.05. Among disabled workers, the discrepancy between the potential wage offer and the actual wage is larger at 1.11, consistent with selection bias.

Figure 3: Selection Bias: Observed Wage vs. Estimated Wage Offers

(a): Non-Disabled Workers

(b): Disabled Workers



Note: The hourly wage rate (2011$) in the PSID data (years 1984 to 2011) is defined as total labor income divided by the annual working hours.

Robustness Analyses. Our benchmark specification allows for the education-dependence of the disability effects (ϕ_L and ϕ_E) and the age-efficiency schedules of labor and experience (λ_L and λ_E). In order to analyze the role of education in the estimation, we re-estimate the wage equation, relaxing the education-dependence. We find, as shown in Table 3, that we might have over-estimated the effects of disability on labor and under-estimated its effects on experience for high school graduates, had we not included education-dependent wage profiles. We further check whether the accumulated experience function $g(e)$ impacts the estimation outcomes once we allow its coefficients to be education-specific. We find that the estimated effects of disability are similar to the benchmark outcomes. We also test the significance of the coefficient estimation results under alternative clustering assumptions and find that the results are significant at the 1% level under various assumptions, which are presented in Appendix B.2.

Table 3: Effects of Disability with Alternative Specifications

Coefficients		(1) $\phi_X(s) = \phi_X$	(2) $\lambda_X(s) = \lambda_X$	(3) Benchmark	(4) $g(e; s)$
Labor Profile	$\phi_L(HS)$	0.6883	0.7319	0.7347	0.7372
	$\phi_L(Col)$		0.6710	0.6305	0.6336
Experience Profile	$\phi_E(HS)$	0.9351	0.9802	0.8704	0.8662
	$\phi_E(Col)$		0.9253	0.9838	0.9882
Education-Specific	ϕ_L and ϕ_E		$\times$	$\times$	$\times$
Components	λ_L and λ_E	$\times$		$\times$	$\times$
	$g(e)$				$\times$

Note: The complete list of estimated coefficients is reported in Appendix B.2.

The main findings from this wage estimation are, first, that a disability lowers the efficiency of labor and experience and, second, that the effect is larger on labor than it is on experience. Accounting for these heterogeneous effects of disability serves as an important input in understanding not only the sources of the loss in productivity of disabled workers but also the potential interaction between workers who exit the labor force and those who stay. For the latter, we now estimate the aggregate production function with labor and experience as inputs.

2.4 The Elasticity of Substitution between Labor and Experience

We use the time-series variations in total amounts of labor and experience and their price estimates to identify the aggregate production function parameters that determine the degree of complementarity between labor and experience and the relative efficiency of experience.

In the aggregate economy, a representative firm has an access to a production technology specified as $Y_t = A_t F(L_t, E_t) = A_t (L_t^\rho + \theta E_t^\rho)^{1/\rho}$.[12] This function features constant elasticity of substitution (CES)

[12] While this production function is parsimonious, it captures the interaction between heterogeneous human capital, which is the main focus of our paper. Within this production function, we take into account the worker differences by education, by allowing for education-dependent coefficients on labor and experience in individual-level wage estimation procedure. We then aggregate the individual-level labor and experience to construct L and E. However, we do not directly model the potential interaction between workers of low and high education. We could, for example, extend the production function to features CES between workers of

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 and concave 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, and the relative price of experience is $\Pi_{E,t} \equiv F_{E,t}/F_{L,t}$.

We construct the total amount of efficiency units of labor and experience based on the wage estimation results from Section 2.3, along with the estimated relative price of experience. Using the observed working hours ($\bar{\eta}_{it}$) in the PSID, we can aggregate individual-level labor and experience in efficiency units by summing up $\hat{l}_{it}$ and $\hat{e}_{it}$, the estimated amounts of labor and experience for individual i in time t: $\hat{L}_t = \sum_i \hat{l}_{it} = \sum_i \hat{\lambda}_L(s,j,h)\,\hat{z}_{it}\bar{\eta}_{i,t}$ and $\hat{E}_t = \sum_i \hat{e}_{it} = \sum_i \hat{\lambda}_E(s,j,h)\,\hat{g}(e_{it})\,\hat{z}_{it}\bar{\eta}_{i,t}$. Figure 4 illustrates the evolution of these two time-series variables: the relative supply of experience to labor ($\hat{E}_t/\hat{L}_t$) and the relative price of experience ($\hat{\Pi}_{E,t}$) from 1985 to 2009.

Figure 4: Relative Price and Supply of Experience

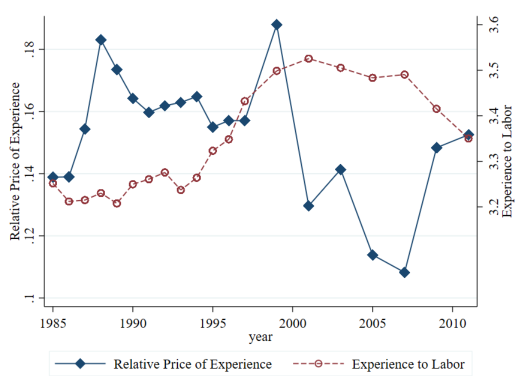


Table 4: Parameter Estimates

Parameters	Coefficient
ρ	-1.413
	(0.011)
θ	2.782
	(0.01)
Time periods	1985 to 2011
Adjusted R^2	0.263

We then use this data to estimate the two production technology parameters. From

$$\Pi_{E,t} \equiv \frac{F_{E,t}}{F_{L,t}} = \theta\left(\frac{E_t}{L_t}\right)^{\rho-1}, \tag{3}$$

we have $\ln \Pi_{E,t} = \ln\theta + (\rho-1)\ln(E_t/L_t)$. Therefore, a linear regression using the aggregate time-series data of relative price and quantity delivers the values for θ and ρ. As documented in Table 4, we find that the elasticity of substitution between the two inputs is 0.41, suggesting that labor and experience are complementary in production.

high and low education as well as between labor and experience, with more aggregate parameters to be estimated. We also abstract from the role of capital in aggregate production, but our identification strategy can be expanded to more generalized production functions with capital, such as the Cobb-Douglas in capital and composite labor ($Y = K^{\alpha}\left(\left(L^{\rho}+\theta E^{\rho}\right)^{1/\rho}\right)^{1-\alpha}$). We view the benchmark specification as a reasonable starting point.

3 The Model

We construct a stochastic life-cycle, general equilibrium model of labor supply and savings, with agents who are subject to health and earnings risks. Our model framework extends those used in the study of health, labor supply, and social insurance programs (e.g., French, 2005; Kitao, 2014) by incorporating the interaction of heterogeneous inputs in the labor market as described in Section 2. We further capture the key features of the DI program similarly to Low and Pistaferri (2015).

3.1 The Model Environment

In this section, we introduce the model environment. We omit education superscripts for brevity, but we allow for education-dependence of some parameters when we quantitatively implement the model.

Preferences, Endowments, and Risks. Individuals' periodic utility is determined by the amount of consumption c, time spent on work $\tilde{l}$, and their health status h, given by

$$u\left(c, l; h\right) \;=\; \frac{\left(c \cdot \exp\left(\eta_h\right) \cdot \tilde{l}\right)^{1-\gamma}}{1-\gamma}. \tag{4}$$

The utility specification follows that of Low and Pistaferri (2015), which allows for health-specific disutility from work through η_h. We assume $\eta_D < \eta_{ND} < 0$, implying that work reduces utility, and more so for disabled workers. Also, disability increases the marginal utility of consumption. Together with the work disutility, individuals incur a health-dependent monetary cost of F_h when working.[13]

An individual is either non-disabled ($h = ND$) or disabled ($h = D$), and health status evolution follows an age-specific Markov chain, $\pi_j^{ab} = \mathbf{Pr}\left(h_{j+1} = b | \, h_j = a\right)$, where j denotes his age.[14] During their life time, individuals face two exogenous risks associated with their health status: medical expenditures and survival. Individuals are subject to medical expenditure risks m, which follows an age- and health-specific stochastic process with mean $\bar{m}_j^h$. An individual of age j and health status h survives the period with probability $\delta_j^h \in (0, 1)$, and the survival probability at maximum age J is zero. The assets of the deceased are distributed equally to all surviving members of the economy in the form of bequest transfer, beq. We

[13] Low and Pistaferri (2015) shows that these components are necessary for replicating the employment patterns.

[14] We assume that the health process is first-order Markov, a commonly used assumption in the literature (e.g., French, 2005; Kitao, 2014). A recent paper De Nardi et al. (2018) captures both the short- and long-run dynamics of health by allowing for history-dependence of health shocks. While the rich modeling would be preferable, we adopt a simpler health transition technology for tractability. Accounting for the rich dynamics could amplify the degree of heterogeneity in the labor market responses to the DI program across health statuses.

focus on the adult phase of the life cycle and assume that an individual must exit the labor market after reaching the mandatory retirement age of $j^R < J$.

All individuals in the model make consumption and savings choices given the risks they face. Further, a working-age individual makes a labor supply decision that is subject to labor market risks. He receives a job offer with probability χ_h, where the wage rate is determined following the specification in Section 2.1. Along with the individual's age, health, and years of work experience, the wage rate w is subject to an idiosyncratic productivity shock ν, which follows a log-normal distribution $\log \nu \sim \mathcal{N}\left(0, \sigma^2_{\nu,h}\right)$ with health-dependent variance.[15] After observing the wage offer, the worker decides whether to work or not. While there is an extensive margin of labor supply, we abstract from the intensive margin and exogenously set the hours worked as health and age-dependent hours l^h_j. Then, the labor income is $w \cdot l^h_j$. Importantly, working-age individuals who are disabled can also decide whether to apply for DI benefits. When they do, they need to forego some of their current income; in particular, we assume that a DI applicant's earnings, disutility of work, and fixed costs are a $\kappa < 1$ share of those of employed workers.

Health Insurance and Government Policies. The access of individuals to the health insurance system depends on their age and labor market status. First, consistent with the employer-sponsored health insurance system in the U.S., employed individuals and DI applicants have access to health insurance with an insurance premium of p_{HI} and a coverage rate of q_{HI}. Second, working-age individuals who are unemployed and DI beneficiaries not yet qualified for Medicare benefits have no access to insurance. Third, qualified DI beneficiaries and retirees are eligible for the Medicare program, a public health insurance program with a premium of p_M and a coverage rate q_M.

We further explicitly model the four government programs—DI, Social Security, Medicare, and Unemployment Insurance (UI)—and implicitly capture other welfare programs (e.g., the Supplemental Nutrition Assistance Program) by assuming that the government provides a consumption floor of amount $\underline{c}_f$. These government programs are funded by labor income tax τ_y, capital income tax τ_k, Social Security tax τ_{ss}, and Medicare tax τ_{med}, which we collectively denote as $\boldsymbol{\tau}$. Further, all other government expenditures are denoted as G.

Working-age individuals can apply for the DI program if they are disabled, which, however, does not guarantee the receipt of DI benefits. The application process is successful with probability π^{DI}, and the worker receives DI benefits that replace the recipient's foregone labor income proportional to his previous

[15] In order to keep the exposition concise, we currently denote job offer arrival to be worker-characteristic and worker-status independent. However, as will be more clear in the following, these health-specific job offer arrival probabilities χ_h will also be allowed to differ according to worker's education and past labor market status (e.g., employed, unemployed, DI applicants, or DI recipients).

earnings, $DI\left(\omega_{DI}\right)$.[16] Further, DI recipients become eligible for Medicare after they receive DI benefits for 24 months. To be consistent with the institutional feature, we assume that DI recipients receive Medicare benefits with probability π^{M} with an expected waiting period of two years. The beneficiary may receive a reassessment of health status with probability π^{RE}. If the individual is not deemed eligible to receive DI (i.e., he is non-disabled), his benefits will be terminated.

Unemployed workers receive UI benefits proportional to their labor market income $UI\left(y\right)$. Retired workers are eligible for Medicare and Social Security benefits of the amount $SS\left(\omega_{SS}\right)$.

Production Technology. A representative firm produces output using labor and experience. The production technology is represented by a CES production function, $Y = A\left(L^{\rho} + \theta E^{\rho}\right)^{1/\rho}$, as discussed in 2.4, and firms trade efficiency units of labor and experience in a competitive market at prices R_L and R_E.

3.2 The Individual's Problem

In this section, we characterize individuals' problems in recursive forms. For working-age individuals, the worker may be of four types—employed (W), unemployed (U), DI applicants (A), and DI beneficiaries (B)—and if retired, he is denoted as a Retiree (R). These individuals make optimal consumption, saving, labor supply, and DI application decisions (in the latter two cases, only if they are of working age) to maximize their discounted utility, given their state variables ($\boldsymbol{x}_{i,}$ for status $i \in \{W, U, A, B, R\}$) and policy parameters of the government. We consider the time-invariant interest rate r to be determined from an external capital market, and all individuals can trade risk-free bonds with an after-tax return of $\tilde{r} \equiv \left(1 - \tau_k\right) r$, while facing a borrowing limit of $\underline{A}$.

The timing of events is as follows. At the beginning of the period, each individual with assets and health status has his medical shock realized. Then, DI reassessment and application results are determined, after which the labor market opens for working-age agents, and labor productivity is realized. Workers then make labor supply and DI application decisions. The workers receive income, UI, DI benefits, or Social Security payments, after which they pay medical and tax bills, consume, and save. Mortality shock is then realized, and agents receive bequests. In the following, we present the value functions for each type of workers.

Employed Workers. An employed individual of age j enters a period with asset level a, health status h, years of work experience e, medical expense m, and idiosyncratic productivity shock ν, and solves the

[16] Under the current DI system in the U.S., DI payments are determined by the worker's average indexed monthly earnings (AIME), the average of the worker's highest 35 years of annual earnings.

following problem:

$$W(\boldsymbol{x}_E) = \max_{c\geq 0, a'\geq \underline{A}} u(c+tr,1;h) + \beta\delta_j^h \pi_j^{h,ND} \left[\begin{array}{c} \chi_h^W \mathbb{E}_{m',\nu'} L(j+1,a',ND,e+1,m',\nu') \\ +\left(1-\chi_h^W\right) \mathbb{E}_{m'} U(j+1,a',ND,e+1,m') \end{array} \right] \tag{5}$$

$$+\beta\delta_j^h \pi_j^{h,D} \left[\chi_h^W \max \left\{ \begin{array}{c} \mathbb{E}_{m',\nu'} L(j+1,a',D,e+1,m',\nu'), \\ \mathbb{E}_{m',\nu'} A(j+1,a',D,e+1,m',\nu') \end{array} \right\} \right] \tag{6}$$

$$+\beta\delta_j^h \pi_j^{h,D} \left[\left(1-\chi_h^W\right) \max \left\{ \begin{array}{c} \mathbb{E}_{m'} U(j+1,a',D,e+1,m'), \\ \mathbb{E}_{m',\nu'} A(j+1,a',D,e+1,m',\nu') \end{array} \right\} \right] \tag{7}$$

$$s.t. \quad c + a' + F_h + p_{HI} + (1-q_{HI})\, m = \tilde{y}\left(w\nu l_j^h; \boldsymbol{\tau}\right) + (1+\tilde{r})\, a + beq, \tag{8}$$

where $\boldsymbol{x}_E \equiv (j,a,h,e,m,\nu)$. His utility today is drawn from consumption and (dis)utility from work. The government's welfare program ensures that the worker is able to consume at least the amount of the consumption floor so that $tr = \max\left\{\underline{c}_f - c, 0\right\}$ (for all individuals in the economy). Given the price of labor (R_L) and experience (R_E) in the market, the base wage of the worker is $w(j,h,e) = R_L\lambda_L(j,h) + R_E\lambda_E(j,h)\, g(e)$ (as in Section 2.1), which depends on the worker's current age, health status, and the years of experience. Total labor earnings therefore are $y(j,h,e) = w(j,h,e)\,\nu l_j^h$, where ν is the iid productivity factor and l_j^h is the hours worked. Conditional on working, we assume that hours are exogenously determined by age and health status: with this notation, total hours worked is then $\tilde{l} \cdot l_j^h$, in which $\tilde{l}$ reflects a choice of not working ($\tilde{l} = 0$); working ($\tilde{l} = 1$); or applying to the DI program (which exogenously implies $\tilde{l} = \kappa$, as described in Section 3.1). As seen in the budget constraint (Equation (8)), the total income of the individual consists of after-tax labor income ($\tilde{y}(w\nu l_h; \boldsymbol{\tau})$), capital income, and bequests. The individual spends these resources on consumption, cost of work, savings, and medical expenditures that consist of premiums and out-of-pocket costs.

In the next period, if he survives (with probability δ_j^h) and turns out to be non-disabled ($\pi_j^{h,h'=ND}$), there are two possibilities (line (5)). He may receive a job offer, which happens with probability χ_h^W. Note that the job offer arrival rates χ is dependent on health and labor market status ($\{W,U,A,B\}$) in order to potentially capture the impacts of attachment to the labor market in their future labor market opportunities. When he receives the offer, then he either makes his labor market decision denoted by the value $L \equiv \max\{W,U\}$ or else does not receive a labor market offer ($1-\chi_h^W$) and becomes unemployed. If the worker becomes disabled, then, his choice set expands as he can also choose to apply (A) or enter the labor market if an offer is received (lines (6) and (7)). As an employed worker today, his experience increases to $e+1$ at the start of his next period.

Unemployed Workers. The unemployed workers' problem looks similar to that of the employed, with the state vector of $\boldsymbol{x}_U \equiv (j, a, h, e, m)$.

$$
\begin{aligned}
U(\boldsymbol{x}_U) \quad = \quad & \max_{c\geq 0, a'\geq \underline{A}} u(c+tr, 0; h) + \beta\delta_j^h \pi_j^{h,ND} \left[\begin{array}{c} \chi_h^U \mathbb{E}_{m',\nu'} L(j+1, a', ND, e, m', \nu') \\ +\left(1-\chi_h^U\right) \mathbb{E}_{m'} U(j+1, a', ND, e, m') \end{array} \right] \\
& + \beta\delta_j^h \pi_j^{h,D} \left[\chi_h^U \max \left\{ \begin{array}{c} \mathbb{E}_{m',\nu'} L(j+1, a', D, e, m', \nu'), \\ \mathbb{E}_{m',\nu'} A(j+1, a', D, e, m', \nu') \end{array} \right\} \right] \\
& + \beta\delta_j^h \pi_j^{h,D} \left[\left(1-\chi_h^U\right) \max \left\{ \begin{array}{c} \mathbb{E}_{m'} U(j+1, a', D, e, m'), \\ \mathbb{E}_{m',\nu'} A(j+1, a', D, e, m', \nu') \end{array} \right\} \right] \\
s.t. \qquad & c + a' + m = UI(y) + (1+\tilde{r}) a + beq.
\end{aligned}
$$

The source of income for the unemployed is UI benefits. The individual no longer incurs monetary costs from work and is without health insurance. Further, he does not accumulate experience; thus, tomorrow's experience stays at e.

DI Applicants. Disabled workers have an option to apply for DI benefits,[17] and their value reads

$$
\begin{aligned}
A(\boldsymbol{x}_A) \quad = \quad & \max_{c\geq 0, a'\geq \underline{A}} u(c+tr, \kappa; h) \\
& + \beta\delta_j^h \sum_{h'} \pi_{j+1}^{hh'} \left[\begin{array}{c} \pi^{DI} DI^{i_M=0}(j+1, a', h', e, m') \\ +\left(1-\pi^{DI}\right) \left[\begin{array}{c} \chi_h^A \mathbb{E}_{m',\nu'} L(j+1, a', h', e, m', \nu') \\ +\left(1-\chi_h^A\right) \mathbb{E}_{m'} U(j+1, a', h', e, m') \end{array} \right] \end{array} \right] \\
s.t. \qquad & c + a' + \kappa \cdot F_h + p_{HI} + (1-q_{HI}) m = \tilde{y}\left(\kappa \cdot w\nu l_j^h; \boldsymbol{\tau}\right) + (1+\tilde{r}) a + beq,
\end{aligned}
$$

where $\boldsymbol{x}_A \equiv (j, a, h = D, e, m, \nu)$. The applicant works for κ share of his time, lowering his income, but at proportional disutility and monetary costs. As a partially attached worker, he has access to health insurance.[18] In the next period, if successful (with probability π^{DI}), the worker becomes a DI recipient without Medicare denoted by value $D^{i_M=0}$. If not successful, he becomes unemployed, unless he is given the opportunity to enter the labor market.

DI Beneficiaries with ($i_M = 1$) and without Medicare ($i_M = 0$). The value of being a DI beneficiary differs according to whether the recipient also qualifies for Medicare benefits or not. Thus, we differentiate the value of DI recipients according to their receipt of Medicare benefits ($i_M = 1$ if the individual is a

[17] We do not allow non-disabled workers to apply; however, it may be that endogenously, it is not in their best interest to do so. In some sense, our notion of disability (from the PSID at least) may extend beyond those who actually receive DI.

[18] Under the Consolidation Omnibus Budget Reconciliation Act (COBRA), workers have the right to continue group health benefits after leaving work for limited periods of time.

Medicare beneficiary) as follows:

$$
\begin{aligned}
B^{i_M}(\boldsymbol{x}_B) \quad = \quad & \max_{c\geq 0, a'\geq \underline{A}} u(c+tr,0;h) \\
& +\beta\delta_j^h\left(\left(1-\pi^{RE}\right)+\pi^{RE}\pi_{j+1}^{h,D}\right)\ \mathbb{E}_{m'}\boldsymbol{E}B^{i_M}(j+1,a',h',e,m') \\
& +\beta\delta_j^h\pi^{RE}\ \ \pi_{j+1}^{h,ND}\left[\begin{array}{c} \chi^B\mathbb{E}_{m',\nu'}L(j+1,a',ND,e,m',\nu') \\ +\left(1-\chi^B\right)\mathbb{E}_{m'}U(j+1,a',ND,e,m') \end{array}\right] \\
s.t. \qquad & c+a'+i_M\left(p_M+(1-q_M)\,m\right)+(1-i_M)\,m=DI(\omega_{DI})+(1+\tilde{r})\,a+beq,
\end{aligned}
$$

with $\boldsymbol{x}_B \equiv (j,a,h,e,m)$. Whether the beneficiary receives Medicare impacts his medical expenditures through the budget constraint. In the following period, if the worker is not reassessed $(1-\pi^{RE})$ or is reassessed and passes (i.e., he is disabled, $\pi^{RE}\pi_{j+1}^{h,h'=D}$), he remains a DI recipient, with expected value $\boldsymbol{E}B^{i_M}(a',h',e,m')$. The expected value is $\boldsymbol{E}B^{i_M=1}=B^{i_M=1}$ for already qualified Medicare beneficiaries and $\boldsymbol{E}B^{i_M=0} = \pi^M B^{i_M=1}+\left(1-\pi^M\right)B^{i_M=0}$ for not-yet-qualified Medicare beneficiaries, the latter of which reflects the future probability of receiving Medicare benefits. If the beneficiary does not pass the reassessment (i.e., he is non-disabled when reassessed), his benefits are terminated. Then, he either becomes unemployed or enters the labor market with probability χ^B. Unlike χ_h^W, χ_h^U, or χ_h^A, which are health-dependent, all workers leaving DI after reassessment are non-disabled and thus homogeneous in their job offer arrival rates.

Retirees. Once retired, individuals receive Social Security benefits based on their earnings history ω and make optimal consumption and saving decisions:

$$
\begin{aligned}
R(\boldsymbol{x}_R) \quad = \quad & \max_{c\geq 0, a'\geq \underline{A}} u(c+tr,0;h)+\beta\delta_j^h\mathbb{E}_{m'}R(j+1,a',h',\omega,m') \\
s.t. \qquad & c+a'+p_M+(1-q_M)\,m=ss(\omega_{SS})+\{1+(1-\tau_t^a)\,r\}\,a+beq
\end{aligned}
$$

with $\boldsymbol{x_R} \equiv (j,a,h,\omega,m)$. In the last period of their working lives, individuals' SS benefits are determined by their past average earnings ω, which becomes an individual's state variable that does not change for the rest of his life.

3.3 Competitive Equilibrium

Let the vector of the state space of individuals (as defined in Section 3.2) be denoted as $\boldsymbol{x} \equiv \{\boldsymbol{x}_W, \boldsymbol{x}_U, \boldsymbol{x}_A, \boldsymbol{x}_B, \boldsymbol{x}_R\}$. Given the government's policy parameters, the competitive equilibrium of the economy consists of individuals' policy functions and value functions; factor prices of labor and experience; the size of bequest transfers;

and the distribution of individuals over the state space $\mu(\boldsymbol{x})$ such that the following conditions are held.

1. The individual policy functions solve their optimization problems as defined in Section 3.2.

2. Factor prices for labor (R_L) and experience (R_E) are determined competitively: $R_L = A\left(L^{\rho} + \theta E^{\rho}\right)^{(1-\rho)/\rho} L^{\rho-1}$ and $R_E = \theta A\left(L^{\rho} + \theta E^{\rho}\right)^{(1-\rho)/\rho} E^{\rho-1}$.

3. Factor markets clear: $L = \sum_{\boldsymbol{x}} \tilde{l}(\boldsymbol{x})\,\lambda_L(j,h)\,l_j^h \mu(\boldsymbol{x})$ and $E = \sum_{\boldsymbol{x}} \tilde{l}(\boldsymbol{x})\,\lambda_E(j,h)\,g(e)\,l_j^h \mu(\boldsymbol{x})$, where $\lambda_L(\cdot)$ and $\lambda_E(\cdot)$ are defined as in Section 2.1.

4. The bequest transfer equals the amount of assets left by the deceased: $beq = \sum_{\boldsymbol{x}} a(\boldsymbol{x})\left(1-\delta_j^h\right)\mu(\boldsymbol{x})$.

5. The government budget is satisfied:

$$\sum_{\boldsymbol{x}} \left\{SS(\boldsymbol{x}) + DI(\boldsymbol{x}) + UI(\boldsymbol{x}) + tr(\boldsymbol{x}) + q_M \bar{m}_j^h \mathbb{I}_M(\boldsymbol{x})\right\}\mu(\boldsymbol{x}) + G$$
$$= \textstyle\sum_{\boldsymbol{x}} T\left(y(\boldsymbol{x}), a(\boldsymbol{x})\right)\mu(\boldsymbol{x}),$$

   in which $\mathbb{I}_M(\boldsymbol{x})$ is an indicator for whether the individual qualifies for Medicare (either DI recipients with Medicare or retirees), and $T\left(y(\boldsymbol{x}), a(\boldsymbol{x})\right)$ denotes the total tax (labor and capital income, SS, Medicare) paid by agents with labor income $y(\boldsymbol{x})$ and assets $a(\boldsymbol{x})$.

4 Calibration

In this section, we describe how we map our model to the data to quantitatively evaluate the impacts of the DI program. For empirical implementation of the model, we allow for two education (s) types: workers with less than or equal to 12 years of education (high school graduates, "HS") and those with more than 12 years of education (college, "Col"). In particular, we allow for education-dependence in iid productivity shock variances $\sigma^2_{\nu,\{h,s\}}$, work disutility $\eta_{h,s}$, fixed costs of work $F_{h,s}$, and offer arrival rates $\chi^X_{h,s}$ for labor market statuses $X \in \{W, U, A, B\}$.

We first document the parameters that are calibrated outside and inside the model, discuss the model's performance on targeted moments, and then validate the model.

4.1 Exogenously Calibrated Parameters

The unit of time in our analysis is a year, and the unit of analysis is an individual.[19] High school graduates start their lives at 18, and college graduates start their lives at 22. All workers retire at the mandatory age of 65 and live at most to 100. We take the 2015 demographic composition of the U.S. population from the National Population Projections by the U.S. Census Bureau.[20] The CRRA parameter γ in the utility function is set exogenously at 2. Workers are borrowing-constrained, and we assume a small open economy where risk-free bonds earns 3% annual returns.

Health, Survival, and Labor Income. Our main source for calibrating health-related parameters is the Panel Study of Income Dynamics (PSID). We classify health status into two categories, non-disabled ($h = ND$) and disabled ($h = D$), based on the binary indicator of work limitation, which is consistent with the empirical specification presented in Section 2. Health status in the model impacts the worker's (i) survival probability; (ii) evolution of health statuses; (iii) medical expenditures; earnings through (iv) hours worked and (v) wage profiles; (vi) job offer arrival rates; and (vii) disutility and fixed costs from work. We exogenously calibrate the first four and endogenously calibrate the last two within the model, which are discussed in Section 4.2. The wage profiles were estimated in Section 2,[21] and we use the residual variances from the wage regression as health- and education-specific variances of the iid productivity shock process. In the following, we discuss how we determine parameters for survival probabilities, evolution of health statuses, and hours worked.

We estimate the impact of health on conditional survival probabilities, using the life table from the Social Security Administration and micro-level data from the PSID. Following the strategy of Attanasio et al. (2011), we obtain age-dependent survival probabilities ($\bar{\delta}_j$) from the life table, and we obtain the empirical health distribution by age (p_j^h) and survival rates by health status and age δ_j^h from the PSID. Then, we use the following equations to obtain health-dependent conditional survival probabilities that are consistent with the life tables: (i) $\bar{\delta}_j = \sum_{h \in \{ND,D\}} p_j^h \delta_j^h$; and (ii) $\Delta_j = \delta_j^{ND} - \delta_j^D$. The second equation represents the survival premium of being non-disabled, relative to having a work limitation (Δ_j). Given the small samples in the PSID, we smooth out survival premia (Δ_j) by fitting polynomials to age, and extrapolate them for individuals older than 90.[22] Figure 5 shows the estimated health-dependent conditional

[19] We abstract from gender in the analysis. This is a simplifying assumption in our quantitative model and is a consistent assumption with the benchmark empirical wage analysis, where we use both genders and control for gender-specific effects.

[20] More detailed information is available in Appendix A.1.

[21] We impose a minimum wage rate of \$5, and all data moments were constructed consistently.

[22] Attanasio et al. (2011) uses Health and Retirement Study (HRS) data to calculate health-dependent survival probabilities. However, the HRS only includes individuals who are older than 50. When we compare our survival premia to theirs, the magnitudes

survival probabilities.

Health status in the model evolves stochastically according to $\pi_j\left(h'|h\right)$, which depends on the worker's age and his current health status. We use the panel dimension of the PSID to find the transition probabilities for five age groups (18–29; 30–41; 42–53; 54–65; and 66 and older) and fit these moments to a quadratic function of age to produce smooth transitions over the life cycle. The estimated transition probabilities are plotted in Figures 6(a) and 6(b). As is clear from the plots, health statuses are persistent, and older workers are more likely to make a transition to having work limitations than young workers.[23]

Figure 5: Survival Probability

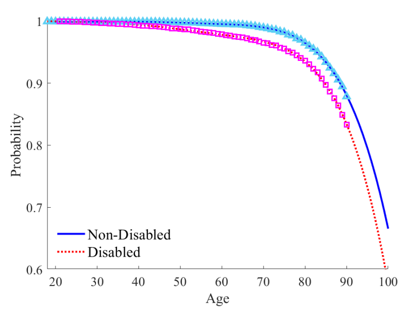
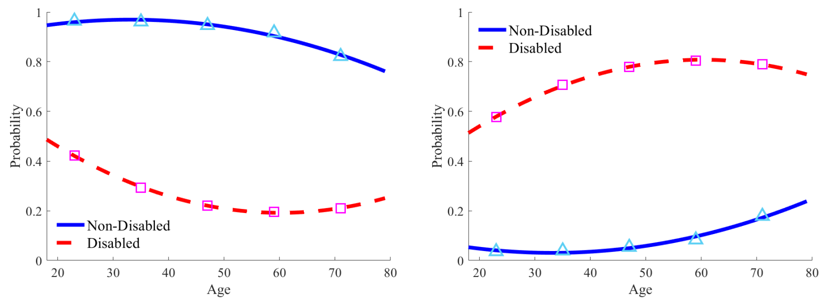


Figure 6: Health Transition Probabilities

(a): Transition to Non-Disabled (b): Transition to Disabled



Note: Markers are data points from the PSID, which we use to estimate survival and transition probabilities by health and age.

Conditional on working, we assume that individuals work fixed hours. For employed workers with age j and health status h, we construct their working hours l_j^h as the average working hours among the employed workers who reported more than 250 working hours per year. In doing so, to smooth out variations from small sample sizes, we compute hours in age j as the average of hours over ages $j-1$, j, and $j+1$. These working hours are plotted in Figure 7.[24]

Medical Expenditures and Health Insurance. Medical expenditure risks differ by age and health statuses. We use adult-equivalent medical expenditures from the PSID to construct these variables. Following Attanasio et al. (2011), we use three medical expenditure bins representing the averages in the 1st–60th percentile, 61st–95th percentile, and 96th–100th percentile. These bins are chosen to capture the long tail in medical expenditure distributions from catastrophic events. Similar to the approach used for health transition functions, we fit medical expenditures using a quadratic function in age j, which are plotted in Figures 8(a) and 8(b).

seem similar.

[23]Due to sample size issues, we use the same parameters for health transition and medical expenditures for those aged 80 or older.

[24]Hours worked are not education-dependent as, empirically, we find small differences in working hours across education (conditional on 250 hours minimum).

Figure 7: Hours Worked

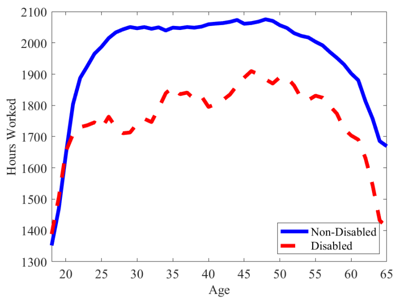
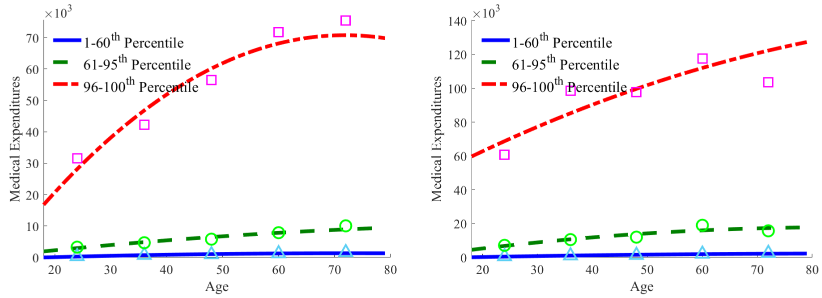


Figure 8: Medical Expenditures

(a): Non-Disabled (b): Disabled



Note: Markers in Figure 8 are data points from the PSID, which we use to estimate survival and transition probabilities by health and age.

We assume that employed workers and DI applicants have access to employer-sponsored health insurance (ESHI) with a constant coverage rate q_{HI} of 60% and a premium p_{HI} of $2,500, values similar to those used in Imrohoroglu and Kitao (2012). The constant health insurance premium captures that ESHI is a group insurance (workers do not pay actuarially fair premium by age and health, for example). However, this is a simplifying assumption as we do not impose the break-even condition of the health insurance system.[25] To deal with this issue, we assume that the differences in expenditures and premia are paid by the government. As these ESHI premia are tax-exempted and thus partially funded by taxes, this may be a reasonable assumption. Moreover, the quantitative magnitude of the differences in expenditures and premia turn out to be small.

Government Policies. Lastly, we discuss parameters for government policies: Disability Insurance, Unemployment Insurance, Social Security, Medicare, welfare programs, and tax policies.

Disability Insurance. There are five parameters that fully describe the DI program: the application penalty parameter for workers κ, the application success probability π^{DI}, the probability of qualifying for Medicare benefits π^{M}, the reassessment probability π^{RE}, and the benefit schedule as a function of previous earnings $DI\left(\omega\right)$.

DI applicants have a waiting period of around five months to receive DI benefits; thus, we assume that applicants earn 60% of their labor income.[26] The DI receipt probability is set at 40%, following the findings from Chen and van der Klaauw (2008). The Medicare receipt probability is 50% to capture that the beneficiary expects to qualify for the benefits after two years. The reassessment probability is set at 6%,

[25]If we imposed the break-even condition, we would need to solve for a fixed-point for equilibrium premium, increasing the computational burden. We choose to simplify the modeling of the ESHI in order to enrich the model in key dimensions, e.g., endogenous experience accumulation and a more detailed DI program.

[26]Disutility and fixed costs of work are also scaled down by the same proportion.

similar to that used in Low and Pistaferri (2015). Lastly, the DI payments are determined by the following Primary Insurance Amount (PIA) formula (in 2011 dollars):

$$PIA(\omega) = \begin{cases} 0.90 \times \omega & \text{if } \omega < \$8,988 \\ \$8,089 + 0.32 \times (\omega - \$8,988) & \text{if } \$8,988 \leq \omega < \$54,204 \\ \$22,559 + 0.15 \times (\omega - \$54,204) & \text{if } \omega \geq \$54,204. \end{cases} \tag{9}$$

where ω reflects the worker's AIME, the average of the worker's 35 highest years of earning. As it is difficult to keep track of each worker's earnings given the large state space,[27] we approximate ω_{DI} using state variables. Specifically, we use the average labor earnings of the worker, given his education, age, and years of experience, such that $\omega_{DI}(j, s, e) = \mathbb{E}_h \left[w(j, s, h, e) \cdot l_j^h \right]$. To better reflect the average previous earnings, we disregard the iid shock ν and take the average across the health status distribution at age j. Thus, ω_{DI} reflects the heterogeneity in AIMEs by workers' education, age, and experience. While this is not perfect, it reasonably approximates the past earnings of individuals with heterogeneous earnings profiles with reduced computational burden. Finally, we follow the policy cap on AIME for benefit calculation, imposing $DI(\omega) = \min\{PIA(\omega_{DI}), \$30,448\}$.

Unemployment Insurance. UI benefits are paid to unemployed workers. With about a 45% replacement rate that pays up to six months, the overall yearly replacement rate is set at 23% of the worker's annual income.

Social Security and Medicare Benefits. Social Security payments are also determined by the PIA in Equation (9). For ω_{SS}, we use a similar approximation as for DI but require 35 years of work experience.[28] Medicare benefits are provided to all retirees and to qualified DI recipients. Beneficiaries pay a premium of $p_M = \$1,157$, and its coverage rate q_M is 50%. The Medicare tax rate is $\tau_M = 0.029$, levied on labor earnings.

Other Taxes and Welfare Programs. Labor income is taxed at rate $\tau_y = 0.26$; the capital income tax rate is $\tau_k = 0.1$.[29] We set the consumption floor as $\underline{c}_f = \$3,150$ to capture other un-modeled welfare programs provided by the government.[30] Social Security taxes are set at $\tau_{ss} = 0.104$, levied on labor earnings, with a maximum taxable earnings of $y_{ss} = \$106,800$.

[27] To be more accurate, one could keep AIME as an additional state variable, an approach taken by Kitao (2014). However, as we keep track of the years of work experience, the additional state variable would be too burdensome computationally. Thus, we choose to exploit experience as an additional observable reflecting workers' previous earnings and show in Section 4.3 that we are able to match the average DI benefit amounts of workers over the life cycle in the calibrated model.

[28] As years of experience is our state variable, we can capture the impact of years of experience on workers' SS benefit determination: if a worker worked for 20 years, for example, we use (as does the U.S. policy) zero as earnings for 15 years. The work requirement for DI benefits are a lot more relaxed; thus, we do not impose such experience restrictions for the approximation of ω_{DI}.

[29] We assume a constant capital income tax rate, similar in level to the long-term capital gains tax rate.

[30] This is within the range used in the literature: the annual consumption floor is set at \$4,000 in Kitao (2014) and estimated to be \$1,540 (in 2003 dollars) in Pashchenko and Porapakkarm (2017) and \$3,593 in De Nardi et al. (2018).

Production Technology. The values for ρ and θ in the aggregate production function $Y = A\left(L^{\rho} + \theta E^{\rho}\right)^{1/\rho}$ are taken from the estimated values reported in Section 2.4.

We summarize the values of all exogenously calibrated parameters in Table 5.

Table 5: Parameters Calibrated Outside the Model

Parameters	Description	Values	Parameters	Description	Values
Demographics, Preferences, Technology			Policies: UI, SS, Medicare, Tax		
$\{n_j\}$	Population share	Appendix A.1	b	UI replacement rate	0.23
$\{\delta_j^h\}$	Survival rates	Fig. 6	τ_y	Labor income tax	0.26
γ	Risk aversion	2	τ_k	Capital income tax	0.10
r	Interest rate	0.03	τ_{SS}	SS tax	0.104
$\{\rho, \alpha\}$	Agg. production	-1.41; 2.87	y_{SS}	Max. taxable earnings	\$106,800
Wage and Hours			τ_M	Medicare tax	0.029
$w(j,h,s,e)$	Wage process coefficients	Table 2	$\{p_M, q_M\}$	Medicare prem., coverage	\$1,157; 0.5
$\sigma^2_{\nu,\{ND,s\}}$	iid shock var., non-disabled	0.42; 0.50	$\underline{c}_f$	Consumption floor	\$3,200
$\sigma^2_{\nu,\{D,s\}}$	iid shock var., disabled	0.74; 0.64	Policy: Disability Insurance		
l_j^h	Hours worked	Fig. 8	κ	Application penalty	0.6
Health, Medical Expenditures, and Health Insurance			π^{DI}	DI receipt prob.	0.4
$\{\pi_j(h'\|h)\}$	Health transition	Fig. 6	π^M	Medicare benefit prob.	0.5
$\{m_j^h\}$	Medical expenditures	Fig. 8	π^{RE}	Re-examination prob.	0.06
$\{p_{HI}, q_{HI}\}$	HI prem., coverage	\$2,500; 0.6	$\{PIA(\omega)\}$	Primary Insurance Amount	Eq. (9)

4.2 Parameters Calibrated within the Model

There are 25 remaining parameters: $\left\{A, \beta, \eta_{h,s}, F_{h,s}, \chi^W_{h,s}, \chi^U_{h,s}, \chi^A_{h,s}, \chi^B_s\right\}$ for $h \in \{ND, D\}$ and $s \in \{HS, Col\}$. We calibrate these parameters to match employment rates[31] by health, education, and age group[32]; share of DI recipients by age group[33]; average consumption of non-disabled workers; and average hourly wage rate by health statuses (48 moments). Given the rich modeling of the labor market, the preferences and labor market parameters jointly match the life-cycle employment rates by health and education. We also directly target the share of DI recipients by age group to ensure that the model replicates the life-cycle share of DI recipients. This latter pattern is determined by disabled workers' labor market opportunities controlled by offer arrival rates for DI recipients and applicants. Further, while we take the relative efficiency of aggregate experience θ and the elasticity of substitution between labor and experience $1/(1-\rho)$ as exogenous, we use the average TFP parameter A to match the level of the wage rate in the

[31] To calculate employment rates in the simulated model, we include employed workers and 60% of applicants as our model assumes that applicants work 60% of their time, while we are not able to identify applicants as a separate share of the workforce from the data.

[32] We use nine age groups: under 25, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-60, 61-65.

[33] We obtain this data from the Social Security Administration (2013), where only the specified age group (the nine age groups that we use) level data is available.

model. Lastly, the time preference of individuals β informs the consumption level of workers.

The values of calibrated parameters are presented in Table 6. We observe that disabled workers have higher disutility of work and higher fixed costs of work. The job offer arrival rates (tomorrow) vary across workers' labor market statuses (today): highest for the employed, lower for the unemployed and DI applicants, and the lowest for DI recipients. The lowest offer arrival rates for DI recipients potentially captures the difficulty of returning to the labor market after being a DI recipient. These trade-offs are key determinants in workers' decisions to apply for the DI program. The estimated offer arrival rates differ across health status and education group. Although we see that across-education differences in work disutility parameters are small, college educated workers experience significantly more favorable job offer rates relative to high school graduates.

Table 6: Parameters Calibrated within the Model

Parameters	Description	Value			
A	Aggregate productivity	0.675			
β	Time discount factor	0.953			
		High School		College	
		Non-Disabled	Disabled	Non-Disabled	Disabled
$\eta_{h,s}$	Disutility of work	-0.104	-0.193	-0.106	-0.191
$F_{h,s}$	Fixed cost of work	752	925	988	1,046
$\chi^W_{h,s}$	Offer arrival rates: Employed	0.926	0.798	0.995	0.891
$\chi^U_{h,s}$	Offer arrival rates: Unemployed	0.787	0.685	0.910	0.741
$\chi^A_{h,s}$	Offer arrival rates: Applicants	0.769	0.608	0.891	0.744
χ^B_s	Offer arrival rates: DI beneficiaries	0.333	-	0.586	-

Figure 9: Employment Rates: Data vs. Model

(a): Employment: High School (b): Employment: College

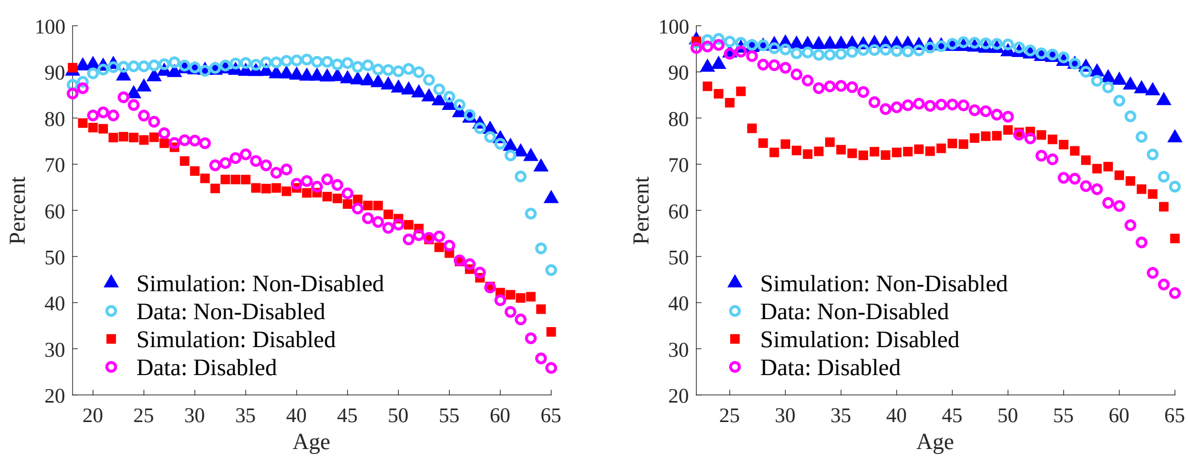


Figure 10: DI Recipient Share:
Data vs. Model

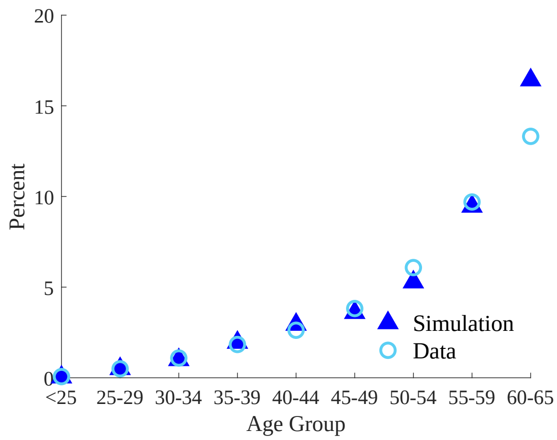


Table 7: Consumption and Wage: Data vs. Model

	Data	Simulation
Avg. consumption, non-disabled	\$20,302	\$21,923
Hourly wage, non-disabled	\$24.80	\$22.86
Hourly wage, disabled	\$21.95	\$20.90

We plot the model-generated employment rates by health status and education over the life cycle with data in Figures 9(a) and 9(b).[34] The model is able to replicate the employment rates of workers quite well. Figure 10 and Table 7 present the performance on the remaining targets. The DI recipient shares are well-replicated in the model, as are the average consumption and average hourly wages by health statuses.

4.3 Model Validation

Before conducting counterfactual analyses, we validate the model by evaluating its performance on non-targeted moments.[35]

Life-Cycle Patterns in Worker Outcomes: Experience, Earnings, and Consumption. We first document the model's performance in replicating the empirical life-cycle patterns of workers' years of experience, earnings, and consumption, none of which served as targets in calibration. Figure 11(a) shows the average years of experience over the life cycle in the simulated model and in the PSID. We find that the simulated model exhibits higher average experiences because we do not model additional heterogeneity to generate those with near zero experience as in the data. Another important aspect of the model is whether it is able to generate differences in experience accumulated by workers with heterogeneous health status. For this analysis, we first categorize workers into two groups in the data–those who experienced disabilities more than 30% of the periods, and those who did not–to compare their experience patterns over the life

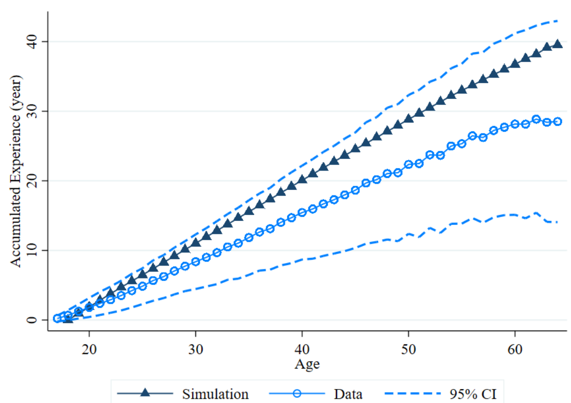
[34] While we target employment rates by health and education for nine age groups, we here present the full life-cycle pattern to show the model's performance at a more disaggregated age level.

[35] While we model retirement periods, we primarily focus on individual's behaviors during their working lives, the main focus of our analyses.

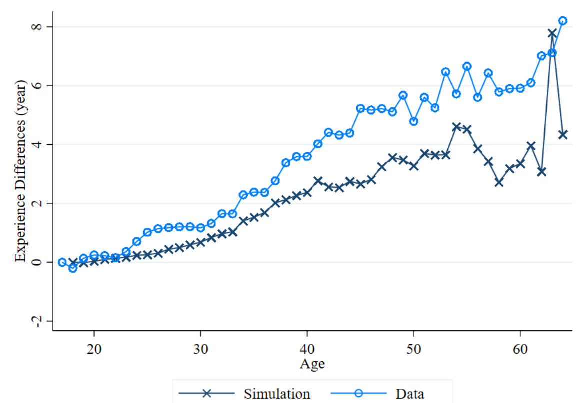
cycle.[36] Figure 11(b) plots the differences in years of experience across these two worker categories. It confirms that the overall growth rates of experience and the difference in accumulated experiences by health status are similar to the patterns from the PSID.

Figure 11: Accumulated Experience: Simulation vs. Data

(a): Experience over the Life Cycle



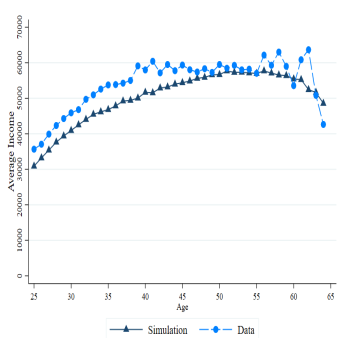
(b): Experience Difference by Health



Note: Dashed lines in Figure 11(a) indicates the 95% confidence interval of PSID data.

Figure 12: Earnings over the Life Cycle: Data vs. Model

(a): Non-Disabled



(b): Disabled

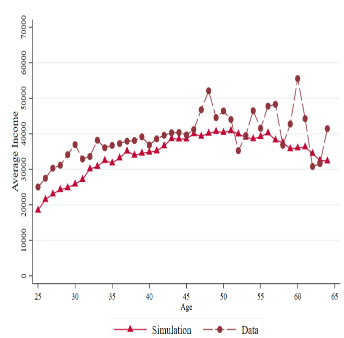
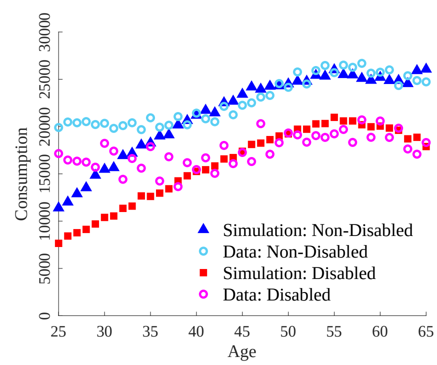


Figure 13: Non-Medical Consumption over the Life Cycle: Data vs. Model



Note: Consumption data in Figure 13 is constructed using the PSID survey data for years 1999 through 2013 and includes food, utilities, transportation, education expenses, childcare, clothing, trips, and recreation categories. We use those aged older than 25 for sample size, and control for family size using an equivalence scale with 0.5 weight on an additional adult and 0.3 on an additional child.

[36]Documenting, for example, the years of experience of individuals by health status at a certain age is misleading as these health statuses are not permanent (though persistent). Thus, we compare the accumulated experience differences of individuals who reported having a disability "a lot" of times with those who did not during their lifetime. On average, a respondent in the PSID has 18.23 periods of observations with a standard deviation of 8.25 in the PSID. We restrict the sample to individuals with at least 10 observed periods and compute the number of periods with a reported disability up to the current age. We apply the same sampling criteria to simulated data to generate the results in Figure 11(b).

Second, we compare workers' earnings[37] and consumption by health statuses over the life cycle.[38] As shown in Figures 12 and 13, our model broadly matches the life-cycle patterns in the data. While the model under-estimates consumption at earlier ages, it is able to match the consumption of workers starting in the late 30s. One of the reasons might be that our model lacks sources of insurance in consumption for the young (e.g., inter vivos transfers from parents), unlike in the data.

Characteristics of DI Applicants and Beneficiaries. Now, we examine our model's performances on DI recipients. In Figure 14, we show the model-predicted share of DI applicants that leads to the DI recipient share in Figure 10. Despite the fact that DI receipt probability is constant over the life cycle, there is a steep increases in the applicant share after the age of 45. This shows that the model is able to capture the trade-offs that workers face in their decision to apply for DI that depends on their labor market opportunities over the life cycle. Further, Figure 15 shows the average DI payment by age in the simulated model compared to those reported by the Social Security Administration (2013). The report documents the average DI benefit amounts by age group, which the model is able to match quite well: the average DI benefit amount is \$14,000 in the model and \$13,100 in the data.[39]

Figure 14: DI Applicant Share

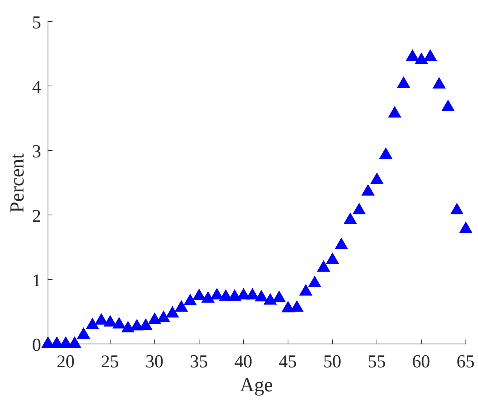
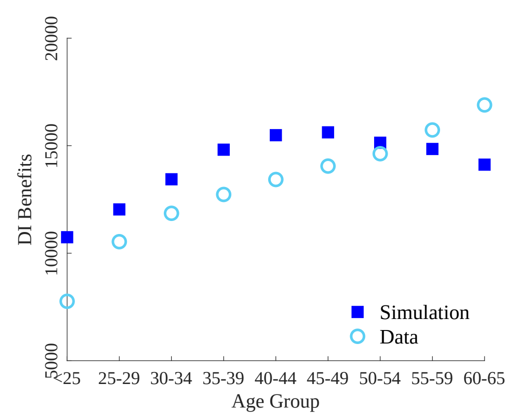


Figure 15: DI Benefit Amounts



In Figure 16, we compare the average experience of DI recipients and non-DI recipients in the simulated model and in the data. We obtain empirical experience profiles from the PSID data.[40] The profile shows that DI recipients near the retirement age have work experiences around 10 years lower than non-DI recipients, and our model is able to replicate the pattern quite well. Further, the ratio of average assets of non-DI

[37] We further include the model's fit on wage profiles by education and health in Appendix C.1.

[38] We target employment rates over the life cycle but only target average wage by health status and average consumption of non-disabled workers during their working lives.

[39] This also implies that the way we approximate the PIA for DI is reasonable.

[40] During the years 1984-1992 and from 2005 onwards, PSID asks respondents, the type of Social Security received, one of which is SSDI. These statistics are based on these survey data. We also report summary statistics of DI recipients in Appendix A.

recipients relative to DI recipients is around 2.5 in the data and 2.3 in the model,[41] indicating that DI recipients have lower assets compared to the rest of the population.

Figure 16: Average Experience by SSDI Status

(a): Non-SSDI Recipients

(b): SSDI Recipients

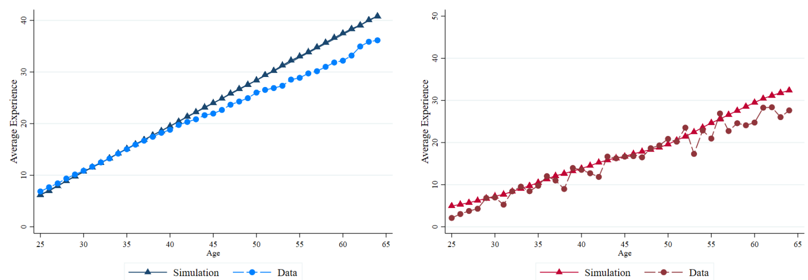
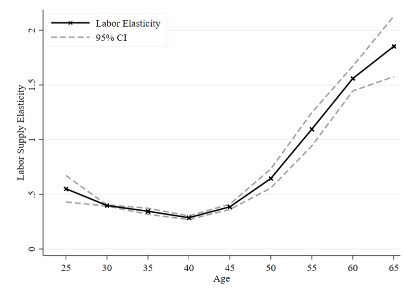


Figure 17: Labor Supply Elasticity



Note: We identify DI recipients using the PSID's question on the type of Social Security received, which is available during the years 1984-1992 and from the year 2005 onwards. We use those aged 25 and older for sample size issues.

Lastly, DI recipients in the model are strongly attached to the program: about 4% of surviving working-age DI recipients (thus excluding exiting due to retirement or death) exit the program. As we assume that benefits are terminated upon failing a reassessment and that workers do not leave the program voluntarily, this rate is determined exogenously by reassessment and health transition probabilities. According to the Social Security Administration (2017), around 10% of DI recipients had their benefits terminated, 87% of them due to reaching retirement age or death (1.3% termination rate for reasons other than death or retirement). Moreover, about 2% of recipients has their benefits withheld yearly. If we include both as flows off the DI program, the exit rate among surviving working-age beneficiaries is around 3%.[42] While we note that the exit rates are exogenously determined by the parameters of the model and that ours is higher,[43] we emphasize that the model endogenously generates the rate at which workers return to the labor market (through job offer arrival rates, wage processes, and the participation choice of workers). In the model, on average, conditional on rejection, about 63% of applicants become employed the next period. This is similar to the rate documented in Maestas et al. (2013) (see Figure 2 for the denied applicants' employment probabilities). Further, the model generates that these rates are higher for younger workers,

[41]We plot the model-simulated asset distribution of DI recipients and non-DI recipients in Appendix C.1. It is worth noting that the asset variables in the PSID are limited and have many missing variables. Thus, we do not have very reliable asset data, especially on the small sample of DI recipients.

[42]The termination rates among workers have been around 8-10% in the recent years. The benefit may be withheld for reasons such as administrative issues (address unknown) or due to pending determination of continuing disability. These statistics are drawn from Tables 48, 49, and 50 of Social Security Administration (2017).

[43]We could potentially target this exit rate by assuming that a share of the population is permanently disabled and thus never leaves the DI program. Although we do not take that approach, we think that overall, we are able to broadly replicate the strong attachment to the DI program and the behaviors of DI recipients as we discuss in this section.

similar to findings by Maestas et al. (2013).[44] Thus, overall, our model is able to capture some of the key behaviors of DI applicants and recipients.

Labor Supply Elasticity. As one of the primary focuses of our paper is analyzing labor market effects, we verify whether the model is able to generate reasonable labor supply elasticities. To compute the labor supply elasticity, we conduct experiments in which individuals of age j experience an unanticipated increase in wage for one period. Figure 17 illustrates the simulated labor supply elasticities. As we do not model intensive margin changes (hours are assumed to be fixed by age and health status), these elasticities are extensive margin elasticities. The average labor supply elasticity is 0.65 and U-shaped over the life cycle, consistent with recent findings in Erosa et al. (2016).[45]

5 Quantitative Analysis

We now use the calibrated model to study the labor market effects of the DI program, the role of accounting for imperfect substitutability, and the value of the policy.[46]

5.1 Labor Market Effects of DI

To evaluate the labor market impact of the DI program in the U.S., we simulate an economy without DI, imposing budget-neutrality using lump-sum transfers. In Figure 18, we plot the percentage point (*pp*) changes in the employment rates of workers. When the DI program is removed, employment rates increase, with magnitudes larger for older workers whose employment rates in the benchmark economy (with DI) are low. Depending on age and health status, we see that the magnitude of increases in employment range widely, with the highest increase around 15*pp*. As a result, Figure 19 that plots experience distribution of workers aged between 60 and 65 in the benchmark and counterfactual economies shows a significant shift of the experience distribution to the right.

[44] Maestas et al. (2013) uses the behaviors of rejected DI applicants to estimate the labor supply effects of DI. They suggest that applicants who are rejected are more likely to work after two years.

[45] Erosa et al. (2016) study the aggregate labor supply elasticities in a rich heterogeneous agents model. Additional to features similar to ours (e.g., life cycle, labor productivity shocks, fixed costs), they also model preference heterogeneity and non-linearity in earnings with respect to hours worked. Our average model-implied extensive margin elasticity from a temporary wage change is smaller than theirs (1.08). According to their decomposition exercises, this may well be due to the lack of preference heterogeneity in our model. Overall, however, we believe that our model is able to broadly replicate the key features of the labor elasticities, similar to their findings.

[46] In the counterfactual analyses, we focus on individuals older than 25, as there are very few disabled workers younger than 25 (in the data and therefore, simulated model).

Figure 18: Employment Changes by Health

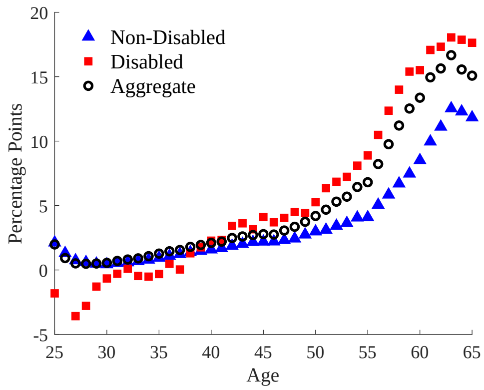


Figure 19: Distribution of Experience for Workers Older than 60

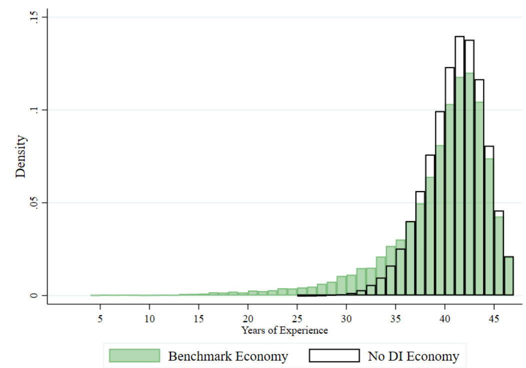


Table 8: Labor Market Effects of DI

	Change from DI to No-DI Economy		Change from DI to No-DI Economy
Relative Supply (E/L)	+1.01%	Employment	+4.88pp
Effective experience, E	+3.86%	Non-disabled	+3.00pp
Effective labor, L	+2.82%	Disabled	+13.81pp
Relative Price (R_E/R_L)	-2.60%	Wage	+1.25%
Price of experience,R_E	-2.00%	Non-disabled	+0.79%
Price of laborR_L	+0.61%	Disabled	+3.68%
Effective labor per worker	-2.96%	Output per worker	-2.61%
Effective experience per worker	-1.94%	Output	+3.06%

Note: This table reports the changes in labor market variables when the DI program is removed. The aggregate statistics are constructed based on simulation data with $n = 50,000$ for each age and aggregated using the share of population so that the size of the population is normalized to one.

Therefore, as summarized in Table 8, the complete removal of the DI program increases the aggregate employment rate by 4.88pp, 3pp for the non-disabled and 13.81pp for the disabled. These in turn lead to higher supplies of both effective labor and experience in the economy.[47] With the disproportionate increase of older workers in the labor force, the relative supply of experience increases due to their high amount of effective experience, even for disabled older workers. The latter effect can be observed from comparing the changes in the average effective labor and average effective experience per worker. While per-worker effective labor decreases by 3% in the no-DI economy due to the entry of less productive disabled workers,

[47]While we model the concurrent disutility of work and its impact on future wages (through increased experience), we do not model that work might impact the health of workers. The research on this issue is inconclusive. Case and Deaton (2005), for example, reports that self-reported health status worsens for workers, particularly for those in manual occupations and have lower health. On the other hand, there are others (e.g., Schaller and Stevens, 2015; Sullivan and von Wachter, 2009) that find that job loss leads to higher mortality and worse self-reported health. We acknowledge that our abstraction from this additional channel could bias our quantitative results: over (under)-estimating the increase in aggregate supply of inputs, if working deteriorates (improves) health.

the decrease in per-worker effective experience is smaller at less than 2%. This is driven by the smaller impact of disability on experience, one of our empirical findings from Section 2.3. As a result, under the economy without DI, the price (marginal product) of experience drops, while the price of labor increases.

Figure 20 presents the impacts of these price changes on the value (e.g., supply evaluated at equilibrium price) of workers' human capital over the life cycle. In it, we plot both the percent changes in the supply of experience and labor relative to the DI economy and the percent changes in their values. While the average supply of both inputs increase by similar magnitudes, the changes in their values diverge, differentially impacting workers' values of human capital as seen from the wage effects in Figure 21. The impact of DI removal on non-disabled workers' wages is relatively mild but it increases as they get older, reflecting the increased amount of effective experience (see Figure 19). On the other hand, changes in disabled workers' wages vary over the life cycle: highest at early ages thanks to the increased price of labor, and increasing again when old thanks to the accumulated experience.[48]

Figure 20: Supply and Value of Labor and Experience

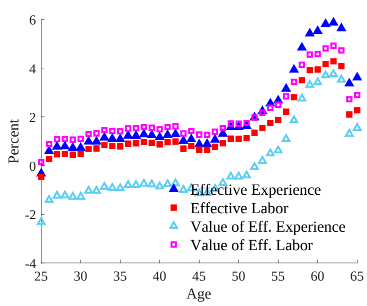


Figure 21: Wage Changes by Health

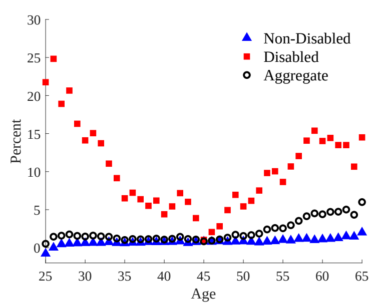
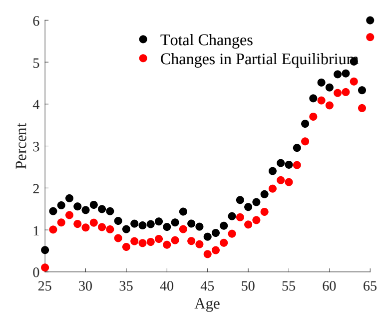


Figure 22: Decomposition of Wage Changes



To decompose the effects on wages, in Figure 22, we compare the total percent changes in the average wage plotted in Figure 21 with the changes in the average wage in the economy without DI, but keeping the prices of inputs as fixed (partial equilibrium). The partial equilibrium effects are dominant, explaining around 74% of the total wage changes. However, for young workers, with most of human capital in labor, a larger fraction of wage changes are driven by the price changes. For older workers, it is their increased accumulated experience that dominantly determine their wages, and thus, the price effect has a smaller role.[49] These show how the changes in the relative supplies and relative price of inputs, resulting from

[48]While the percent change in wage for disabled workers at early ages is large, the absolute magnitude is around \$2, from \$10 to \$12.

[49]Under our wage specification, work experience directly increases wages (through $g(e)$ in the wage equation). Thus, employment in the current period not only benefits workers through income this period but through higher wages in the future. In order to qualitatively understand the impact of the latter (future wage benefits of work), we conducted a counterfactual experiment where the worker i's experience level does not directly impact his wage, i.e., by letting $g(e_i) = g(\bar{e}(j,h,s))$ for all individuals i of age j, health h, and education s (thus, accumulated experience level is equalized across all workers with the same age, health, and

the removal of the DI program, impacts workers of all ages and health spectrum. The average effects of removing the DI program are further summarized in Table 8. The average wage effects are positive, smaller for non-disabled workers compared to disabled workers.[50]

Lastly, with the increased supply of workers, the total output of the economy increases by 3%, but aggregate productivity, as measured by output per employed worker, decreases by 2.61%, which is reflected in the decrease in the average per-worker productivities in labor and experience.

5.2 Role of Imperfect Substitutability

Now, we aim to evaluate the role of the imperfect substitutability between labor and experience on analyzing the labor market effects of DI. To do so, we abstract from the assumption that labor and experience are imperfectly substitutable in aggregate production.[51] Instead, we assume that these two inputs are perfectly substitutable by imposing that ρ, the parameter controlling the elasticity of substitution, is equal to one; thus, the aggregate production function now reads $Y = A\,(L + \theta E)$. We then re-calibrate the economy and conduct the counterfactual experiment of removing the DI program.[52]

Table 9: Labor Market Effects of DI under Perfect Substitutability between Inputs

	Change from DI to No-DI			Change from DI to No-DI	
	Benchmark	$\rho = 1$		Benchmark	$\rho = 1$
Relative Supply (E/L)	+1.01%	+1.06%	Employment	+4.88pp	+4.84pp
Effective experience, E	+3.86%	+3.57%	Non-disabled	+3.00pp	+2.91pp
Effective labor, L	+2.82%	+2.49%	Disabled	+13.81pp	+14.00pp
Relative Price (R_E/R_L)	-2.60%	-	Wage	+1.25%	+0.82%
Price of experience,R_E	-2.00%	-	Non-disabled	+0.79%	+0.37%
Price of laborR_L	+0.61%	-	Disabled	+3.68%	+3.16%
Effective labor per worker	-2.96%	-3.12%	Output per worker	-2.61%	-2.92%
Effective experience per worker	-1.94%	-2.09%	Output	+3.06%	+2.69%

education). We find that under such an environment, the wage increases of old, non-disabled workers are negligible—when wages depend on experience, the removal of DI induces more work and thus endogenously increases wages. However, as this experience channel on wage is removed, wage benefits for the old disappear. On the other hand, it may benefit disabled workers, who faced lower wages due to their lack of experience under the benchmark wage specification. When DI is removed under experience-independent wage profile, young disabled workers' wages increase more as the price of labor increases, and they are not penalized for the lack of experience. But the old disabled workers' wage increase is smaller as the price of experience decreases without them being compensated for their accumulated experiences.

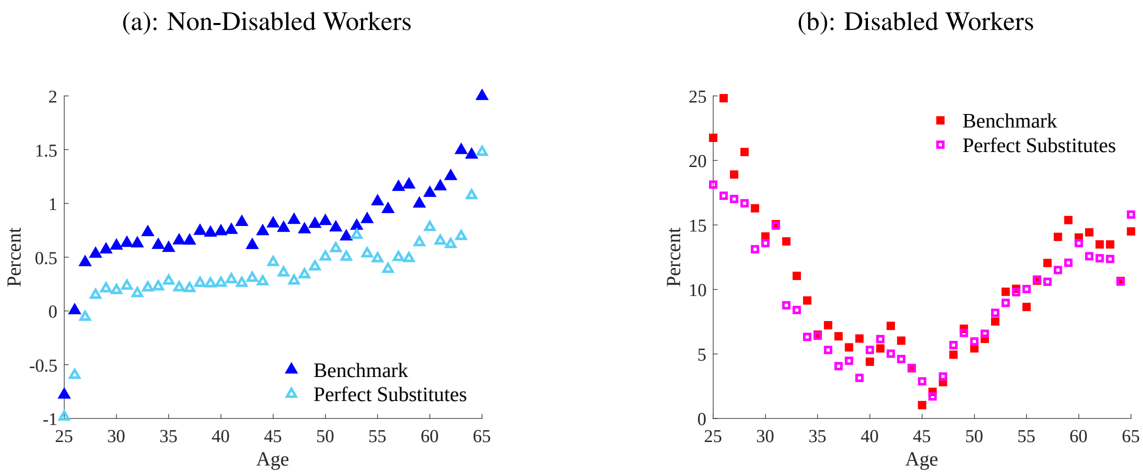
[50]We show these results by education status in Appendix C.2. In general, old, high school workers experience the largest employment and wage effects from the removal of DI.

[51]Although the two inputs are perfectly substitutable, we still maintain the assumption that individuals are endowed with labor and experience, which evolve over the life cycle.

[52]We start out by re-estimating the wage equation (Equation (2)), with the restriction that the experience premium, $\Pi_{E,t}$, is equal to one for all years, the implication of imposing $\rho = 1$. Then, we use the re-estimated wage equation parameters and re-calibrate the model with an additional parameter θ, the relative efficiency of experience in aggregate production (which, for the benchmark model, we obtained from Equation (3)). We summarize the calibrated parameters of this economy in Appendix C.3.

Table 9 compares the counterfactual results under $\rho = 1$ with those under the benchmark production specification. First, note that in the economy with perfect substitutability of inputs, the marginal products (prices) of experience and labor are independent of the supply changes. With no adjustment in prices, the relative supply of experience increases more with the per-worker efficiency of inputs dropping further. In both economies, we observe similar employment effects.

Figure 23: Wage Effects by Health Status



In Figure 23(a), we plot the wage changes in the benchmark and under the perfect substitutes economy. Compared to the benchmark counterfactual analysis, non-disabled workers' wage changes are smaller, especially at younger ages, as the increase in the price of labor is absent. In comparison, the differences are smaller for old workers and disabled workers. This implies that the lack of the input complementarity may result in the under-estimation of the impact of the DI program on young, healthy workers.

Lastly, we compare the output and output per worker effects of removing the DI program. Despite similar employment effects, the output increase from removing the DI program is smaller, while output per worker decreases more in the perfect substitutes economy. This implies that when we account for the complementarity between labor and experience, the productivity of the workforce does not decline as much as when we ignore the linkages between these two inputs. Put differently, thanks to the complementarity between old (experience) and young (labor) workers, the increased supply of old workers, relative to the perfect substitutes economy, may lead to the relatively higher output per worker in equilibrium.

5.3 Value of DI

We now analyze the value of the DI program by conducting the following counterfactual experiment. For each worker of age j, we make the DI program unavailable for one period so that the worker's labor market choices are restricted to between working and not working. Further, the wage rate of the individual becomes what it would have been in the no DI economy presented in Section 5.1. All other aspects of the model are identical to the benchmark environment up to age $j-1$ and starting again at age $j+1$.[53] Then, we calculate the consumption equivalent variation (CEV) as the percentage of consumption a worker of age j needs to be compensated to be as well off as in the economy with DI[54]: thus, the CEV represents how valuable the DI program is for a worker of age j, accounting for both the insurance benefit of DI and the labor market effects of removing the DI program.

Figure 24: Value of DI by Disability

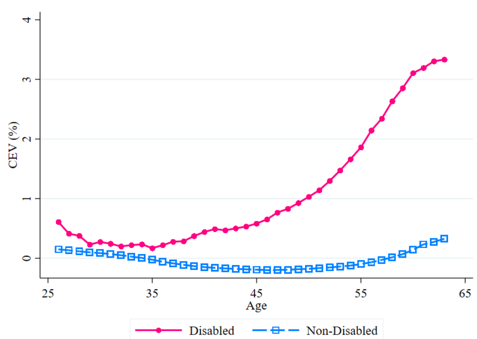


Table 10: CEV (%) by Subgroups

A. By health and education			
Non-disabled		Disabled	
High School	College	High School	College
0.12	-0.15	2.55	0.40

B. Disabled workers only: by labor market status at $t-1$			
Employed	Unemployed	DI status	
		Applied	Received
0.44	0.57	4.54	15.37

In Figure 24, we illustrate the CEVs by age and health status. We find that the overall welfare change is small for those who are not disabled, but it increases with age as they become more exposed to disability risks. For disabled workers, the value of DI is higher, increasing to 4% as they get older.

We further show that these average statistics mask large heterogeneities in the value of DI, with older and poorer workers in general, having higher welfare benefits from having the DI program.[55] Panel A of Table 10 reports the CEVs by education and health status. Non-disabled, college-educated workers have negative CEVs, implying that they prefer not having the DI program. This effect is mostly driven by their preference towards the wage structure in the economy without DI. Other workers, especially the low-educated, disabled workers have higher welfare from living in the economy with DI. Further, in Panel B are the CEVs by labor market status in the previous period of disabled workers. Overall, workers who are involved in the DI program, either as an applicant or a recipient, have higher CEVs. For DI recipients, losing the program

[53]We consider this experiment more suitable for measuring the value of DI, rather than calculating the welfare in the economy with complete removal of the DI program. The complete removal of the program is a large reform that also leads to, for example, significant changes in government budget and the lump-sum transfers that workers receive. These confounding factors make it difficult to isolate the welfare effects from the removal of DI alone.

[54]Let the utility of worker age j in the benchmark economy be $\bar{V}_j$; and in the counterfactual economy, $\tilde{V}_j$, where in age j, the DI program is removed and wage rate is adjusted. Then, consider a proportional consumption increase of Δ_j to this worker in every period (from today onwards) in the counterfactual economy, which given our utility preferences equals $(1+\Delta_j)^{1-\gamma}\,\tilde{V}_j$. Then we solve for Δ_j such that $\bar{V}_j = (1+\Delta_j)^{1-\gamma}\,\tilde{V}_j$.

[55]In Appendix C.4, we show the CEVs by asset, health, and age.

leads to a large average welfare loss reaching 15% of consumption.

6 Conclusion

The Social Security Disability Insurance program is an important social safety net for workers facing disability risks. However, empirical findings suggest that it creates sizable disincentives for the labor supply of workers. Our goal in this paper is to understand the aggregate implications of DI. Toward that goal, we estimated the productivity effects of disability regarding the two kinds of human capital possessed by a worker, (pure) labor and experience, and the interaction of these two inputs in aggregate production. One of the key empirical findings is that while disability lowers overall productivity, it is less detrimental to the productivity of older workers whose human capital primary consists of experience (than labor). Our counterfactual analyses from a calibrated life-cycle model of workers are used to evaluate the impact of removing the DI program and to measure the value of DI to workers of heterogeneous characteristics. Removal of the DI program has broad effects on the labor market, increasing the wages of young workers through general equilibrium effects, and those of older workers through an increase of accumulated experiences. Also, the aggregate productivity effects of removing the DI program may be smaller, when we account for heterogeneous human capital, thanks to the complementarity between human capital. The welfare benefits of the DI program are heterogeneous, with higher valuations from the poor and older workers.

Analyzing the effects of policies with aggregate interactions between heterogeneous human capital (inputs) modeled in this paper is not limited to the context of DI. The recent demographic changes from the aging of population in the U.S. would impact the relative supply of labor and experience, affecting workers across all ages and the aggregate productivity of the workforce. Thus, it would be interesting to study the role of policies that influence the labor supply decisions of workers such as an increase in the mandatory retirement age or changing the Social Security payment schedule within our model framework. We leave these important questions to future research.

References

Attanasio, O., S. Kitao, and G. L. Violante (2011). *Demography and the Economy*, Chapter Financing Medicare: A General Equilibrium Analysis, pp. 333–366. University of Chicago Press.

Autor, D., A. Kostol, M. Mogstad, and B. Setzler (2019, July). Disability benefits, consumption insurance, and household labor supply. *American Economic Review 109*(7), 2613–2654.

Bound, J. (1989). The health and earnings of rejected disability insurance applicants. *American Economic Review 79*, 482–503.

Braun, R. A., K. Kopecky, and T. Koreshkova (2017). Old, sick, alone and poor: A welfare analysis of old-age social insurance programs. *The Review of Economic Studies 84*, 580–612.

Card, D. and T. Lemieux (2001, May). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics 116*(2), 705–746.

Case, A. and A. Deaton (2005). *Broken Down by Work and Sex: How Our Health Declines*. Chicago: University of Chicago Press.

Chen, S. and W. van der Klaauw (2008). The work disincentive effects of the disability insurance program in the 1990s. *Journal of Econometrics 142*, 757–784.

Congressional Budget Office (2016). Social Security Disability Insurance: Participation and spending. CBO Publication No.51443.

Currie, J. and J. Gruber (1996a). Health insurance eligibility, utilization of medical care, and child health. *The Quarterly Journal of Economics 111*(2), 431–466.

Currie, J. and J. Gruber (1996b, December). Saving babies: The efficacy and cost of recent changes in the medicaid eligibility of pregnant women. *Journal of Political Economy 104*(6), 1263–1296.

De Nardi, M., E. French, and J. B. Jones (2016, November). Medicaid insurance in old age. *American Economic Review 106*(11), 3480–3520.

De Nardi, M., S. Pashchenko, and P. Porapakkarm (2018). The lifetime costs of bad health. *NBER Working Paper No. 23963*.

Erosa, A., L. Fuster, and G. Kambourov (2016, July). Towards a micro-founded theory of aggregate labour supply. *Review of Economic Studies 83*(3), 1001–1039.

Feenberg, D. R. and E. Coutts (1993). An introduction to the taxsim model. *Journal of Policy Analysis and Management 12*(1), 189–194.

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

French, E. and J. Song (2014, May). The effect of disability insurance receipt on labor supply. *American Economic Journal: Policy 6*(2), 291–337.

Gruber, J. and K. Milligan (2010). *Social Security Programs and Retirement around the World: The Relationship to Youth Employment*, Chapter :Do Elderly Workers Substitute for Younger Workers in the United States?, pp. 345–360. University of Chicago Press.

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

Hubbard, R., J. Skinner, and S. P. Zeldes (1995). Precautionary savings and social insurance. *Journal of Political Economy 103*, 360–399.

Imrohoroglu, S. and S. Kitao (2012, July). Social Security reforms: Benefit claiming, labor force participation, and long-run sustainability. *American Economic Journal: Macroeconomics 4*(3), 96–127.

Jeong, H., Y. Kim, and I. Manovskii (2015). The price of experience. *American Economic Review 105*(2), 784–815.

Karabarbounis, L. and B. Neiman (2014, January). The global decline of the labor share. *The Quarterly Journal of Economics 129*(1), 61–103.

Katz, L. F. and K. M. Murphy (1992, February). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics 107*(1), 35–78.

Kitao, S. (2014). A life-cycle model of unemployment and disability insurance. *Journal of Monetary Economics 68*, 1–18.

Krusell, P., L. E. Ohanian, J.-V. Rios-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica 68*, 1029–1054.

Low, H. and L. Pistaferri (2015, October). Disability insurance and the dynamics of the incentive-insurance tradeoff. *American Economic Review 105*(10), 2986–3029.

Low, H. and L. Pistaferri (2019). Disability insurance: Error rates and gender differences. *Working Paper NBER Working Paper Series*(26513).

Maestas, N., K. J. Mullen, and A. 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, A. H. and A. Y. Wu (2012, October). Will delayed retirement by the baby boomers lead to higher unemployment among younger workers? *Working Paper, Center for Retirement Research at Boston College*.

Pashchenko, S. and P. Porapakkarm (2017). Work incentives of Medicaid beneficiaries and the role of asset testing. *International Economic Review 58*(4).

Schaller, J. and A. H. Stevens (2015). Short-run effects of job loss on health conditions, health insurance, and health care utilization. *Journal of Health Economics 43*, 190–203.

Social Security Administration (2013). Annual statistical report on the Social Security Disability Insurance program.

Social Security Administration (2017, October). Annual statistical report on the Social Security Disability Insurance program. SSA Publication No.13-11826.

Sullivan, D. and T. von Wachter (2009). Job displacement and mortality: An analysis using administrative data. *Quarterly Journal of Economics 124*(3), 1265–1306.

Appendix

A Data Appendix

A.1 Census Population Estimates

We use the 2014 version of the Population Projections Program, obtained from the U.S. Census Bureau. The Population Projections Program provides projected estimates of demographic composition 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 25: Population Share by Age

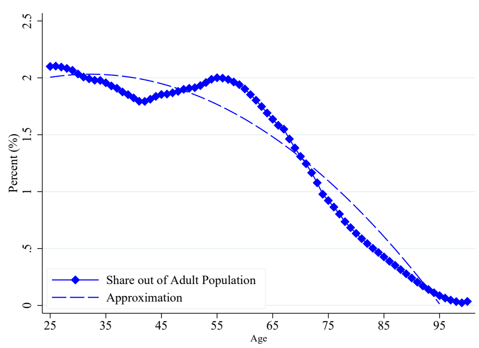


Table 11: Summary Statistics

	Non-Disabled		Disabled	
	Mean	SD	Mean	SD
Female	0.52	0.50	0.56	0.50
Age (yr)	41.16	12.20	47.80	12.17
Schooling (yr)	13.70	2.10	12.83	2.40
Subjective health	2.12	0.89	3.46	1.06
Employment	0.80	0.40	0.47	0.50
Annual working hours	1,743.04	922.67	975.94	1,037.46
Experience (yr)	11.46	10.08	11.73	10.59
Number of obs.	90,713		14,511	

Note: Table 11 presents the summary statistics, weighted by individual survey weights. Subjective health measure is a category variable ranging from 1 (excellent) to 5 (very poor).

A.2 Variable Construction for Wage Process Estimation

Sample Selection Criteria. We use the PSID as our main source of data for wage process estimation. Our sample consists of individuals of working age (between 18 to 65), both male and female, at all education levels. Since Jeong et al. (2015) focus on *employed* workers' wage process, their sample consists of individuals with more than 700 working hours per year. However, our sample includes working-age individuals, regardless of their employment status, because we are interested in changes in extensive margin with respect to policy variation.

We exclude observations missing key information on disability status. We take the self-reported measure of one or more work-limiting health problems in the PSID as our indicator of disability status. We also drop observations missing key information such as age, schooling, and years of experience if we were unable to fill those gaps even after exploring the past observations in panel data. Table 11 reports the summary statistics by disability status. Note that even though disabled workers are 6.6 years older than the non-disabled, their the average work experience exhibits no statistically significant difference.

Construction of the Experience Variable. One of the key variables in our empirical analysis is the years of work experience. Similar to Jeong et al. (2015), we take as its basis the number of years reported by the

Table 12: Other Statistics by SSDI Status

Variable	SSDI Recipients	Non-SSDI Recipients
age	51.17 (10.87)	40.53 (12.08)
schooling (year)	11.94 (2.55)	13.58 (2.13)
marital status (married)	0.44 (0.50)	0.60 (0.49)
health score in scale 1(excellent) to 5 (poor)	3.80 (1.04)	2.25 (0.99)
work limitation	0.823 (0.38)	0.11 (0.31)
experience (year)	12.71 (10.57)	11.00 (10.07)
mean assets	71,559 (270,189.1)	175,274.1 (971,452.2)
median assets	10,266.93 (270,189.1)	38,469.67 (971,452.2)
No. of obs.	1,290	50,115

PSID (which directly asked respondents for years of prior work) and construct the experience variable by adding experience when an individual reported working hours above a threshold.[56]

Jeong et al. (2015) take 700 hours as the cutoff value for accumulating a year of experience, which is consistent with their sample selection criteria. In our analysis, we use a less strict measure and consider that a worker earns a year of experience when he works at least 250 hours per year. When we observe an individual whose first experience variable was larger than one he had at the age of 18 or older, we construct the experience variable retrogressively in time for his younger working life. Starting from 1999, the PSID changed its survey frequency from annual to biennial. Accordingly, we adjust the gap and add two years of experience when he worked full time in the past year. Figure 26 presents the distribution of the experience variable in our sample by education, gender, and disability status. Again, Figure 26 confirms that the similarity in average experience for non-disabled and disabled workers is mainly driven by composition: conditional on age, we see sizable differences in work experience by health status. We also construct the accumulated working hours, and compare the yearly measured experience variables. Table 13 shows that the measures share quantitatively similar features in their life-cycle properties by disability status.

Table 13: Accumulated Experience and Hours in the PSID by Disability Trajectory

Variable		Accumulated Working Hours			Accumulated Years of Experience		
Disability History		less than 20%	more than 20%	ratio (%)	more than 20%	less than 20%	ratio (%)
Age	18-29	6,730	5,946	88.4	4.6	3.8	82.6
	30-39	18,708	14,992	80.1	10.8	8.9	82.4
	40-49	35,433	26,827	75.7	18.3	14.0	76.5
	50-59	48,123	34,523	71.7	25.1	17.7	70.5

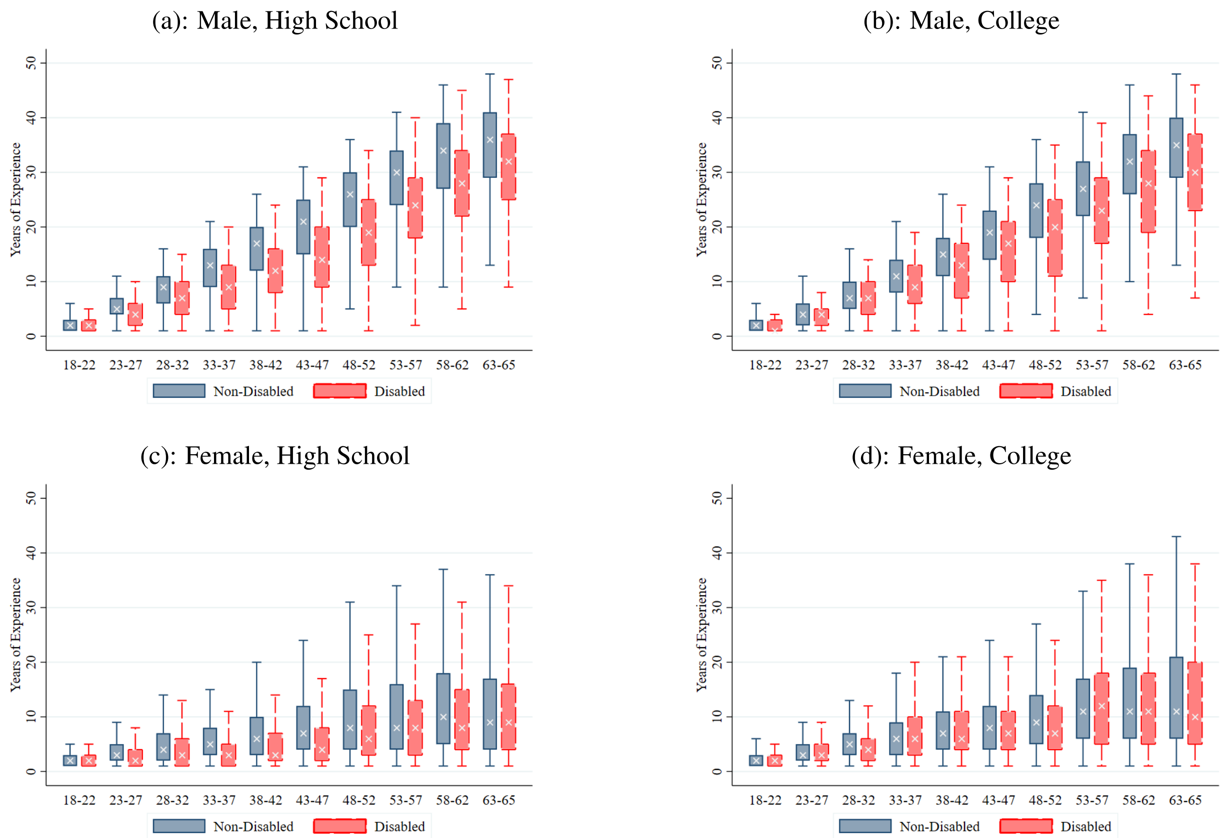
B Estimation of Wage Equation

B.1 First Stage

Construction of Potential Benefits and Taxes. To address the selection bias problem in our wage estimation, we adopt Heckman's two-stage estimation (Heckman, 1979) and run a probit regression using potential

[56] The PSID asked for the number of years of experience in 1974, 1975, 1976, and 1985 for every head of household and wife of household. In subsequent sample years, the PSID has collected this information for new heads and wives.

Figure 26: The Experience Variable in the PSID



Note: The charts in Figure 26 illustrate the life-cycle patterns of the years of experience by disability, education, and gender. Solid lines (gray) and dashed lines (red) represent the 95% distribution of the non-disabled and the disabled observations, respectively. Boxes show the range of samples between 25% and 75% percentiles. Marker × denotes the median years of experience for each age group.

government transfers and taxes as our exclusion restriction. Similar to the simulated IV method in Currie and Gruber (1996a,b) and Low and Pistaferri (2015), we construct the "magnitude of potential benefits" from the state government and the interaction of those benefits with disability status as our exclusion restrictions. Unlike the actual transfer amounts, which are endogenous, these potential benefits are exogenous by default.

Following Low and Pistaferri (2015), we compute the potential benefits for a representative household enrolled in each federal or state-level welfare program based on the following welfare programs: the Earned Income Tax Credit (EITC), Unemployment Insurance (UI), the Supplemental Nutrition Assistance Program (SNAP), Aid to Families with Dependent Children (AFDC), and Temporary Assistance for Needy Families (TANF). We start from the welfare benefit calculations available in Online Appendix C.1 of Low and Pistaferri (2015) and update the benefit formulas when more recent policy changes occurred. We apply each policy formula to a representative household, and compute the potential benefit amounts for the years from 1983 to 2016.

To construct the potential tax liabilities by state and year, we use the NBER TAXSIM program v.27, which calculates federal and state income taxes given a household's financial circumstances. As a first step, we construct a financial statement of a representative household. To do so, we merge 11 waves of the Survey of Consumer Finances (SCF) from the years 1983 to 2013 with the PSID.[57] The SCF is a triennial cross-sectional survey providing rich information on the financial status of U.S. households. This survey provides information on earnings (including business income, dividends, and capital gains) by source. The SCF also includes respondents' mortgage balance and payment records, which we use to approximate mortgage interest payments.[58] We combine the SCF with the PSID, which contains variables such as childcare expenses, UI and SSI benefit payments, rents, and house prices. Conjointly, we construct a profile of a representative household for tax-filing via NBER TAXSIM. For tax liability calculations, we use the nominal values of expenditures and earnings variables from both the SCF and PSID. We convert these tax liabilities into 2011 U.S. dollar using the CPI before we estimate a probit regression.

Over-Identification Test. We include potential tax liabilities as our instrumental variables along with potential welfare benefits, because we believe that having a wide range of potential benefits/taxes as our exclusion restrictions could be useful because our sample includes both high school and college graduates.[59]

Intuitively, we can infer the validity of our instruments based on the coefficient estimation results. Our benchmark analysis includes two types of instruments, potential benefits and taxes by disability status. Table 14 shows that the potential tax liabilities are statistically significant across education and disability groups. We also find that more generous benefit programs are negatively related to employment.

Table 14: Coefficient Estimation Results from the First Stage: Instrumental Variables

Dependent variable: employment	high school + college		high school only		college only	
potential benefits	-0.0086***	(0.0017)	-0.0072**	(0.0029)	-0.0091***	(0.0021)
potential benefits×disability	-0.0031	(0.0071)	0.0032	(0.0102)	-0.0088	(0.0100)
potential taxes	0.000032***	(8.3e-06)	0.00003**	(0.00001)	0.00004***	(0.00001)
potential taxes×disability	0.000054***	(0.00002)	0.00006**	(0.00003)	0.00005*	(0.00003)

Table 15: Over-Identification Test for Labor Supply Decision

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	high school + college			high school only			college only		
J-test	3.2973	0.1567	0.2471	0.9563	0.1801	0.0014	3.7006	0.0081	0.2551
p-value	0.5094	0.6923	0.6191	0.8118	0.6713	0.9706	0.2957	0.9283	0.6135
potential benefits	×	×		×	×		×	×	
potential taxes	×		×	×		×	×		×
number of obs.		83,523			36,223			47,300	

We examine whether our choice of exclusion restrictions is proper using the J-test, which evaluates the null hypothesis that our additional instrument is structurally correlated with error terms. For computational simplicity, instead of the nonlinear wage equation discussed in the main text, the test statistics are derived based on a standard linear log-wage equation. As reported in Table 15, we find that the null hypothesis is

[57] Since the PSID has been conducted biennially since 1997, these two surveys are simultaneously available every six years. In our case, except for the years 2001 and 2007, we decided to merge the two data sets by matching the most recent SCF to the PSID. Thus, some components of taxable incomes from the SCF may have, at most, a one-year gap with the variables in the PSID. This gap has no specific direction in the sense that it could be either proceeding or lagging.

[58] Specifically, income variables include WAGEINC, BUSSEFARMINC, INTDIVINC, KGINC, INCOME, and SSRETINC. These are wage and salary income, business income, interest, capital gains/losses, family income, and pensions, respectively. Mortgage balances, house value, and mortgage payments (MORTPAY, HOUSES, and NH_MORT) are conjointly used to predict mortgage interest payments, assuming a standard 30-year mortgage schedule.

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

rejected, indicating that our instruments are jointly valid. We also check the validity of the instrumental variables by using the residuals from the second-stage wage estimation results.

Probit Estimation Results. Table 16 reports the probit regression results.

Table 16: The First-Stage Probit Regression Results

Variable	Coefficients		Variable	Coefficients	
disability	-0.8089***	(0.0319)	age	0.0026	(0.0030)
experience (e)	0.2117***	(0.0065)	age^2	-0.0009***	(0.0001)
e^2	-0.0092***	(0.0004)	married	-0.0827***	(0.0223)
e^3	0.0001***	(6.64e-6)	male	0.1158***	(0.0220)
years of schooling	0.0706***	(0.0048)	black	-0.0637***	(0.0241)
Number of obs.			101, 414		
Pseudo R^2			0.2252		

Note: Table 16 reports the first-stage probit regression results of Heckman's two-stage estimation. The dependent variable is the employment status of an individual. Independent variables also include year dummies. We use individual-level survey weights for our analysis. Standard errors are clustered at the individual level and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

B.2 Nonlinear Wage Equation Estimation

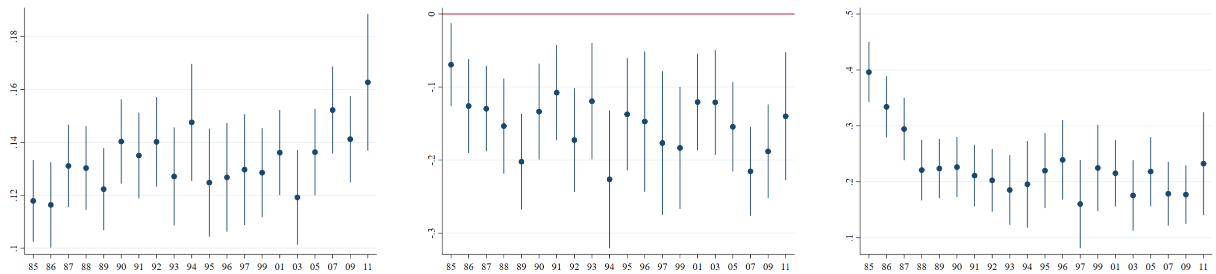
Table 17: The Effect of Disability on Wage: With and Without Selection Control

Coefficients		(1) benchmark	(2)
Inverse Mils Ratio		0.2662 (0.0891)	
Labor Profile	$\lambda_{L,1}\,(HS)$	0.0213 (0.0052)	0.0198 (0.0050)
	$\lambda_{L,2}\,(HS)$	-0.0004 (0.0001)	-0.0002 (0.0001)
	$\lambda_{L,0}\,(Col)$	-0.2467 (0.0555)	-0.2455 (0.0547)
	$\lambda_{L,1}\,(Col)$	0.0526 (0.0056)	0.0508 (0.0001)
	$\lambda_{L,2}\,(Col)$	-0.0010 (0.0001)	-0.0008 (0.0001)
	$\ln\phi_L\,(HS)$	-0.3083 (0.0838)	-0.1464 (0.0578)
	$\ln\phi_L\,(Col)$	-0.4614 (0.0787)	-0.3094 (0.0566)
Experience Profile	$\lambda_{E,1}\,(HS)$	0.0048 (0.0137)	0.0034 (0.0203)
	$\lambda_{E,2}\,(HS)$	-0.0003 (0.0003)	-0.0003 (0.0004)
	$\lambda_{E,0}\,(Col)$	-0.3748 (0.1785)	-0.3762 (0.2456)
	$\lambda_{E,1}\,(Col)$	0.0088 (0.0184)	-0.0070 (0.0267)
	$\lambda_{E,2}\,(Col)$	-0.0003 (0.0004)	-0.0001 (0.0005)
	$HS : \phi_E/\phi_L$	1.1847 (0.1941)	1.0611 (0.2318)
	$Col : \phi_E/\phi_L$	1.5606 (0.2602)	1.5037 (0.3320)
Accumulated	ζ_2	-0.0485 (0.0080)	-0.0544 (0.0111)
Experience	ζ_3	0.0012 (0.0004)	0.0017 (0.0005)
	ζ_4	-0.00001 (5.51e-6)	-0.00002 (7.64e-6)
Number of Obs.			83,532
Adjusted R^2		0.9981	0.9981

Note: Table reports the second-stage nonlinear wage equation estimation results with and without selection correction. The dependent variable is the log hourly wage rate of an individual. The control variables (other than those that are part of the functional specification) also include region and year-specific dummy variables for gender, race, and schooling (college). We use individual-level survey weights for our analysis. Standard errors are clustered at the individual level and reported in parentheses.

Figure 27: Estimation Results: The Dummy Variables

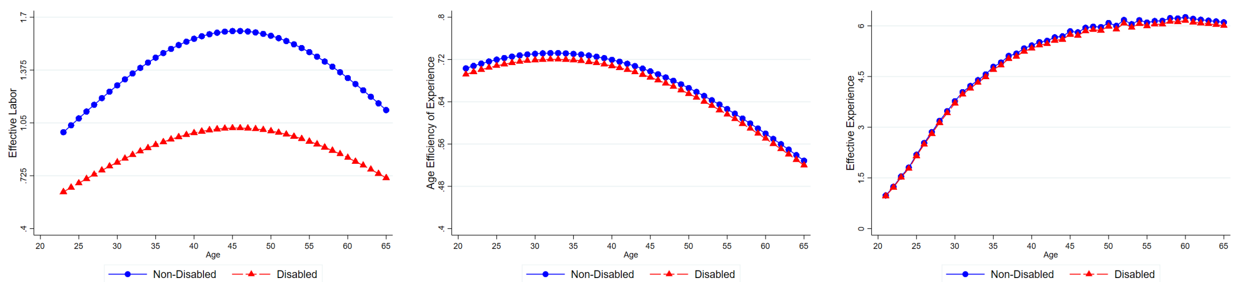
(a): Schooling: College (b): Race: Black (c): Gender: Male



Note: Figure 27 plots the estimated coefficients of dummy variables for race (black), education (college or more), and gender (male). The x-axis is the year of the sample, and the y-axis is the productivity measured by log hourly wage rate in 2011 U.S. dollar. Circular dots represent point estimates, and the lines are their 95% CI. Standard errors are clustered at the individual level.

Figure 28: The Efficiency Schedules over the Life-cycle: Workers with College Education

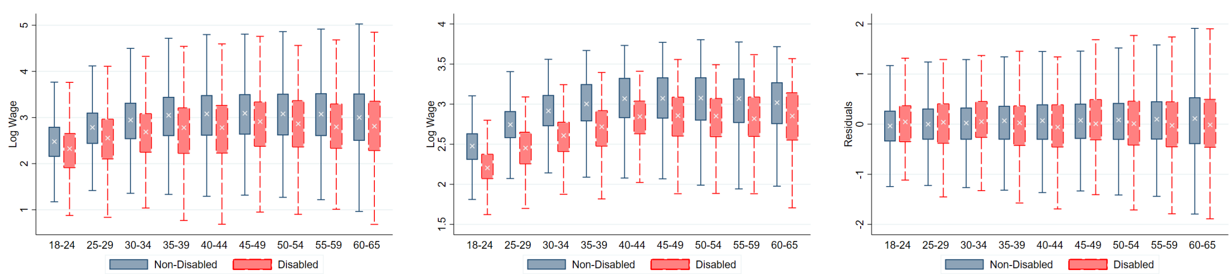
(a): $\lambda_L(Col, j, h)$ (b): $\lambda_E(Col, j, h)$ (c): $\lambda_L(Col, j, h)g(\bar{e}_j)$



Note: Figure 28 (a) and (b) plot the college graduates' life-cycle patterns of labor and experience by disability using the estimated coefficients of Equation (2). Figure 28(c) shows the empirical pattern of total amount of effective experience, applying the average years of experience to the estimated $\lambda_E\left(Col, j, h\right) g\left(\bar{e}_j\right)$ for each age. Circular (blue) and triangular (red) markers represent the profiles of the non-disabled and the disabled, respectively.

Figure 29: Estimation Results: Residuals

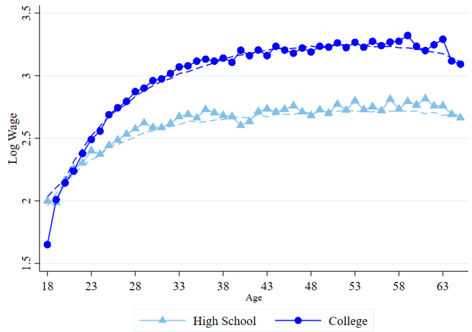
(a): Data: $\log(w)$ (b): Estimation: $\log(\hat{w})$ (c): Residuals



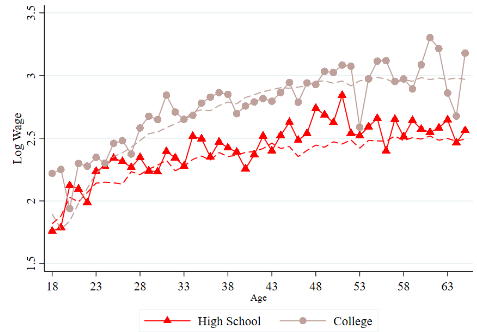
Note: Figure 29 presents the predicted log wages and residuals, along with the original data. The x-axis is the age groups, divided into nine from 18-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, and 60-65. The y-axis is the productivity measured by log hourly wage rate in 2011 U.S. dollar. Solid lines (grey) and dashed lines (red) represent the 95% distribution of the non-disabled and the disabled observations, respectively. Marker $\times$ denotes the median years of experience for each age group.

Figure 30: Estimation Results: Data Fit

(a): Non-Disabled Workers



(b): Disabled Workers



Note: Figure 30 plots the predicted log wages along with the original data. Lines with markers denotes data, and the dotted lines, the estimated wages by age.

Table 18: Robustness Analyses: Estimated Labor and Experience Efficiencies with Alternative Specifications

Coefficients		(1)	(2)	(3) benchmark	(4)
Labor Profile	$\lambda_{L,1}\,(HS)$	0.0396 (0.0041)	0.0393 (0.0041)	0.0213 (0.0052)	0.0214 (0.0053)
	$\lambda_{L,2}\,(HS)$	-0.0007 (0.0001)	-0.0007 (0.0001)	-0.0004 (0.0001)	-0.0004 (0.0001)
	$\lambda_{L,0}\,(Col)$	-	-	-0.2467 (0.0555)	-0.2384 (0.0565)
	$\lambda_{L,1}\,(Col)$	-	-	0.0526 (0.0056)	0.0524 (0.0058)
	$\lambda_{L,2}\,(Col)$	-	-	-0.0010 (0.0001)	-0.0009 (0.0001)
	$\ln\phi_L\,(HS)$	-0.3735 (0.0643)	-0.3121 (0.0818)	-0.3083 (0.0838)	-0.3049 (0.0836)
	$\ln\phi_L\,(Col)$	-	-0.3990 (0.0746)	-0.4614 (0.0787)	-0.4564 (0.0784)
Experience Profile	$\lambda_{E,1}\,(HS)$	0.0047 (0.0090)	0.0033 (0.0150)	0.0048 (0.0137)	0.0044 (0.0146)
	$\lambda_{E,2}\,(HS)$	-0.0003 (0.0002)	-0.0003 (0.0003)	-0.0003 (0.0003)	-0.0003 (0.0003)
	$\lambda_{E,0}\,(Col)$	-	-	-0.3748 (0.1785)	-0.4514 (0.1882)
	$\lambda_{E,1}\,(Col)$	-	-	0.0088 (0.0184)	0.0086 (0.0222)
	$\lambda_{E,2}\,(Col)$	-	-	-0.0003 (0.0004)	-0.0003 (0.0004)
	$HS:\phi_E/\phi_L$	1.3585 (0.1671)	1.3393 (0.2413)	1.1847 (0.1941)	1.1750 (0.1916)
	$Col:\phi_E/\phi_L$	-	1.379 (0.2235)	1.5606 (0.2602)	1.5596 (0.2633)
Accumulated	$\zeta_2\,(HS)$	-0.0485 (0.0086)	-0.0488 (0.0088)	-0.0485 (0.0080)	-0.0528 (0.0091)
Experience	$\zeta_3\,(HS)$	0.0012 (0.0004)	0.0013 (0.0004)	0.0012 (0.0004)	0.0015 (0.0004)
	$\zeta_4\,(HS)$	-0.00001 (5.83e-6)	-0.00001 (6.00e-6)	-0.00001 (5.51e-6)	-0.00001 (6.13e-6)
	$\zeta_2\,(Col)$	-	-	-	-0.0430 (0.0147)
	$\zeta_3\,(Col)$	-	-	-	0.0009 (0.0007)
	$\zeta_4\,(Col)$	-	-	-	-5.82e-6 (0.00001)
Education-Specific	ϕ_L and ϕ_E		$\times$	$\times$	$\times$
Components	λ_L and λ_E	$\times$		$\times$	$\times$
	$g\,(e)$				$\times$
Number of Obs.		83,532			
Adjusted R^2		0.9980	0.9981	0.9981	0.9981

Note: Table 18 reports the coefficient estimation results of the nonlinear wage equation with alternative specifications. The control variables (other than those that are part of the functional specification) include region and year-specific dummy variables for gender, race, and schooling (college). We use individual-level survey weights for our analysis. Standard errors in parentheses are clustered at the individual level.

Table 19: Robustness Analyses: Estimated Coefficients with Alternative Clustering

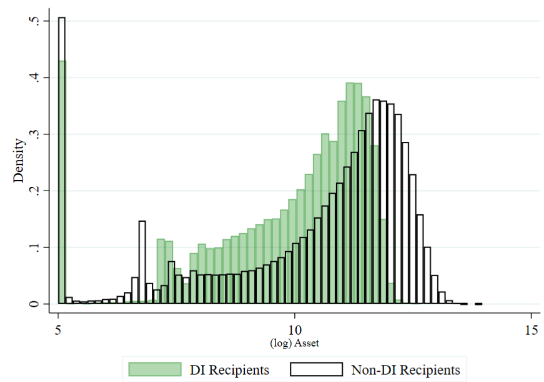
Coefficients		(1)	(2)	(3)	(4) benchmark
Inverse Mils Ratio		0.2662 (0.5579)	0.2662 (0.1270)	0.2662 (0.0554)	0.2662 (0.0891)
Labor Profile	$\lambda_{L,1}\,(HS)$	0.0213 (0.0041)	0.0213 (0.0053)	0.0213 (0.0037)	0.0213 (0.0052)
	$\lambda_{L,2}\,(HS)$	-0.0004 (0.0001)	-0.0004 (0.0002)	-0.0004 (0.0001)	-0.0004 (0.0001)
	$\lambda_{L,0}\,(Col)$	-0.2467 (0.0420)	-0.2467 (0.0622)	-0.2467 (0.0396)	-0.2467 (0.0555)
	$\lambda_{L,1}\,(Col)$	0.0526 (0.0041)	0.0526 (0.0053)	0.0526 (0.0033)	0.0526 (0.0056)
	$\lambda_{L,2}\,(Col)$	-0.0010 (0.0001)	-0.0010 (0.0001)	-0.0010 (0.0001)	-0.0010 (0.0001)
	$\ln \phi_L\,(HS)$	-0.3083 (0.0711)	-0.3083 (0.1102)	-0.3083 (0.0579)	-0.3083 (0.0838)
	$\ln \phi_L\,(Col)$	-0.4614 (0.0624)	-0.4614 (0.0806)	-0.4614 (0.0522)	-0.4614 (0.0787)
Experience Profile	$\lambda_{E,1}\,(HS)$	0.0047 (0.0090)	0.0047 (0.0129)	0.0047 (0.0085)	0.0048 (0.0137)
	$\lambda_{E,2}\,(HS)$	-0.0003 (0.0002)	-0.0003 (0.0002)	-0.0003 (0.0002)	-0.0003 (0.0003)
	$\lambda_{E,0}\,(Col)$	-0.3748 (0.1052)	-0.3748 (0.1584)	-0.3748 (0.1194)	-0.3748 (0.1785)
	$\lambda_{E,1}\,(Col)$	0.0088 (0.0133)	0.0088 (0.0202)	0.0088 (0.0108)	0.0088 (0.0184)
	$\lambda_{E,2}\,(Col)$	-0.0003 (0.0003)	-0.0003 (0.0003)	-0.0003 (0.0002)	-0.0003 (0.0004)
	$HS : \phi_E/\phi_L$	1.1847 (0.1564)	1.1847 (0.2326)	1.1847 (0.1354)	1.1847 (0.1941)
	$Col : \phi_E/\phi_L$	1.5606 (0.1731)	1.5606 (0.2145)	1.5606 (0.1796)	1.5606 (0.2602)
Accumulated	ζ_2	-0.0485 (0.0049)	-0.0485 (0.0066)	-0.0485 (0.0046)	-0.0485 (0.0080)
Experience	ζ_3	0.0012 (0.0003)	0.0012 (0.0003)	0.0012 (0.0002)	0.0012 (0.0004)
	ζ_4	-0.00001 (4.17e-6)	-0.00001 (4.87e-6)	-0.00001 (3.45e-6)	-0.00001 (5.51e-6)
Cluster	year	$\times$		$\times$	
	state		$\times$		
	id			$\times$	$\times$
Number of Obs.		83,532			
Adjusted R^2		0.9980	0.9981	0.9981	0.9981

Note: Table 19 reports the coefficient estimation results of the nonlinear wage equation with alternative clustering choices. The control variables (other than those that are part of the functional specification) also include region and year-specific dummy variables for gender, race, and schooling (college). We use individual-level survey weights for our analysis.

C Quantitative Analysis

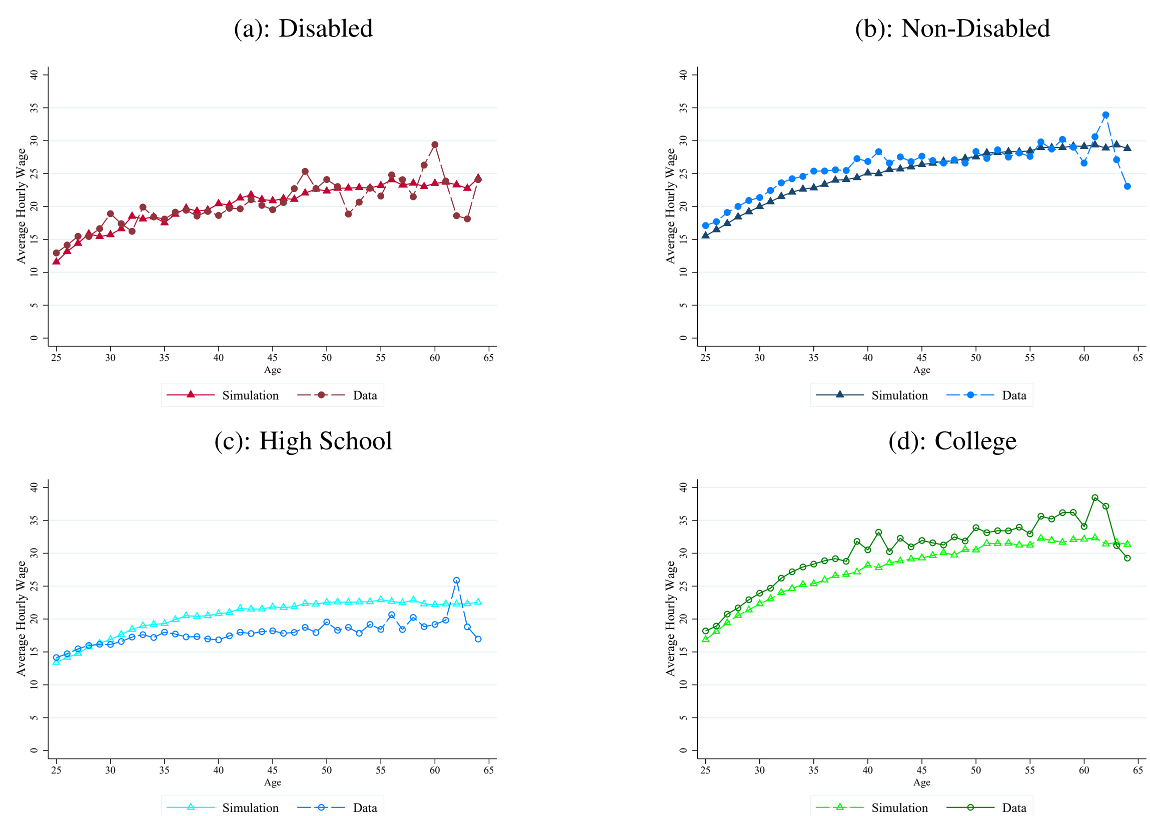
C.1 Calibration

Figure 31: Asset Distribution by DI Status



Note: In Figure 31, assets less than \$150 has been collapsed to \$150 for plotting purposes only.

Figure 32: Wage: Simulation vs. Data



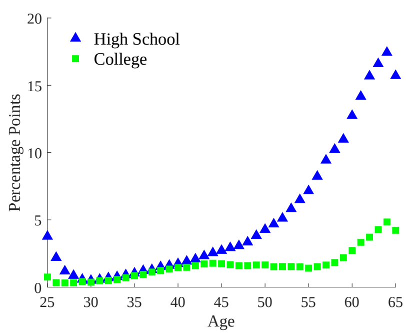
Note: Panels (a) and (b) compare the simulation results of wage over the life cycle to their empirical counterparts from the PSID by disability status. Panels (c) and (d) illustrate the wage profile by education group. For all panels, we use circular and triangular markers for data and simulation, respectively.

C.2 Labor Market Effects of DI: Effects by Education

In this section, we show the effects of removing the DI program by education, as measured by percentage point differences from the benchmark economy (DI economy) for employment and percent differences from the benchmark economy for wages. Overall, we see large employment effects from disabled workers with high school education (Figure 33). Similarly, the changes in wages are larger for disabled workers with low education (Figure 34). Thanks to accumulated experience as they approach old age, their wage increases in percentage terms increase as they near retirement and outpace those of college graduates. Next, we plot both the total changes in workers' wages and change in workers' wages when we remove DI but do not account for change in factor prices in Figure 35. The difference between the two lines, thus, represents the changes explained by the relative price effects of removing the DI program. As we see, while most of the changes in disabled workers' wages are due to the changes in their own labor inputs (mostly their accumulated experience), a relatively larger share of non-disabled workers' wage changes are explained by price effects, and more so for non-disabled workers with a college degree (45% for high school graduates, and 70% for college graduates).

Figure 33: Employment Effects by Education

(a): Employment: Non-Disabled Workers



(b): Employment: Disabled Workers

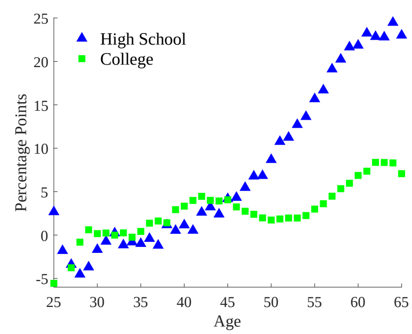
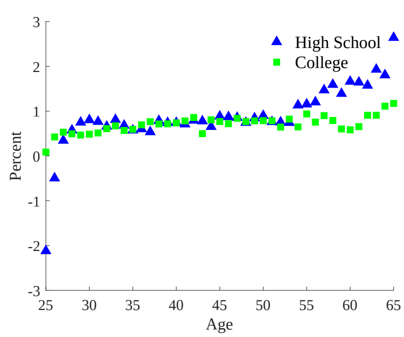
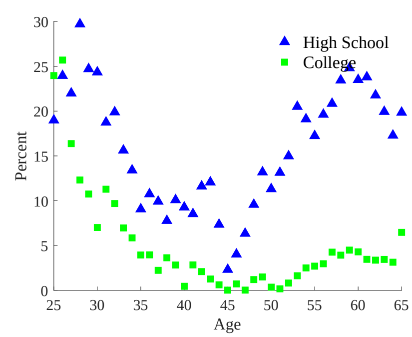


Figure 34: Wage Effects by Education

(a): Non-Disabled Workers



(b): Disabled Workers

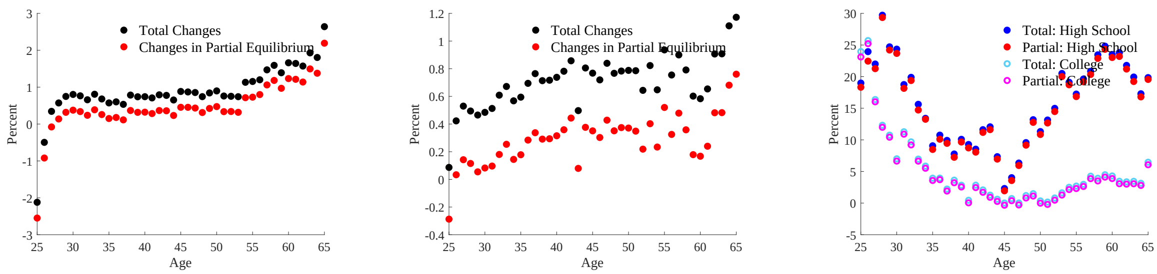


C.3 Role of Imperfect Substitutability

Our goal in Section 5.2 is to study the effects of incorporating the role of heterogeneous inputs into aggregate production to analyze how DI impacts labor market outcomes. To do so, we re-estimate the model with the assumption that $\rho = 1$, i.e., that labor and experience are perfect substitutes in aggregate production. We

Figure 35: Decomposition of Wage Effects

(a): High School, Non-Disabled (b): College, Non-Disabled (c): Disabled Workers



first, start out by re-estimating the wage equation (Equation (2)), with the restriction that the experience premium, $\Pi_{E,t}$, is equal to one for all years, which is the implication of imposing that $\rho = 1$. Then, we use the estimated wage equation parameters and re-calibrate the model with an additional parameter θ, the relative efficiency of experience in aggregate production (which, for the benchmark model, we obtained from Equation (3)). Tables 20 and 21 summarize the calibrated parameters in this economy.

Table 20: Estimated Coefficients of Wage Profile with Perfect Substitutability

Labor λ_L, High School		Experience λ_E, High School		Inverse Mills Ratio	
$\lambda_{L,1}\,(HS)$	0.0226 (0.0053)	$\lambda_{E,1}\,(HS)$	0.0021 (0.0139)		0.2639 (0.0904)
$\lambda_{L,2}\,(HS)$	-0.0004 (0.0001)	$\lambda_{E,2}\,(HS)$	-0.0002 (0.0002)		
$\ln\phi_L\,(HS)$	-0.2959 (0.0845)	$\phi_E\,(HS)\,/\phi_L\,(HS)$	1.1502 (0.1983)		
Labor λ_L, College		Experience λ_E, College		Accumulated Experience: $g\,(e)$	
$\lambda_{L,0}\,(Col)$	-0.2379 (0.0563)	$\lambda_{E,0}\,(Col)$	-0.3757 (0.1760)	ζ_2	-0.0489 (0.0079)
$\lambda_{L,1}\,(Col)$	0.0520 (0.0058)	$\lambda_{E,1}\,(Col)$	0.0087 (0.0184)	ζ_3	0.0013 (0.0004)
$\lambda_{L,2}\,(Col)$	-0.0009 (0.0001)	$\lambda_{E,2}\,(Col)$	-0.0003 (0.0004)	ζ_4	-0.00001 (0.000)
$\ln\phi_L\,(Col)$	-0.4510 (0.0800)	$\phi_E\,(Col)\,/\phi_L\,(Col)$	1.5157 (0.2612)		
Number of Obs.					83,532
Adjusted R^2					0.998

C.4 Value of DI

Figure 36 plots the CEV distributions by asset, and the share of population in specific asset levels by age and health. As seen in the figure, there is a significant share of workers with low assets, for whom DI is more valuable (high CEV).

Table 21: Parameter Values with Perfect Substitutability

Parameters	Description	Value			
A	Aggregate productivity	0.675			
θ	Efficiency of experience	0.049			
β	Time discount factor	0.953			
		High School		College	
		Non-Disabled	Disabled	Non-Disabled	Disabled
$\eta_{h,s}$	Disutility of work	-0.105	-0.193	-0.104	-0.190
$F_{h,s}$	Fixed cost of work	773	922	991	1,031
$\chi^W_{h,s}$	Offer arrival rates: employed	0.931	0.812	0.995	0.899
$\chi^U_{h,s}$	Offer arrival rates: unemployed	0.824	0.692	0.926	0.745
$\chi^A_{h,s}$	Offer arrival rates: applicants	0.768	0.600	0.899	0.690
χ^B_s	Offer arrival rates: DI beneficiaries	0.372	-	0.600	-

Figure 36: Value of Disability Insurance by Asset

(a): CEVs by Asset (b): Asset

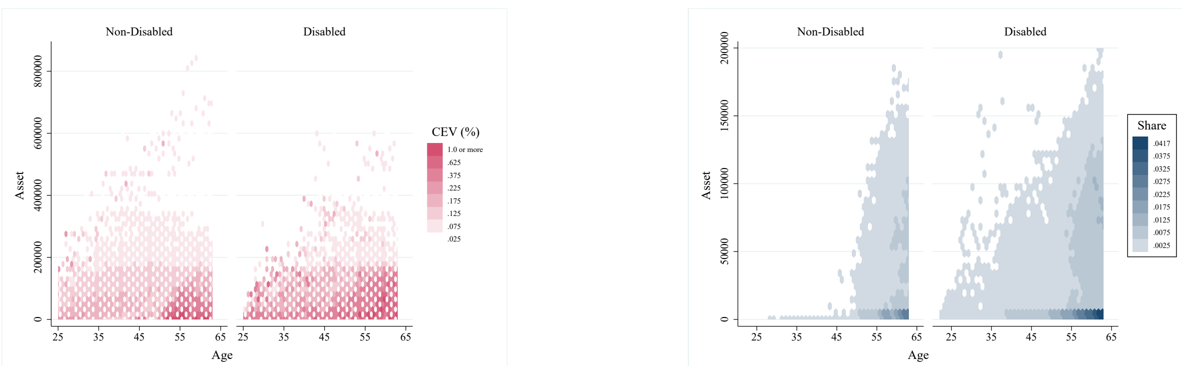


Figure 37: Value of Disability Insurance by Education

(a): Disabled (b): Non-Disabled