Institute for Fiscal Studies

Maxwell Kellogg

Working paper

22/11

Household self-insurance and the value of disability insurance in the United States

Economic and Social Research Council

Household Self-Insurance and the Value of Disability Insurance in the United States

By Maxwell Kellogg*

Draft: February 17, 2022

This paper uses a life cycle model to study the welfare implications of reforms to U.S. Disability Insurance (DI) while accounting for household self-insurance. In addition to crowding out the insurance value of DI, household self-insurance may drive negative selection into DI by reducing implicit application costs. Allowing for such interactions, I find that expansionary DI reforms do not necessarily improve welfare. However, an asset test reduces negative selection and improves the welfare effects of DI expansions. Household self-insurance crowds out the value of DI expansions, but abstracting away from insurance value can deliver erroneous policy recommendations.

* National Bureau of Economic Research, 1050 Massachusetts Ave, Cambridge, MA 02138 (email: mkellogg@nber.org). This research was supported by funding from the National Institute on Aging (T32AG000243). The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. I am grateful to Stéphane Bonhomme, Manasi Deshpande, and Magne Mogstad for their guidance and support. I would like to thank Jack Light for his assistance with the PSID data, and Bradley Setzler for his computational advice. I thank Derek Neal, Alessandra Voena, and participants of the Public/Labor Lunch at the University of Chicago Department of Economics for valuable feedback and suggestions.

The Disability Insurance system (DI) is a large component of the U.S. social safety net. In 2018, it provided over $185 billion in cash transfers to 12.5 million disabled workers, representing approximately 6.2 percent of the working-age U.S. population (Social Security Administration, 2019). Despite substantial growth over the past 50 years and several attempts by policymakers at retrenchment, there is limited research to date on the social value of DI reform. Through the lens of a life cycle model of single earners, Low and Pistaferri (2015) concludes that DI has high social value to individuals compared to its costs. At the same time, other recent work suggests that accounting for spousal labor supply and pooled family savings is important for understanding the insurance value of DI: in Norway, they buffer the effects of DI allowance on income and consumption, at least among individuals who apply for DI and then appeal an initial rejection (Autor, Kostøl, Mogstad, and Setzler, 2019). To the best of my knowledge though, no research to date has incorporated spousal labor supply and family savings in examining the life cycle utilization—and the social value—of the DI system.

The purpose of this paper is to assess the welfare implications of alternative reforms to the U.S. DI system while accounting for its interactions with household self-insurance. These interactions can take at least two potentially important forms. First, by themselves insuring against risks, household self-insurance mechanisms may crowd out the insurance value of DI. Second, households may negatively select into DI on the basis of self-insurance capacity due to the burden imposed by the lengthy application process, over which poorly self-insured households that highly value income support are unable to effectively smooth consumption. Indeed, motivating empirical evidence documented here is consistent with selection into DI on self-insurance capacity: partnered workers, compared to unpartnered workers, are roughly 50 percent more likely to take up DI benefits after a severe work-limiting disability. Furthermore, they smooth their consumption by supplementing those benefits with higher family savings and stable spousal work income. Unpartnered workers, by contrast, more heavily rely on general social safety net transfers for income support, reduce their work less, and face large declines in consumption.

The main analyses of this paper are derived from a life cycle model that incorporates DI and household self-insurance, on top of the key health, productivity, and life cycle dynamics that have been previously documented in the DI literature. Following Low and Pistaferri (2015), the model includes savings and early life cycle choices; non-separable preferences over work, consumption, and health; different degrees of work limitation; fixed costs of work that depend on work limitation status; permanent shocks to productivity; and interactions with the general tax and transfer system.

Building the model further, households pool their savings and coordinate their labor supply decisions, though correlated risks may limit household members' ability to mutually self-insure. The DI system partially insures workers against risk, with a year-long application process during which the household must fund consumption with accumulated savings, spousal work income, and general safety net transfers. Heterogeneity in the timing of disability risk, the presence of a work-capable partner, and earnings potential within the household generate differences in self-insurance capacity across households by which (negative) selection into DI may arise. DI examiners do not observe the health or self-insurance capacity of applicants or beneficiaries. They instead make allowance and reassessment decisions based on noisy signals of individual work capacity, which depend on the individual's true health, age, and marital status.

The model is empirically fit using panel data on joint consumption, savings, work, health, earnings, and household composition from the Panel Study of Income Dynamics (PSID). I first estimate parameters of exogenous processes governing health, potential wages, and marital status in reduced form, accounting for joint selection into work and latent heterogeneity across households. The model then translates data related to consumption, work, health, and DI utilization into three key sets of parameters that are estimated by simulated method of moments: latent allowance rules for the DI system, work frictions including health-dependent pecuniary costs of work, and health-dependent household preferences over leisure and consumption. To validate the fitted model, I show that it reproduces key features of the data including life cycle patterns in savings and earnings, qualitative patterns in joint work decisions, and dynamic household responses to disability onset. Benchmarking the behavior of households in the model against the literature, it generates plausible responses to variation in wages and DI benefit generosity.

The estimated model provides several novel insights into the welfare implications of DI reform. First, I consider the welfare effects of expansionary reforms when implemented alongside revenue-neutral tax rate adjustments. I find that simple expansionary reforms are not necessarily welfare-improving, because the work disincentives of such reforms can be substantial and their benefits accrue to the married households that disproportionately select into DI, who are comparatively well-insured against risks. However, households can benefit substantially from expansions to DI when they are implemented alongside an asset test, which limits fiscal costs and reverses negative selection patterns. That said, an asset test does not serve as an adequate substitute for the lengthy application process as a screening device: providing income support to all current DI applicants (i.e., through a temporary DI program with trivial health certification criteria) would expand the scope of the DI system, incur substantial fiscal costs, and ultimately reduce welfare even after an asset test has been imposed.

Second, I assess how the value of expansionary DI reform is crowded out by household self-insurance and driven by health risks. The broad implicit insurance provided by marriage can crowd out half or more of the value of DI expansions, though spousal labor supply alone plays at most a modest role in that result. Furthermore, alternative reforms differ in their alignment with the DI system's statutory purpose of insuring against disability risk. Conventional adjustments to the leniency of DI allowance or the generosity of DI benefits maintain 26 to 38 percent of their value in the absence of health risks, which can instead be attributed to reallocation across the life cycle and across non-health dimensions of the state space. By contrast, a temporary DI program maintains over 90 percent of its value in the absence of health risks, behaving in this way more like an expansion to the Food Stamp Program than one to the DI system.

Third and finally, I compare the welfare gains of each reform to their fiscal costs and I consider the role that insurance value plays in driving those gains. On a per-dollar basis, households would prefer public spending in an asset-tested DI system be directed toward an increase in the leniency of the DI application process. However, relying on an approach to welfare analysis that abstracts away from the insurance and redistributive value of DI reform would guide policymakers toward two types of erroneous conclusions, in spite of crowd-out by household self-insurance. First, policymakers would incorrectly conclude that households are unwilling to pay for the costs of expansionary DI reforms, with or without an asset test. Second, policymakers would fail to correctly distinguish relative household preferences over the DI reforms, concluding that households prefer spending on an increase in benefit generosity or the roll-out of a temporary DI program.

The findings of this paper contribute first and foremost to the empirical literature on the DI systems. The motivating evidence—the apparent links between household self-insurance and take-up of DI after disability onset—has not been documented in this literature, to my knowledge. Relative to existing structural analyses, this paper stresses the importance of household behaviors, not only because they may crowd out the insurance value of DI, but also because they may drive negative selection into the application process. This negative selection into DI is one reason to consider supplemental (or alternative) screening devices such as asset tests, which have been theoretically motivated in DI by Golosov and Tsyvinski (2006) but have not been evaluated in the context of a realistic life cycle model of individual or household behavior. Furthermore, the result that gains from DI reform are driven both by health and non-health risks corroborates recent evidence pointing to that conclusion from the reduced-form literature (Deshpande and Lockwood, 2020).

By assessing the role of insurance value in driving the welfare gains of DI reform, this paper also connects to a strand of the public finance literature popularized by recent work including Hendren and Sprung-Keyser (2020) and Finkelstein and Hendren (2020). This strand has developed an alternative approach to welfare analysis that is based on causal estimates of the economic effects of a policy change. As noted by Hendren and Sprung-Keyser (2020), this approach is transparent and straightforward to implement, but in practice often relies on ex-post effect estimates that abstract away from insurance value.[1] This limitation may be a minor concern in other settings, but the results of this paper suggest that it has important consequences for guiding DI policy.

The remainder of the paper proceeds as follows. Section I provides background information on the U.S. DI system and summarizes the institutional reasons that negative selection on the basis of household self-insurance may arise. Section II introduces the core data that will be used in the model and presents empirical evidence from event studies of disability onset that is consistent with negative selection into DI. Section III describes the setup of my life cycle model of the household. Section IV describes the mapping of model parameters to data and describes the process of empirically fitting the model. Section V fits the model parameters to data and shows that it reproduces the empirical behaviors of households well. Section VI documents the key insights from the model regarding the welfare value of DI reforms and their relationships with household self-insurance. Finally, Section VII concludes.

I. Disability Insurance and the Household

The U.S. Disability Insurance system (DI) consists of two programs, both administered by the Social Security Administration. Benefits from these programs consist of cash transfers and health insurance coverage. Social Security Disability Insurance (SSDI) is the larger of two, with its eligibility and benefit amounts linked to a worker's earnings history, mimicking Social Security retirement benefits. The second program, Supplemental Security Income (SSI), provides support for non-workers and can supplement SSDI for workers who had low earnings. Unlike SSDI, SSI is explicitly asset-tested but eligibility does not require a work history.[2] SSI provided $750 per month to enrollees in 2018, though these benefits are reduced by other household income, including work income from family members.

[1] Hendren (2021) is a notable exception. In the setting of insurance markets, he proposes a method that estimates ex-ante welfare effects based on willingness to pay and the gap in the marginal utility of consumption for those who do versus do not purchase insurance.

[2] Couples can own no more than $3,000 in assets, excluding one's home and the household's first car. Since 2014, the ABLE Act allows beneficiaries to protect up to $100,000 in savings for a broad set of qualifying disability-related expenses.

Both DI programs impose strict work limitations on applicants and beneficiaries, requiring that they earn no more income than "substantial gainful activity" ($1,180 per month in 2018). Additionally, they share a common application process that is intended to limit take-up to individuals with both poor health and poor employment prospects, consistent with their statutory purposes. However, the application process is potentially complicated and lengthy, and work limitations apply throughout. After submitting an application, applicants receive an initial decision from examiners at a state Disability Determination Service (DDS) office. If denied, applicants may then apply for reconsideration by the DDS office and await a second decision. Applicants who are denied again may appeal their case once more and wait for a review by an administrative law judge (ALJ). If denied at the ALJ stage, applicants have additional opportunities to appeal that are rarely utilized in practice: first to the SSA Appeals Councils and lastly to the federal courts. Among applications filed from 1989 to 1999, roughly 39 percent of applicants were allowed in the initial stage of their first application. This rate ultimately climbs to 67 percent with appeals and reapplications, with most additional allowances coming at the ALJ phase (French and Song, 2014).

Each stage of the application process takes, in expectation, several months or more. Including the appeals process and re-applications, eventual DI beneficiaries wait on average 15.3 months for an allowance decision (Autor, Maestas, Mullen, and Strand, 2017). The application process consequently makes the work restrictions of DI more onerous, as applying for DI requires workers to undergo a prolonged period with little work income and no DI support. However, while both SSDI and SSI impose strict work limitations on applicants and beneficiaries, SSDI does not impose any limitations on household savings or the work income of *other* family members. This leads to a key observation of this paper: that the work restrictions and potentially lengthy application process may perversely screen out workers who value insurance highly but lack the household self-insurance capacity to smooth consumption on their own. To my knowledge, there exists little evidence on how individuals select into DI on the basis of household self-insurance, or any household characteristics for that matter.

II. Data and Motivating Empirics

The analyses of this paper rely primarily on data from the Panel Study of Income Dynamics (PSID), which follows a nationally representative longitudinal sample of households and collects particularly detailed information on the person (or heterosexual couple) most financially responsible for the household (the "reference person"). I use data collected from 1986 (when major reforms to the DI system were last enacted) to 2017, where the PSID runs annually until 1997 and biannually afterward. I impose three restrictions on the sample that reflect the focus of the modeling exercise. First, I restrict to households in which the reference person has no more than a high school education. These reference persons are more than 2.5 times more likely to receive DI than their counterparts with more education. They are plausibly the observable subgroup in which the welfare value of DI is most concentrated, as well as the most relevant subgroup for considering the effects of marginal reforms to DI. Secondly, I further restrict to households with a male reference person who does not engage in self-employment. The only households excluded by the restriction on sex consist of female reference persons with no male partner. While women compose almost half of DI rolls, women also tend to have less consistent labor force attachment over the life cycle, which introduce concerns about eligibility for SSDI.

These first two restrictions are identical to the restrictions imposed by Low and Pistaferri (2015) on the PSID. Finally, I further restrict to households in which the reference person reports working at least once with no work limitation, and never reports collecting DI through SSI alone. This restriction is intended to exclude individuals who lack the work history to qualify for SSDI eligibility, who may have suffered a health shock in childhood, and who face a different life cycle problem.

A full description of the key income, consumption, wealth, and health measures used from the PSID can be found in Appendix B. I draw the definition for the key measure of work limitation status is from the literature. It relies on self-reported subjective assessments of health that capture i) the presence of a work limitation, and ii) the extent to which it limits the type or amount of work one can do. I say an individual has a severe work-limiting disability if they i) "can do nothing" (as opposed to some types of work) or ii) claim their disability limits work "a lot" (as opposed to "somewhat," "just a little," or "not at all"). A disability is "moderate" if it does not meet this criteria, but still limits work "somewhat" or "just a little." Any remaining response options (reporting no disability, or one that limits work "not at all") imply no work limitation. The strengths and weaknesses of this measure are discussed in detail in Appendix B. As shown there, this work limitation measure captures variation across workers in functional limitations and medical diagnoses for conditions such as cancer, diabetes, or heart disease. However, it correlates more weakly with those objective criteria among the excluded individuals who report collecting SSI (and not SSDI) or who never report working without a work limitation.

A. Motivating Empirics

While key institutional features of the DI system outlined in Section I may drive negative selection on the basis of household self-insurance, whether and to what extent that negative selection arises empirically is not well known. Toward providing some evidence, Figure 1 plots the share of reference persons that are collecting DI benefits in the years surrounding a particular type of negative health shock: their first disability onset, where that disability is a severe work limitation. I separate partnered reference persons from unpartnered reference persons in order to illustrate the relationship between DI take-up and household self-insurance around this event. In the four years prior to the onset of that condition, virtually no reference persons collect SSDI or SSI benefits.[3] For unpartnered persons, DI receipt rises to 7 percent in the year of disability onset and quickly plateaus a year later at roughly 20 percent. By contrast, partnered reference persons not only take up DI faster in the year of disability onset (10 percent), their take up steadily increases over time, so that over 40 percent of them report receiving DI benefits four years later.

As it is empirically rare for workers to collect DI benefits before *ever* reporting *some* level of work limitation, the averages plotted in Figure 1 represent plausible estimates for the effects of a severe disability on DI take-up. Estimating effects on other important household outcomes, though, requires comparing outcomes among work-limited households to their counterfactuals in the absence of work limitation. Appendix C implements this comparison with an event study design that uses not-yet disabled peer households as a comparison control group, and presents estimates from the PSID. Appendix C also includes analyses of DI application behaviors using the same design and data on older adults from the Health and Retirement Study (HRS).

While noisy, these estimates reveal striking qualitative differences between the consequences of disability for partnered and unpartnered workers that a model for assessing the welfare value of DI should replicate. Partnered workers have comparatively high DI take-up, high savings, and stable spousal labor supply after severe disability onset, allowing them to smooth their consumption. Unpartnered workers, by contrast, rely on other transfer programs and experience a decline in consumption due to disability, despite the fact that they reduce their work less than do partnered workers. Analysis of data from the HRS implies that higher take-up of DI among partnered workers is *not* due to a higher propensity to apply given a work-limiting condition. Rather, partnered applicants are more likely to receive benefits conditional on applying. Their greater success at acquiring DI may be explained by a willingness to persist through appeals and reapplications; on average, eventual DI beneficiaries who are partnered wait 273 days for their final allowance decision, whereas unpartnered beneficiaries wait only 178 days. This explanation is further supported by household responses to back pain in the HRS, which is a notoriously difficult-to-verify condition:

[3]The PSID does not consistently distinguish SSDI receipt from SSI receipt across survey waves.

individuals who claim DI benefits after suffering back pain are almost exclusively married, and they tend to wait almost a full year for eventual benefit receipt.

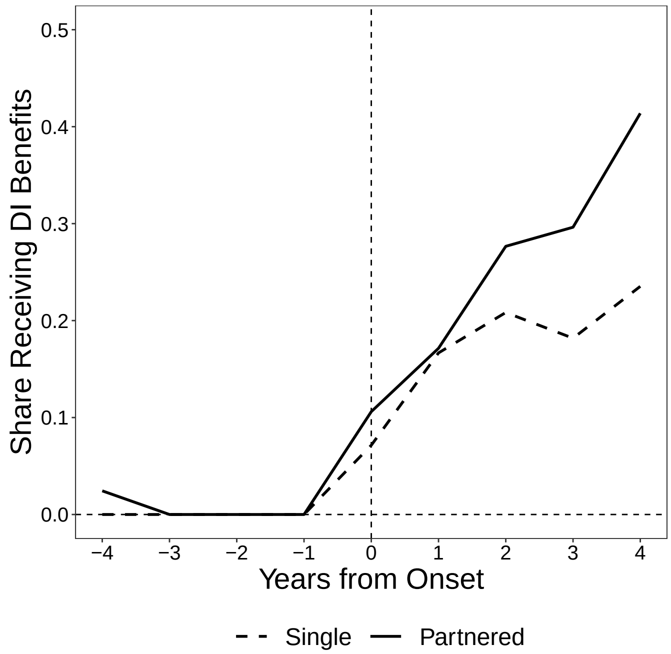


Figure 1. : DI receipt in the years surrounding severe disability onset

III. A life cycle model of the household

While striking, the differences in how partnered versus unpartnered workers respond to disability onset are not sufficient to draw welfare conclusions about the DI system. Those conclusions depend critically on two key features of preferences—the cost of work and the marginal utility of consumption—and how they depend on disability and other household characteristics.

A. The household problem

To both pin down these features of household preferences and infer unobserved details of the DI allowance process, I develop a model of work, health, household formation, savings, and the social safety net. Consistent with the restrictions imposed in Section II, the model concerns the life cycle problem faced by the households of working-age men who have no more than a high school education. Subject to constraints, unitary households choose work, whether to apply for DI (for the reference person), and how much to consume (versus save) to maximize the present discounted value of utility flows:

$$\max_{C_t, L_t^1, L_t^2, App_t^1} E_t \sum_{r=t}^{72} \beta^{r-t} U(C_r, L_r^1, L_r^2; M_r, \boldsymbol{H}_r^1, H_r^2) \tag{1}$$

where β is a discount parameter, L^s is an indicator for full-time employment of partner s ($s \in \{1, 2\}$), M is the marital status of the household, and $\boldsymbol{H}^s$ is a vector of indicators for work limitations for partner s. Work limitations can take one of three levels for the reference person (no disability, $H = 0$, moderate disability, $H = 1$, or severe disability, $H = 2$). Spousal work limitations on the other hand is binary (representing "any" disability, so that $\boldsymbol{H}^2$ reduces to a scalar). Spouses do not exist in unmarried households ($M = 0$), so spousal labor supply is unavailable when reference persons are unmarried.

When making decisions each year, households take expectations over several sources of exogenous risk. Household members face joint disability (and recovery) risks that evolve with age, where health affects potential earnings and preferences over work and consumption. They also face correlated permanent shocks to earnings each working-age year, and a risk of year-long unemployment. Reference persons who who wish to claim DI benefits must apply and face an allowance probability depending on age, disability status, and marital status. DI beneficiaries are randomly selected for reassessment conditional on health status, and are potentially removed from DI. Lastly, households face marriage and divorce risks that evolve with age.

The model allows households to insure dynamically against these risks in three ways. First, households can accrue savings to hedge against the realization of future risks. Second, individuals can adjust labor supply in response to risks. Third, households can receive partial insurance through DI and the general social safety net. The model also approximates systematic spousal behaviors (e.g., home production) that partly insulate households against risks by allowing costs of work and health-dependent labor frictions to differ for married and unmarried reference persons.

A household's state space consists of a fixed "type," their current savings, age in the life cycle, marital status, DI beneficiary status, and individual-specific states for the reference person and their spouse: health status, potential wage, and employment status. The household's type characterizes the profile of marriage and divorce risks, health risks, and potential wages the household will face over the life cycle. It is known to the household but unobserved in the data, and it reflects the already-realized differences across workers as they enter adulthood. It captures the idea that there may be a group of individuals who have low productivity, a low likelihood of marrying a partner, and a high propensity to suffer disability. Through household type, workers may anticipate their future risks and behave accordingly. Selection into DI on the basis of self-insurance capacity may be driven by the systematic cross-type differences and by within-type joint realizations of health, marriage, wage, and employment risks over the life cycle.

Household choices are made subject to several constraints. If exogenously fired, an individual loses the option of working for the year. Reference persons may only apply for DI if they meet eligibility criteria, and the application process takes a full year. While applying for or receiving DI benefits, the reference person faces work restrictions. These work restrictions drive moral hazard in the DI system, in the sense that workers may choose to limit their (otherwise gainful) employment in order to apply for or maintain DI benefits. Households can save but cannot borrow against future income, a restriction which is common in the literature and which prevents households from borrowing against disability or retirement income.

The timing of the model is as follows:

i) Exogenous health, wage, employment shocks, and changes in marital status are realized for year t.

ii) Individuals who have not faced an employment shock may choose to work.

iii) DI beneficiaries face potential reassessment and removal from DI depending on their health status and work decision in year t, where removal leaves them without benefits in year t.

iv) Households then makes DI application choices, and consumption decisions with full knowledge of their current state. Households receives net income from work, along with transfers from DI or other programs.

v) DI allowance decisions are made for applicants based on health information and employment from year t. Those allowed onto DI become beneficiaries, receiving cash payments beginning in year $t + 1$.

The household life cycle begins at age 23, and ends with exogenous retirement at age 62 followed by 10 years of post-retirement life.[4] The life cycle then terminates with no bequest motive.

B. Preferences

Household preferences are non-separable over household consumption and the individual health and work of the two partners. The parametric specification for preferences is an extension of Low

[4]At age 62, households become eligible for early Social Security retirement benefits.

and Pistaferri (2015) and is similar to Autor et al. (2019):

$$(2) \qquad U(C_t, L_t^1, L_t^2; \boldsymbol{H}_t^1, H_t^2, M_t) = \frac{(C_t e(L_t^1, L_t^2; \boldsymbol{H}_t^1, H_t^2, M_t))^{1-\gamma}}{1-\gamma}$$

$$e(L_t^1, L_t^2; \boldsymbol{H}_t^1, H_t^2, M_t) \equiv \exp(\mu M_t + \theta_1' \boldsymbol{H}_t^1 + \theta_2 H_t^2 M_t + \eta_1(M_t, \boldsymbol{H}_t^1) L_t^1 + \eta_2(\boldsymbol{H}_t^1) L_t^2)$$

$$\eta_1(M_t, \boldsymbol{H}_t^1) \equiv \begin{pmatrix} \eta_1^{mod} \\ \eta_1^{sev} \end{pmatrix}' \boldsymbol{H}_t^1 + \eta_1^M \qquad \eta_2(H_t^2) \equiv \eta_2^{mod} H_t^2 + \eta_2^1$$

$$\theta_1 \equiv \begin{pmatrix} \theta_1^{mod} \\ \theta_1^{sev} \end{pmatrix} \qquad \theta_2 \equiv \theta_2^{mod}$$

The fixed cost of work, η_s, is a function of health and marital status. It consists of two additively separable terms: a base disutility of work term η_s^M, which varies by marital status (for the reference person), and an additional disability-related disutility of work term which is common across marital status. The marital intercept μ allows for the possibility that households prefer shifting equivalized consumption toward (or away) from married states of the world. A positive value of μ is consistent with implicit home production performed by spouses (independent of their labor supply) being a Frisch substitute for consumption. This Frisch substitutability with consumption can also manifest in the disutility of work for the reference person (i.e., if $\eta_1^1 > \eta_1^0$).

It is useful to highlight three features of these preferences that arise when work and disability are both "bads" ($\eta_s^M < 0$ and $\theta_s < \mathbf{0}$, for $s \in \{1, 2\}$ and $M \in \{0, 1\}$). First, work and disability will both be Frisch complements with consumption. Due to the Frisch complementarity of disability and consumption, households will prefer to self-insure against disability risk even if disability has no consequences for potential income. Second, work limitation for *one* person increases the disutility of work on the part of *both* household members, even if $\eta_s^{mod} = 0$ (and $\eta_1^{sev} = 0$). Thus, individuals will prefer to avoid working when the household suffers poor health, which is consistent with both disability making work more difficult and with home care arising as a potential response to a partner's disability. Third and lastly, households will prefer to avoid having partners work jointly at the same time. This is consistent with home production (which is not distinguishable from leisure in this model) reducing the cost of work for one's spouse. On the other hand, this is not consistent with leisure complementarities in utility. Instead, positive empirical correlations in work within the household will be driven by life cycle dynamics and by correlated health and wage risks.

C. Savings

Alongside potential adjustments in partner's labor supply, savings are a key mechanism by which households can self-insure against risks. During the working phase of life (when individuals can choose to work and do not receive retirement income), the savings equation takes the form:

$$
\begin{aligned}
(3) \qquad & S_{t+1} = R[S_t + \tau(W_t^1, W_t^2, L_t^1, L_t^2, DI_t, M_t) - C_t^* - F_{t,H_t^1}^{1,M} L_t^1 - F_{t,H_t^2}^{2} L_t^2] \\
& F_{t,H_t^s}^{s,M} \equiv F^{old,s,M} \mathbb{1}\{t \geq 45\} + F^{s,M\prime} \boldsymbol{H}_t^s
\end{aligned}
$$

where S_t are savings at the beginning of year t, R is a fixed interest rate, W_t^s are the hourly wages of partner s, C_t^* is spending on consumption, DI_t an indicator for the reference person receiving DI benefits, and $F_{t,H_t^s}^{s,M}$ is the monetary fixed cost of work that depends on health, current household age, and marital status. The fixed cost of work consists of two additively separable terms: a cost associated with old age, and a cost associated with work limitation. Both costs depend, for the reference person, on marital status, and they reflect how work requires accommodations not captured by PSID consumption categories (e.g., spending on over-the-counter treatments). Spending on consumption is adjusted for family size using the OECD-modified equivalence scale before it enters preferences (the distinction between C_t^* and C_t), reflecting that household consumption is an imperfectly public good. Gross income which includes earnings, DI receipt, and other transfers is mapped into net income by the function τ which approximates the joint tax and transfer system for households.

D. Taxes, and social insurance

The tax and transfer system τ maps earned income and DI beneficiary status into net income, closely following the U.S. tax code and rules for major transfer programs. Those transfer programs consist of the DI system (both SSI and SSDI) and a simple general social safety net consisting of food stamps, Unemployment Insurance (UI), and Social Security retirement. Together, these systems provide partial insurance against the risks of the model. The tax system includes federal income taxes (joint for married partners), child tax credits, FICA taxes, and the EITC. The tax and transfer system is described in full detail in Appendix D, but I continue here with a brief description of the DI system of the model.

The DI system has a year-long application process during which savings, spousal work, and general transfers must support household consumption. This application decision approximates committing to all stages of the process that actual applicants may have to undergo in order to secure a final allowance decision. The only cost of applying for DI is, potentially, forgone work income during the application process and during benefit receipt if allowed onto DI.

Allowance probabilities for applicants depend on marital status, current disability, and age in a manner which is consistent with the Vocational Grid used by examiners:

$$P(DI_t = 1 | App^1_{t-1} = 1, H^1_t, M_t, t) = \begin{cases} \pi^{young}_{H^1_t, M_t} & \text{if } t < 45 \\ \pi^{old}_{H^1_t, M_t} & \text{if } 45 \leq t \leq 62 \end{cases} \tag{4}$$

An applicant who is allowed on DI begins collecting cash benefits in the year after applying. A rejected applicant may return to work. The model includes reapplications; rejected applicants may chose to re-initiate the application process after securing eligibility to re-apply. The model also includes Continuing Disability Reviews (CDRs), whereby DI beneficiaries are reassessed by examiners. Once a CDR is initiated, DI examiners review the beneficiary in the exact same way they review an applicant, with allowance probabilities described in Equation 4. The probability that a beneficiary is subject to a CDR depends on their health status.

By setting up the DI system of the model in this way, I make a series of simplifying assumptions. To avoid the challenging computational issues that surround modeling SSDI eligibility, I assume that DI is available only to the male reference persons (who tend to work regularly), and not their spouses. Furthermore, eligibility for DI requires only that a reference person has worked full-time for at least one year before (re-)applying. For the same reason, the benefit formula for SSDI is a function of potential earnings rather than one's realized earnings history. As I do not allow for an intensive margin to the employment decision, I do not model the substantial gainful activity threshold over which the DI system formally limits work. Instead, I suppose the DI system views all work (regardless of the earnings produced by it) equally. As working while receiving DI benefits is empirically rare, I assume working DI beneficiaries are immediately removed by CDR, so that work and DI receipt are mutually exclusive choices. For the same reason, I assume that DI applications from current workers are rejected.

E. Wage equation

Individuals face potential wages that evolve over time and depend on disability status. Potential wages of partner s at age t are characterized by the equation:

$$\ln W_t^s = \beta_1^{s\prime} \boldsymbol{A}_t^s + \beta_2^{s\prime} \boldsymbol{H}_t^s + f_g^s + \omega_t^s, \quad \omega_t^s = \omega_{t-1}^s + \xi_t^s, \quad \omega_{22}^s = 0 \tag{5}$$

where $\boldsymbol{A}_t^s$ is a square polynomial in age and f is an intercept that is common among all households of the same household type g.[5]

The permanent wage innovations ω_t^s follow a random walk and are correlated within households, reflecting variation in potential earnings which is unrelated to health. Correlations in observed earnings within households will additionally be driven by life cycle dynamics and by correlations in health shocks. With heterogeneity in initial wages being driven by the type-specific intercept f_g, the random walk begins at $\omega_{22}^s = 0$.

F. Employment

Individuals may choose each year whether or not to work full-time. At the end of each year, reference persons who choose to work are subject to exogenous job loss with probability δ_1^M which depends on marital status M. If that occurs, the worker is subject to an unemployment spell for the following year. This job loss probability reflects job destruction and activates Unemployment Insurance benefits (UI). It provides one mechanism by which individuals may choose to apply for DI as a consequence of employment shocks rather than health shocks. Spouses face an unemployment risk δ_2 which contains the same job loss risks faced by their partners, but additionally includes excess unemployment frictions. If these excess frictions prevent a spouse from working (occurring with probability $\delta_2 - \delta_1^1$), the spouse is not rendered eligible for UI. Unlike conventional job destruction, these excess frictions can render spouses unemployed in year $t + 1$ even if not working in year t.

G. Household formation

Marital status evolves over the life cycle as workers exogenously marry and divorce partners. The probability of finding a partner when single, or separating from a partner when married, depends on the age and household type of the reference person. Persons of the same household type match with a common pool of potential spouses, allowing for flexible assortative matching

[5]In estimation, the wage equation also contains calendar year fixed effects.

on risks and characteristics which are realized before early adulthood. The health and wages of a potential spouse are carried with a reference person throughout the life cycle, and they evolve whether the reference person is married or unmarried. This means that divorce does not affect the household's expectations about spousal health or spousal potential wages if the household later re-enters marriage.[6]

The number of adolescent and adult children within the household evolves exogenously and deterministically over the life cycle, depending on household type and marital status. Variation in the number of children over the life cycle, by marital status, and by household type serves two purposes. First, it drives variation in the cost of consumption, since children must be financially supported. All else equal, households will prefer to spend more when they have more children. Second, children enter into the tax and transfer system. EITC payments, tax deductions, tax credits, and personal tax exemptions all depend on the number of children in the household. Similarly, food stamp eligibility and payments both depend on the number of children in the household.

H. Health process

Individual health can take one of three states for reference persons (healthy, moderately work-limited, severely work-limited) or two states for partners (healthy, work-limited). Health within a household evolves jointly over time, according to an exogenous first order Markov process. The transition probabilities depend flexibly on age and household type. This allows health shocks to be correlated within households in a manner that may change over time, and it allows for types to face systematic differences in exposure to disability risk.

One implication of the model is that health risks are exogenous conditional on age and household type, which requires that work, income, or savings do not affect health. This restriction seems empirically reasonable, as effects of non-employment on health found in the literature have been economically small (Stern, 1989; Bound and Waidmann, 2007). Furthermore, shocks to wealth appear to have little effect on health (Kim and Ruhm, 2012; Schwandt, 2018; Adda, Banks, and von Gaudecker, 2009).

IV. Taking the model to data

The model is estimated using the household data from the PSID. Relative to comparable longitudinal and publicly available datasets, the PSID is uniquely well-suited to this purpose. It contains

[6] Savings follow the male partner across marriage and divorce, as I do not model behaviors of unmarried spouses separately.

joint information on key household outcomes related to disability, DI receipt, and well-being across the life cycle: work, income, self-reported health, consumption, program take-up, and assets. Aside from its small sample size, the main limitation of the PSID is that it does not collect information about applications for DI benefits—only receipt of benefits. Unfortunately, no publicly available US datasets, to my knowledge, contain information about household characteristics the timing of health changes, DI application behaviors, and DI allowance over the life cycle. While the HRS contains much of this information, it only samples from the population of workers above age 50. This limits the value of the HRS for this model, as there are well-known differences in how stringently the DI system treats younger versus older applicants (Chen and van der Klaauw, 2008). Like Low and Pistaferri (2015), I circumvent the problem of missing DI application data using indirect inference: I simulate out the decision to apply for DI in the model, and assess how well the model performs at replicating moments related to DI receipt, which can be readily obtained in the actual and simulated data.

Lacking an analytic solution, the model must be solved numerically, working backward from the last year of life. More information on the solution method for the model can be found in Appendix E. Parameters are fit in a three-step approach that is common for life cycle models of this form.

In the first step, some model parameters are set externally. I fix the probabilities of reassessment by CDR for DI beneficiaries to match SSA guidelines. These guidelines are based on likelihood of recovery ("expected to improve", "not permanent", or "permanent"), but I follow Low and Pistaferri (2015) in mapping them to current health status in the model (giving probabilities of 1, 1/3, and 1/7 in order of decreasing health). I also fix some preference parameters at values commonly used in the literature (e.g., Low and Pistaferri 2015; Autor et al. 2019). Specifically, I fix the interest rate R at 1.016, the discount factor β at 0.9756, and the coefficient of relative risk aversion γ at 1.5.

The remaining two steps fit parameters of the model to data. In the second step, reduced-form estimators are constructed for health and marriage transition probabilities, household types, and wage equation parameters. Finally, the third step estimates remaining parameters using the structure of the model. These parameters (parameters governing preferences, work frictions, and DI allowance probabilities) are estimated by simulated method of moments (SMM). While all these parameters are estimated jointly and their mappings into empirical moments are indirect, each parameter can be linked to a specific moment which intuitively provides identifying information. Details on the numerical procedure used to implement SMM can be found in Appendix F, whereas the links between parameters and moments are discussed below in Section IV.D.

A. Ex-ante heterogeneity across households

Households in the model have some ex-ante information about their self-insurance capacity, their exposure to health risks, their household formation prospects, and their potential wages. To estimate this heterogeneity empirically, I assume that households draw one of three potential types at the beginning of the life cycle. These types are then estimated by k-means clustering on reference persons' average log wages when healthy and their share of healthy years spent working full-time.[7] With this approach, estimating the model amounts to the two-step procedure proposed by Bonhomme, Lamadon, and Manresa (2019), except that the second step here is moment-based. In classifying household types, I restrict to households observed in at least three survey waves. Reference persons with no earnings in any survey wave are assigned to the lowest type, as these households all tend to have low marriage rates and high disability risk.

The differences across types arise non-parametrically from the data. The estimates presented in the sub-sections to follow show that households with the least fortunate childhood draw ("low-type" households) face high risks of disability, are unlikely to get married, and have low expected potential wages. High-type households are likely to be married, have low disability risks, and have high expected potential wages. Mid-type households tend to have lower wages like low-type households, but face lower risks of disability and singlehood. Only 15 percent of households in the analytic sample are classified as low-type households and 38 percent are classified as mid-type households, leaving 47 percent classified as high-type households.

The k-means classification procedure imposes some restrictions on the data generating process. For instance, it requires that conditional expected wages and employment ($E[W^1_{it}|H^1_{it} = 0, g(i)], E[L^1_{it}|H^1_{it} = 0, g(i)]$) vary as a function of $g(i)$, so that the individual sample means can be used to differentiate the groups. Likewise, it requires that W^1_{it} and L^1_{it} are ergodic processes, at least conditional on $H^1_{it} = 0$, so that within-person averages converge to individual expectations. Additionally, it requires an assumption related to sampling and non-response in the PSID: namely, that when households appear in the survey (with respect to age or calendar time) is unrelated to their type g.

Because household-level averages are taken over potentially short panels, the average characteristics used to classify households may be noisily measured. Consequently, k-means may partition

[7] Characteristics are residualized with respect to age of the reference person and calendar year, to control for differences arising due to when the household was sampled by the PSID.

household types with excess error, as it constructs these partitions by minimizing within-group variances rather than using information from the model to maximize likelihoods. Table A.1 of the Appendix assess this concern using households simulated from the fitted model. It shows that k-means clustering assigns 74 percent of households to the type most likely to generate its observed average characteristics. Most of the classification error comes from mistakenly assigning low-type and mid-type households to instead be high-type households.

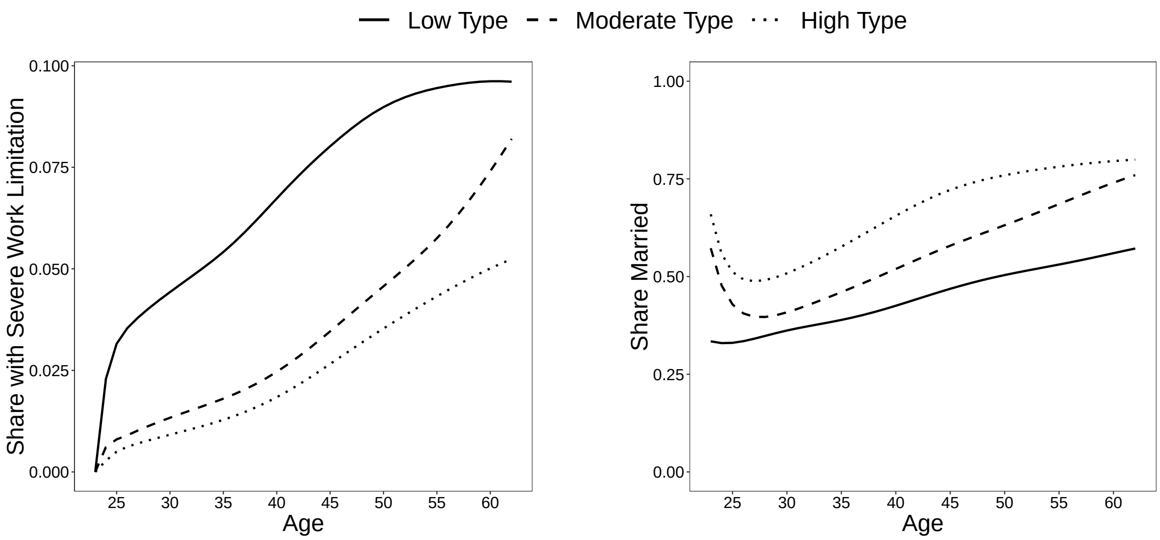
B. Health risks, household formation risks, and household demographics

Taking household type as given, the size of the household, marital transition probabilities, and health transition probabilities are independent of any household choices and can be estimated by their natural sample analogues.[8] After calculating the age-specific transition probabilities in the data, I smooth them using local regression over age for each household type. To illustrate heterogeneity in the health risks faced by households, Figure 2a depicts the resulting estimated probability that a reference person suffers a severe work limitation over the life cycle given their household type. Reference persons in low-type households are more likely than other types to experience a severe work limitation over the entire working-age life cycle. At age 35, for instance, low-type reference persons are more than twice as likely to suffer from a severe work limitation than mid-type or high-type reference persons. Mid-type households are subsequently more likely than high-type households to suffer severe work limitation, with their risks rising particularly as retirement approaches. Figure 2b depicts the probability that a reference person is married at each age, reflecting flows into and out of marriage across the life cycle. Low-type households are roughly half as likely as other types to start the life cycle with a married partner. Mid-type and high-type households experience net divorce in the first 3 years of adulthood, but their net marriage rates outpace low-type households for the remainder of the life cycle.

C. Wage process

To estimate the wage equation, I follow Low and Pistaferri (2015) in augmenting Equation 5 with an idiosyncratic error term which can be interpreted as measurement error. In addition, for

[8]The PSID survey moves from annual to biannual waves in 1996. In estimating transition probabilities for marriage and health, I follow Low and Pistaferri (2015) in abstracting away from the change in panel structure.



(a) Severe Work Limitation

(b) Marriage

Figure 2. : Exposure to marital and health risks for different household types

reference persons, I augment the equation with marriage effects:

$$\ln W_{it}^{s} = \beta_1^{s\prime} \boldsymbol{A}_{it}^{s} + \beta_2^{s\prime} \boldsymbol{H}_{it}^{s} + \beta_3 M_{it} + f_{g(i)}^{s} + \omega_{it}^{s} + \epsilon_{it}^{s} \tag{6}$$

$$\omega_{it}^{s} = \omega_{it-1}^{s} + \xi_{it}^{s} \tag{7}$$

$$\begin{pmatrix} \xi_{it}^{1} \\ \xi_{it}^{2} \end{pmatrix} \sim N \left(0, \begin{bmatrix} \sigma_{\xi}^{1\,2} & \theta \sigma_{\xi}^{1} \sigma_{\xi}^{2} \\ \theta \sigma_{\xi}^{1} \sigma_{\xi}^{2} & \sigma_{\xi}^{2\,2} \end{bmatrix} \right), \quad \epsilon_{it}^{s} \sim N(0, \sigma_{\epsilon}^{s\,2}) \tag{8}$$

Estimation of the wage equation then proceeds in two steps. In the first step, I estimate the coefficients β_1^s and β_2^s using first differences:

$$\Delta \ln W_{it}^{s} = \beta_1^{s\prime} \Delta \boldsymbol{A}_{it}^{s} + \beta_2^{s\prime} \Delta \boldsymbol{H}_{it}^{s} + \beta_3 \Delta M_{it} + \Delta \omega_{it}^{s} + \Delta \epsilon_{t}^{s} \tag{9}$$

These coefficients are estimated by OLS, using a standard selection correction (Heckman, 1979). To implement the correction, I follow Low and Pistaferri (2015) by instrumenting for selection into work with a simulated measure of the generosity of potential welfare benefits, in the event that an individual has low work earnings. This measure is in the spirit of simulated IV, and uses variation over states and time in welfare program policies (see Appendix G). With household types identified separately from the parameters of the wage equation, I allow individuals to differently select into work as a function of their household type in the first stage.

The type-specific fixed effects of the wage equation are estimated in the second step. This step relies on the residual log wage, removing the effects of observable characteristics:

$$\widetilde{\ln W_{it}^s} = \ln W_{it}^s - \hat{\beta}_1^{s\prime} \boldsymbol{A}_{it}^s - \hat{\beta}_2{}' \boldsymbol{H}_{it}^s - \hat{\beta}_3 M_{it} \tag{10}$$

Estimates for the type-specific fixed effects $\hat{f}_g^s$ are taken for each g by averaging $\widetilde{\ln W_{it}^s}$ over all observations for which $\hat{g}(i) = g$, and applying a selection correction to account for the truncation of wage innovations ω_{it}^s due to selection into work.[9]

Appendix H provides more information regarding identification and estimation related to the wage equation, including the selection correction procedure and modeling of the work decision. It also presents and derives an estimator of the covariance matrix of the wage innovations. This estimator accounts for selection into work on the part of both partners, extending how the standard Heckman correction for the wage equation accounts for selection on the part of a single individual. The covariance estimator relies on the assumption that measurement error terms are exogenous, and it imposes some restrictions on the joint selection problem of partners: it requires that shocks to preferences for work are uncorrelated across partners, and that one's wage shocks are unrelated to the latent preferences for work of one's partner.

Panel A of Table 1 reports estimates for regression coefficients. Disability has at most a small effect on the potential wages of reference persons or their partners; a severe disability for the head reduces his work income by about 3.6 percentage points. The coefficients on the age polynomials indicate that the profile of wages over age is slightly steeper for spouses than for reference persons. Consistent with assortative matching, Panel B of Table 1 shows that high-type households tend to have substantially higher reference person and spousal wages (f_3) than mid-type or low-type households (f_2 and f_1).

Panel C of Table 1 reports estimates for the covariance matrix characterizing the wage innovation process for reference persons and their partners. Year-to-year wage innovations are large, with a standard deviation of 14.2 log points for reference persons and 18.1 log points for spouses. Wage innovations are moderately and positively correlated within households, with a correlation coefficient of 0.203.

[9]The selection correction in the estimator $\hat{f}_g$ is consistent with the correction applied in Equation 9 and the distributional assumptions imposed on ω_{it}^s. It is provided in Appendix H.H2. As mentioned in Appendix H.H3, I assume workers do not select on measurement error.

Table 1—: Wage parameters

	Head	Partner
Panel A: Wage Coefficients		
H = 2	-0.0357	—
	(0.0512)	
H = 1	-0.0024	0.0236
	(0.0171)	(0.0441)
Age	0.0560***	0.0645***
	(0.0057)	(0.0100)
$Age^2/100$	-0.0516***	-0.0504***
	(0.0072)	(0.0126)
Married	0.0041	—
	(0.0129)	
Panel B: Type Effects		
f_1	0.895	0.550
f_2	0.889	0.511
f_3	1.516	0.824
Panel C: Variance Parameters		
σ_z	0.1418***	0.1807***
	(0.0417)	(0.0535)
ρ_z	0.2031*	
	(0.1147)	

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the household level.

D. Preferences, frictions, and Disability Insurance parameters

The remaining model parameters (household preferences, work frictions, and DI allowance probabilities when not working) are estimated by indirect inference, using simulated method of moments. The moments related to work and DI take-up which I target are reported in the first column of Table 2 and 3 respectively. The moments related to consumption which I target are reported in the first column of Table 4. While all the parameters are estimated jointly and their mappings into data are indirect, each parameter can be linked to specific moments which intuitively provide identifying information.[10]

It is straightforward to see how these moments partially identify their corresponding parameters, fixing the remaining parameters of the model. The fixed costs of work associated with health (the vector $F^{s,M}$) drive down employment among both younger and older disabled workers, by reducing

[10]To form the objective function minimized by SMM, I use a diagonal weighting matrix which is equivalent to weighting each deviation between an observed and simulated moment by the inverse of the standard deviation of the observed moment. This is the same weighting matrix used by Autor et al. (2019) and Blundell, Costa-Dias, Meghir, and Shaw (2016a), and it is motivated by evidence from Altonji and Segal (1996) that the asymptotically efficient weighting matrix has poor small-sample properties. To numerically solve the objective by SMM, I use a particle swarm algorithm (see Appendix F)

Table 2—: Targeted employment moments and their simulated fit

	Data	Simulation	Diff	St. Error
Panel A: Employment Shares, Married Workers				
$H_t^1 = 0, age_t \leq 45$	0.935	0.935	0.001	0.003
$H_t^1 = 0, age_t > 45$	0.880	0.880	0.000	0.008
$H_t^1 = 1, age_t \leq 45$	0.785	0.797	-0.012	0.023
$H_t^1 = 1, age_t > 45$	0.629	0.607	0.022	0.033
$H_t^1 = 2, age_t \leq 45$	0.381	0.342	0.039	0.045
$H_t^1 = 2, age > 45$	0.186	0.238	-0.052	0.027
Panel B: Employment Shares, Single Workers				
$H_t^1 = 0, age_t \leq 45$	0.844	0.845	-0.001	0.009
$H_t^1 = 0, age_t > 45$	0.788	0.769	0.019	0.023
$H_t^1 = 1, age_t \leq 45$	0.610	0.693	-0.083	0.058
$H_t^1 = 1, age_t > 45$	0.472	0.509	-0.038	0.081
$H_t^1 = 2, age_t \leq 45$	0.336	0.321	0.015	0.049
$H_t^1 = 2, age > 45$	0.219	0.202	0.017	0.065
Panel C: Employment Shares, Spouses				
$H_t^2 = 0, age_t \leq 45$	0.588	0.594	-0.006	0.010
$H_t^2 = 0, age_t > 45$	0.567	0.532	0.036*	0.018
$H_t^2 = 1, age_t \leq 45$	0.343	0.308	0.036	0.024
$H_t^2 = 1, age_t > 45$	0.191	0.270	-0.079*	0.021

Note: * $p < 0.05$ Standard errors are scaled by a factor $(1 + 1/S)$, where $S = 10$ is the number of times the PSID dataset is re-simulated (McFadden, 1989).

the payoff from work when one suffers a disability. The fixed cost of work associated with old age ($F^{old,s,M}$) mediates the difference in employment rates between younger workers and older workers fixing health. The exogenous job loss rate δ_s^M drives down employment among young and healthy workers specifically, who face no fixed costs of work and typically prefer to work if given the option. These parameters are all intuitively pinned down by employment shares for reference persons and their partners, conditional on health status and age bin.

The DI allowance probabilities for individuals who do not engage in work are naturally pinned down by the flow of reference persons into DI receipt, conditional on health status, marital status, and age. They can alternatively be pinned down by the composition of DI beneficiaries by health status, conditional on marital status and age. For estimation, I fit the allowance probabilities to target both of these sets of moments, to improve precision.

The parameters for disutility of work and poor health (the functions η_s and the parameters θ_s) are linked to empirical moments from a regression of log household consumption on work and

Table 3—: Targeted DI moments and their simulated fit

	Data	Simulation	Diff	St. Error
Panel A: Flows into DI, Married Workers				
$H^1_t = 0, age_t \leq 45$	0.000	0.000	0.000	0.000
$H^1_t = 1, age_t \leq 45$	0.008	0.009	-0.001	0.004
$H^1_t = 2, age_t \leq 45$	0.107	0.109	-0.002	0.027
$H^1_t = 0, age_t > 45$	0.004	0.005	-0.001	0.001
$H^1_t = 1, age_t > 45$	0.052	0.043	0.009	0.012
$H^1_t = 2, age_t > 45$	0.249	0.230	0.019	0.034
Panel B: Flows into DI, Single Workers				
$H^1_t = 0, age_t \leq 45$	0.001	0.001	0.000	0.000
$H^1_t = 1, age_t \leq 45$	0.031	0.026	0.005	0.016
$H^1_t = 2, age_t \leq 45$	0.105	0.083	0.022	0.035
$H^1_t = 0, age_t > 45$	0.017	0.015	0.002	0.005
$H^1_t = 1, age_t > 45$	0.048	0.054	-0.006	0.029
$H^1_t = 2, age_t > 45$	0.150	0.161	-0.011	0.049
Panel C: Health Composition of Married DI Beneficiaries				
$H_t = 2, age_t \leq 45$	0.685	0.787	-0.102	0.070
$H_t = 2, age_t > 45$	0.573	0.559	0.014	0.044
$H_t = 1, age_t \leq 45$	0.148	0.213	-0.065	0.050
$H_t = 1, age_t > 45$	0.249	0.283	-0.034	0.036
$H_t = 0, age_t \leq 45$	0.167	0.000	0.167*	0.061
$H_t = 0, age_t > 45$	0.178	0.158	0.020	0.029
Panel D: Health Composition of Single DI Beneficiaries				
$H_t = 2, age_t \leq 45$	0.530	0.452	0.078	0.094
$H_t = 2, age_t > 45$	0.413	0.363	0.050	0.070
$H_t = 1, age_t \leq 45$	0.288	0.373	-0.085	0.093
$H_t = 1, age_t > 45$	0.221	0.244	-0.023	0.045
$H_t = 0, age_t \leq 45$	0.182	0.175	0.007	0.055
$H_t = 0, age_t > 45$	0.365	0.393	-0.028	0.066

Note: * $p < 0.05$ Standard errors are scaled by a factor $(1 + 1/S)$, where $S = 10$ is the number of times the PSID dataset is re-simulated (McFadden, 1989). Values in Panels C and D sum to 1 over levels of H_t, fixing age and marital status.

household characteristics conditional on marital status:[11]

$$\begin{aligned}\ln C_{it} =& \alpha_0 + \alpha_1 M_{it} + \alpha_2^{M_{it}\prime} \boldsymbol{H}^1_{it} + \alpha_3 H^2_{it} M_{it} + \alpha_4^{M_{it}\prime} \boldsymbol{H}^1_{it} DI_{it} + \alpha_5^{M_{it}\prime} L^1_{it} \boldsymbol{H}^1_{it} + \\ &+ \alpha_6 L^2_{it} H^2_{it} + \alpha_7^{M_{it}} t + \alpha_8^{M_{it}} t^2 + \epsilon^{cons}_{it}\end{aligned} \tag{11}$$

The motivation for using these moments comes from the first order condition for optimal consumption each period, taking work decisions and health as given. Differentiating the preferences of Equation 2 with respect to consumption makes clear the log-additive relationship underlying the

[11] When I estimate the consumption regression in sample, I account for part-time work in the same way that full-time work enters the regression equation. I also estimate the consumption regression with calendar year effects.

Table 4—: Targeted log consumption regression coefficients and their simulated fit

	Data	Simulation	Diff	St. Error
$(H_t^1 = 1)X(M_t = 0)$	-0.039	-0.081	0.043	0.102
$(H_t^1 = 2)X(M_t = 0)$	-0.235	-0.341	0.107	0.209
$(L_t^1 = 1)X(M_t = 0)$	0.467	0.547	-0.079	0.080
$(DI_t = 1)X(H_t^1 = 0)X(M_t = 0)$	0.068	-0.164	0.232*	0.103
$(DI_t = 1)X(H_t^1 = 1)X(M_t = 0)$	0.403	0.160	0.244	0.168
$(DI_t = 1)X(H_t^1 = 2)X(M_t = 0)$	0.489	0.261	0.228	0.228
$M_t = 1$	0.062	-0.263	0.325	0.261
$(H_t^1 = 1)X(M_t = 1)$	-0.017	-0.003	-0.013	0.085
$(H_t^1 = 2)X(M_t = 1)$	-0.113	-0.110	-0.003	0.121
$(L_t^1 = 1)X(M_t = 1)$	0.167	0.205	-0.038	0.047
$(L_t^1 = 1)X(H_t^1 = 1)$	0.082	0.198	-0.117	0.090
$(L_t^1 = 1)X(H_t^1 = 2)$	0.156	0.436	-0.281	0.159
$H_t^2 = 1$	-0.104	-0.097	-0.007	0.036
$L_t^2 = 1$	0.142	0.133	0.009	0.021
$(L_t^2 = 1)X(H_t^2 = 1)$	0.153	0.255	-0.102*	0.048
$(DI_t = 1)X(H_t^1 = 0)X(M_t = 1)$	-0.141	-0.219	0.078	0.089
$(DI_t = 1)X(H_t^1 = 1)X(M_t = 1)$	-0.075	-0.037	-0.038	0.119
$(DI_t = 1)X(H_t^1 = 2)X(M_t = 1)$	0.057	0.001	0.056	0.134

Note: * $p < 0.05$ Standard errors are scaled by a factor $(1 + 1/S)$, where $S = 10$ is the number of times the PSID dataset is re-simulated (McFadden, 1989).

consumption regression.

$$(12)\quad \frac{\partial U(C_t, L_t^1, L_t^2; \boldsymbol{H}_t^1, \boldsymbol{H}_t^2)}{\partial C_t} = C_t^{-\gamma}(\exp(\mu M_t + \theta_1' \boldsymbol{H}_t^1 + \theta_2 H_t^2 M_t + \eta_1(M_t, \boldsymbol{H}_t^1) L_t^1 + \eta_2(\boldsymbol{H}_t^2) L_t^2))^{1-\gamma}$$

With full insurance, households would equalize marginal utility of consumption (Equation 12) across potential work and health states within each time period, and set them proportionally across time according to the relative size of the interest rate versus the discount factor, R/β. This implies that the work preference parameters defining the function η_1 (taking γ as known) should map to coefficients in the vectors α_5^M in the log consumption regression (Equation 11), and likewise for η_2 and α_6. The same is true for the disability parameters θ_1 and θ_2 and the marital coefficient μ mapping to coefficients in the vectors α_2^M, α_3, and α_1 respectively. However, this identification result relies on full insurance against work and health risks for households.

In reality households are not perfectly insured against disability (or wage and work risks), so log consumption is contaminated with income effects driven by both health and work. These can threaten the identification of the disutility parameters if the income effects break the injective mapping of model parameters to observed empirical moments. While it is difficult to formally dismiss this threat to identification, it can be instructive to focus one parameter at a time. As shown for select parameters in Figure A.1 of the Appendix, key preference parameters tend to have a monotonic relationship with their linked empirical moment, fixing all other model parameters.

V. Indirect inference results and their implications

Table 5—: Parameter values estimated by simulated method of moments

	Single		Married		Spouse	
	Coef	St. Error	Coef	St. Error	Coef	St. Error
			Panel A: Preferences			
μ	—		—		1.021	(0.328)
θ^{mod}	-0.708	(0.355)	(same as single)		0.095	(0.163)
θ^{sev}	-0.167	(0.3503)	(same as single)		—	
η	-0.822	(0.080)	-0.165	(0.045)	-0.033	(0.163)
η^{mod}	-0.068	(0.145)	(same as single)		-0.487	(0.224)
η^{sev}	-0.407	(0.226)	(same as single)		—	
			Panel B: Frictions and Fixed Costs			
δ	0.016	(0.057)	0.073	(0.008)	0.437	(0.011)
F^1	0.037	(0.136)	0.297	(0.260)	0.001	(0.608)
F^2	0.175	(0.223)	0.308	(0.525)	—	
F^{old}	0.099	(0.039)	0.211	(0.062)	0.411	(0.144)
			Panel C: Allowance Probabilities			
π_0^{young}	0.101	(0.041)	0.222	(1.132)	—	
π_1^{young}	0.174	(0.120)	0.017	(0.241)	—	
π_2^{young}	0.639	(0.201)	0.820	(0.483)	—	
π_0^{old}	0.278	(0.045)	0.369	(0.038)	—	
π_1^{old}	0.231	(0.131)	0.287	(0.067)	—	
π_2^{old}	0.547	(0.181)	0.995	(0.160)	—	

Note: For spouses, δ reflects total risks of job loss (including those that do not lead to UI). The fixed costs of work for moderate and severe disability (F_2 and F_3) and for old age (F^{old}) are reported as a share of median income at age 23. Standard errors are calculated using the formula provided by Gourieroux, Monfort, and Renault (1993).

The parameter estimates that govern household preferences, costs of work, and DI allowance probabilities are reported in Table 5. Panel A reports the estimated disutilities associated with marriage, work, and work-limiting disability. The marriage intercept is large and positive ($\mu = 1.021$), consistent with implicit spousal home producting acting as a Frisch substitute for consumption, so that households prefer consuming relatively more in states of singlehood. The disutility of work is comparatively low for married reference persons ($\eta_1^1 = -0.165$) and very high for unmarried reference persons ($\eta_1^0 = -0.822$). This is consistent with implicit home production of spouses reducing the cost of work, and it means that avoiding work is a potential motive for taking up DI even in the absence of a work-limiting disability, especially for unmarried workers. A severe work-limiting disability has both a direct effect on utility ($\theta_1^{sev} = -0.167$) and on the disutility of work ($\eta_1^{sev} = -0.407$). By contrast, a moderate work limitation has an even larger direct effect on utility ($\theta_1^{mod} = -0.708$) but virtually no effect on the disutility of work ($\eta_1^{mod} = -0.068$), implying it cannot be characterized simply as a "weaker" version of a severe limitation. The point estimate for spousal work-limiting disability, on the other hand, implies a slightly *positive* direct effect on utility ($\theta_2^{mod} = 0.095$), though like θ_1^{mod} and θ_1^{sev}, it is imprecisely estimated and statistically indistinguishable from zero.

These parameter estimates imply that work and disability are Frisch complements with consumption for reference persons. Disability being a Frisch complement with consumption, in turn, provides households with a motive for insuring against disability risk—even if it has no effect on household income. Whereas severe work-limiting disabilities also affect the utility cost of work, this Frisch complementarity with consumption is the main way that moderate work-limiting disabilities directly enter preferences.

Estimates for the exogenous job loss risk and the fixed costs of work are reported in Panel B of Table 5. I find a 7.3 percent and 1.6 percent chance of job loss per year for married and unmarried reference persons respectively, compared to a 43.7 percent chance for spouses. The comparatively high risk of job loss suggests that spouses face substantial informal frictions, on top of the job destruction risk that renders one eligible for unemployment insurance. These frictions could reflect unmodeled spousal home production processes, which serves to reduce the utility cost of work for the reference person and explains why η_1^1 is substantially lower than η_1^0. However, fixed costs of work associated with disability and older age are both larger for married workers than unmarried workers, though this difference weighs against the higher average earnings of married workers (who tend to be higher-type) when affecting work behavior.

Panel C of Table 5 reports the fitted allowance probabilities for DI applicants by age and health status. These allowance probabilities reflect how households undergo the application and appeals process conditional on initiating it. For the most part, married applicants are more likely to receive DI allowance than their unmarried peers, which is consistent with married applicants being more willing to fully commit to the application and appeals process. Furthermore, allowance probabilities tend to increase with age as implied by the Vocational Grid used by examiners (Chen and van der Klaauw, 2008). However, the DI system struggles to properly distinguish healthy and moderately work-limited applicants; while severely work-limited applicants have the highest allowance probabilities, healthy applicants tend to have *higher* allowance probabilities than their otherwise identical moderately work-limited peers, especially among applicants who are married. This suggests that the DI system is vulnerable to "false" claims, and that screening mechanisms may play an important role in discouraging healthy individuals from initiating them.

A. Fit of model to targeted moments

The complete set of moments targeted by the simulation procedure are reported in Table 2, Table 3, and Table 4. Table 2 presents the work-related moments, Table 3 presents DI-related moments, and Table 4 presents the consumption regression moments. Overall, they suggest that the model is able to fit these key features of the data very well.

Panels A and B of Table 2 show that the model closely replicates employment patterns over the life cycle for both married and unmarried workers. Even absent any work-limiting disability, unmarried workers are less likely to work full-time than married workers. Employment declines with age and with work limitations, and the model fits these patterns as well. The only statistically significant deviations from the data arise when fitting the employment patterns of spouses, reported in Panel C. Most substantively, the model does not replicate the age gradient in employment that spouses with a work limitation exhibit. This gradient is steeply downward-sloping in the data, whereas it is flatter in simulation.

Empirical moments describing how individuals enroll in DI as a function of age, health, and marital status are reported in Panels A and B of Table 3. The model is able to closely fit all of these moments. DI entry increases monotonically with age and work limitation status for both married and unmarried workers, and marital differences in DI flows are most prominent among older, severely work-limited individuals. Among those individuals, the married are more likely to enter DI than the unmarried, consistent with negative selection into DI on self-insurance capacity.

Panels C and D describes the composition of DI beneficiaries in terms of their observable characteristics. These moments more broadly reflect entry into and exit from DI, as well as changes in health and marital status for DI beneficiaries. The model consequently fits these moments less well in absolute terms, but they are also less precisely estimated. A substantial faction of DI beneficiaries in truth suffer no work limitation at all, which together with the low DI entry rates among healthy individuals in Panels A and B suggests that the CDR process does not consistently remove DI beneficiaries after they recover their good health. Among these patterns, the one significant deviation from the data occurs for younger married DI beneficiaries; the model implies that they virtually always suffer at least some work limitation, whereas 16.7 percent of them report no work limitation empirically.

Next, Table 4 shows that the model manages to reproduce the relationships between consumption, work, health, and DI receipt described by the log consumption regression. Only 2 of the 18 simulated coefficients statistically differ from their empirical counterparts. For both married and unmarried households, consumption decreases with work-limiting disability of the reference person. However, households tend to increase their consumption if they manage to secure DI receipt (especially if unmarried), suggesting the Frisch complementarities between consumption and work-limiting disability that arise in the fitted model. Consumption increases with work for both married and unmarried households, and increases to an even larger extent if the worker suffers a work-limiting disability. These patterns drive the estimated Frisch complementarities between consumption and work ($\eta_1^0 > 0$, $\eta_1^1 > 0$, and $\eta_2 > 0$) that are stronger in the event of a work limitation ($\eta_1^{mod} < 0$, $\eta_2^{mod} < 0$, and $\eta_1^{sev} < 0$). Compared to unmarried workers, married workers tend to have weaker consumption responses to work and disability, suggesting they more effectively smooth consumption.

B. Untargeted fit and implications of the model

While the model performs well at replicating the targeted features of the data, additional untargeted features can reveal how realistically the model captures key aspects of the household problem that are relevant for studying the DI system. I first examine how well the model replicates household responses to the onset of a severe work-limiting disability, as studied in Appendix C. Though these effects are not used to fit the model, Table 6 shows that the model successfully replicates the qualitative differences by marital status. Panel A reports the effect estimates for households that continue to report a severe work-limiting disability after the initial negative health shock. Like their actual counterparts, simulated married households behave in a manner consistent with comparatively robust insurance against health risk: they increase their DI take-up and reduce their work to a greater extent than unmarried households, while better smoothing their consumption. Panel B reports effect estimates for households in which the severity of the limitation has improved, showing that disability has some persistent effects on DI receipt and work that the model does not capture. Absent a richer model for the CDR process, this persistence is controlled by the DI allowance probabilities for healthy and moderately work-limited households: the DI take-up of recovered households can be increased (and their employment decreased) in simulation by increasing allowance probabilities for healthy and moderately work-limited DI applicants.

I next compare simulated DI stocks and DI continuation probabilities against their empirical counterparts in Table 7. The model qualitatively replicates patterns in these moments, but the health and DI processes of the model lack the dynamics necessary to fit them exactly. Simulated DI stocks of Panels A and B tend to be too high for individuals without a severe work limitation, regardless of marital status. However, simulated continuation rates on DI for those individuals are

Table 6—: Empirical and simulated effects of a severe work limitation

	Single				Married			
			Data - Simulation				Data - Simulation	
	Data	Simulation	Coef.	St. Error	Data	Simulation	Coef.	St. Error
	Panel A: Effects for Current Severely-Limited Households							
DI Receipt	0.18	0.13	-0.06*	(0.06)	0.34	0.24	-0.10*	(0.05)
Work	-0.41	-0.61	-0.20*	(0.08)	-0.46	-0.62	-0.16*	(0.06)
Spouse Work	-0.07	-0.01	0.06**	(0.03)	-0.06	-0.02	0.04*	(0.05)
Log Spending	-0.32	-0.37	-0.05	(0.19)	0.02	-0.06	-0.09	(0.15)
	Panel B: Effects for Recovered Households							
DI Receipt	0.13	0.03	-0.10*	(0.06)	0.14	0.05	-0.09**	(0.05)
Work	-0.26	-0.03	0.23	(0.13)	-0.29	-0.07	0.21*	(0.06)
Spouse Work	-0.11	-0.01	0.10*	(0.05)	-0.08	-0.05	0.03*	(0.08)
Log Spending	0.05	0.02	-0.03	(0.24)	0.06	0.03	-0.04	(0.12)

Note: * $p < 0.05$ Standard errors are scaled by a factor $(1 + 1/S)$, where $S = 10$ is the number of times the PSID dataset is re-simulated (McFadden, 1989).

generally *too low*, as shown in Panels C and D.

Spending, income, and savings profiles.

Turning to broader household behaviors, I consider three key and closely related life cycle profiles: spending, income, and savings. First, Figure 3a shows that the fitted model generates a hump-shaped profile of household spending in the PSID like the one observed empirically. Moreover, the model successfully replicates the spending gap between married and unmarried households, at least later in life. Both in data and in simulations, married households see an increase in spending over their working years, whereas spending in unmarried households plateaus at middle age. However the simulated spending profile is too steep in the mid-20's for all households, and for married households, does not have the slight downward curve observed empirically in the 50's and 60's.

Second, the model is able to replicate the life cycle profiles of gross income by marital status, as shown in Figure 3b. However, the simulated and actual income profiles deviate in two ways worth pointing out. First, the log gap between married and unmarried profiles is slightly larger in simulations than is observed in the PSID. This is unsurprising, as unmarried households in data occasionally have secondary (or tertiary) earners that contribute to household income, whereas all simulated unmarried households are single-earners. Second, income among married households simulated from the model does not decline in the 50's and 60's to the same extent it does in the data. The model approximates declining income over these years for unmarried households with a discontinuous drop at age 45, driven by an increase in their fixed cost of work (which causes a

Table 7—: Untargeted empirical and simulated DI moments

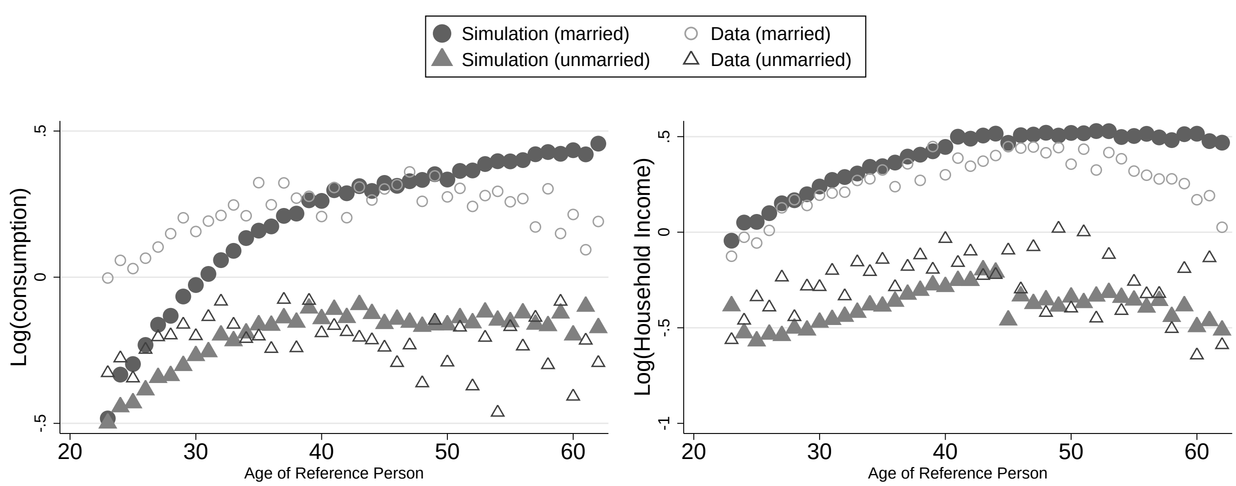
	Data	Simulation	Diff	St. Error
Panel A: Stocks on DI, Married Workers				
$H_t^1 = 0, age_t \leq 45$	0.001	0.000	0.001*	0.000
$H_t^1 = 1, age_t \leq 45$	0.014	0.074	-0.060*	0.006
$H_t^1 = 2, age_t \leq 45$	0.196	0.216	-0.020	0.045
$H_t^1 = 0, age_t > 45$	0.009	0.050	-0.041*	0.002
$H_t^1 = 1, age_t > 45$	0.121	0.197	-0.076*	0.022
$H_t^1 = 2, age_t > 45$	0.462	0.472	-0.010	0.042
Panel B: Stocks on DI, Single Workers				
$H_t^1 = 0, age_t \leq 45$	0.003	0.006	-0.003*	0.001
$H_t^1 = 1, age_t \leq 45$	0.104	0.134	-0.030	0.047
$H_t^1 = 2, age_t \leq 45$	0.287	0.182	0.105	0.060
$H_t^1 = 0, age_t > 45$	0.041	0.080	-0.039*	0.010
$H_t^1 = 1, age_t > 45$	0.258	0.178	0.080	0.066
$H_t^1 = 2, age_t > 45$	0.448	0.333	0.115	0.087
Panel C: Continuing DI Beneficiaries, Married Workers				
$H_t^1 = 0, age_t \leq 45$	0.800	0.000	0.800*	0.210
$H_t^1 = 1, age_t \leq 45$	0.800	0.385	0.415	0.227
$H_t^1 = 2, age_t \leq 45$	0.955	0.977	-0.022	0.046
$H_t^1 = 0, age_t > 45$	0.714	0.077	0.637*	0.101
$H_t^1 = 1, age_t > 45$	0.875	0.577	0.298*	0.056
$H_t^1 = 2, age_t > 45$	0.913	0.957	-0.044	0.042
Panel D: Continuing DI Beneficiaries, Single Workers				
$H_t^1 = 0, age_t \leq 45$	0.889	0.036	0.853*	0.124
$H_t^1 = 1, age_t \leq 45$	0.933	0.730	0.203*	0.078
$H_t^1 = 2, age_t \leq 45$	1.000	0.940	0.060*	0.000
$H_t^1 = 0, age_t > 45$	0.815	0.108	0.707*	0.084
$H_t^1 = 1, age_t > 45$	1.000	0.559	0.441*	0.000
$H_t^1 = 2, age_t > 45$	0.968	0.866	0.102*	0.035

Note: * $p < 0.05$ Standard errors are scaled by a factor $(1 + 1/S)$, where $S = 10$ is the number of times the PSID dataset is re-simulated (McFadden, 1989).

decline in employment).

At least in part, the discrepancies at the two ends of the life cycle profile for both consumption and gross income can be attributed to simplifying assumptions in the model. Fixing income, simulated households may under-consume to compensate for the initial condition of zero savings at age 23; actual households may have already accumulated some savings. Late in life, on the other hand, households lack an early retirement option before age 62. That option is sometimes available to real households through public or private employment, and it may drive down both empirical spending and income in the 50's and early 60's. While these simplifications should affect the entire life cycle profiles, they plausibly distort these tails the most. Broadly though, the model successfully replicates the life cycle trends in consumption and gross income that are observed in the data.

Third, Figure A.2 of the Appendix plots the savings profiles for married and unmarried households. Comparing household savings is particularly useful for gauging how well the model approximates the capacity of married and unmarried households to self-insure against risks at different points in the life cycle. However, two limitations render direct comparisons of the actual and simulated savings profiles difficult. First, the model permits households only a single liquid asset, whereas actual households often accumulate savings in the form of illiquid assets such as homes and pension entitlements. Second, the PSID is known to underestimate wealth compared to arguably more reliable sources, such as the Survey of Consumer Finances (Pfeffer, Schoeni, Kennickell, and Andreski, 2016). Reassuringly, simulated and actual average savings holdings are similar in the final years of the working-age life cycle (when the simulated and actual savings measures are most comparable). While simulated households tend to over-save compared to actual households as expected, it remains true that married households tend to have higher savings than unmarried households. In the model, this is driven by three advantages married households have over unmarried households: they tend to be higher-type (thus having higher earnings potential), they can rely on spousal employment as a secondary source of income, and they tend to have lower costs associated with work (both in terms of its direct disutility and its fixed cost). As a result, higher savings is one important margin that tends to allow married households better self-insurance than unmarried households.



(a) Household Spending

(b) Household Income

Figure 3. : Simulated and empirical life cycle household outcomes

Note: Panel (a) plots non-equivalized, demeaned log household spending. Panel (b) plots non-equivalized, demeaned log average earnings. Both are reported in 1996 US dollars.

Table 8—: Empirical and simulated joint labor supply behaviors in married households

	Data	Simulation	Diff	St. Error
Neither Head nor Spouse Works	0.083	0.117	-0.034*	0.004
Only Spouse Works	0.044	0.052	-0.007*	0.002
Only Head Works	0.463	0.572	-0.109*	0.008
Both Head and Spouse Work	0.410	0.259	0.151*	0.008

Note: * $p < 0.05$ Standard errors are scaled by a factor $(1 + 1/S)$, where $S = 10$ is the number of times the PSID dataset is re-simulated (McFadden, 1989).

Joint work decisions of partners.

While the model is targeted to fit only marginal moments related to the employment of the reference person and spouse, Table 8 shows that it is able to capture the qualitative patterns in *joint* work behaviors within the household well. The most common married work arrangement in married households is for only the reference person to work, though reference persons and spouses often jointly work at the same time as well. Compared to the data, the model implies reference persons work on their own 10.9 percentage points too often, and work jointly with their spouse 15.1 percentage points too infrequently. Households more often do not work at all than do spouses work alone, in both the fitted model and the data.

Elasticities.

One final but important task is to verify that the model generates plausible responses to incentives for work and DI receipt. To that end, the first column of Table 9 reports overall elasticities simulated from the model, which can be compared to elasticity estimates drawn from the literature. These elasticities characterize key response margins at play in DI reforms.

Panel A of Table 9 reports aggregate elasticities of non-employment and DI applications with respect to DI benefit generosity. I find an elasticity of non-employment of 0.09, which lies on the lower end compared to estimates ranging from 0.06 to 0.93 based on the static models of the earlier literature (Bound and Burkhauser, 1999; Haveman and Wolfe, 2000) and which compares closely to the value of 0.056 estimated by Low and Pistaferri (2015). I find an elasticity of DI applications to DI benefits of 0.68 on the other hand, which closely matches the value of 0.62 simulated by Low and Pistaferri (2015) and which fits well within the range of 0.2 to 1.3 appearing in the earlier literature (Bound and Burkhauser, 1999).

Panel B of Table 9 reports estimates for aggregate labor supply elasticities with respect to potential wages. Workers simulated from my fitted model lean toward the inelastic side of the literature.

Table 9—: Simulated and empirical labor supply elasticities

		by Marital Status		by Health		
Elasticity	Overall	Single	Married	Healthy	Mod. Limit	Sev. Limit
Panel A: elasticities with respect to DI benefit generosity						
Non-employment	0.09	0.09	0.09	0.06	0.12	0.19
Applications	0.68	0.64	0.72	0.86	0.79	0.41
Panel B: elasticities with respect to potential wages						
Head, Own Marshal	0.18	0.28	0.12	0.15	0.52	1.42
Spouse, Own Marshal	0.14	—	0.14	0.14	0.17	0.13
Head, Own Frisch	0.61	1.38	0.17	0.51	1.57	4.12
Spouse, Own Frisch	0.20	—	0.20	0.18	0.32	0.26

Note: Elasticities are calculated by finite-difference evaluated with a 10% increase in DI benefits, or one standard deviation of a permanent shock to log wages.

Marshallian elasticities, estimated based on responses to permanent changes in potential wages for the entire life cycle, for simulated men and women (i.e. reference persons and their partners) are 0.18 and 0.14 respectively. In a review of estimates over the past several decades, Keane (2011) finds values that range from -0.47 to 0.51 and from -0.20 to 0.89 respectively. More recently, Blundell, Pistaferri, and Saporta-Eksten (2016b) produces a value for women of 0.42, whereas Blundell et al. (2016a) produces values of 0.22 or 0.55 for women in couples, depending on the presence of children. My simulated Frisch elasticity for the reference person, estimated based on responses to a single-year unanticipated shock, is 0.61. This lies toward the lower end of the range of estimates reviewed by Keane (0.03-2.75), but compares well to more recent work; Blundell et al. (2016b) finds a Frisch elasticity of labor supply of 0.58 for men. I find a Frisch elasticity for women of 0.20, which is somewhat lower than recent estimates of 0.88 and 0.63 found by Blundell et al. (2016b) and Blundell et al. (2016a) respectively, but again well within the range of 0.03 to 3.05 from Keane (2011).

Overall, the first column of Table 9 reveal empirically reasonable responses by simulated households to key incentive margins. The remaining columns detail how, according to the model, these responses differ based on marital status (Columns 2 and 3) or reference person health status (Columns 4, 5, and 6). The small response of non-employment to DI benefits is driven more by the moderately and severely work-limited than the healthy. DI applications are conversely more elastic for healthy workers. This in part reflects the fact that few healthy workers apply for DI at baseline, but it also suggests the inefficiencies that arise when expanding the DI system through simple increases in the leniency of the DI allowance process. Conversely, labor supply responses to potential wages are more driven by unmarried individuals and individuals who suffer a work limitation.

VI. Welfare analysis of reforms to Disability Insurance

I now use the fitted model to consider the welfare implications of alternative policy experiments that reform the DI system. Regarding how to interpret these experiments, note that they represent local deviations from the status quo; I do not attempt to make claims about the design of optimal policy. Furthermore, the policy experiments reveal only partial effects of DI reform. They hold fixed potentially important general equilibrium effects, including any that would affect the exogenous processes of the model such as potential wages or wage risks, health risks, marriage/divorce risks, or DI allowance probabilities.

I estimate households' ex-ante willingness to pay (WTP) for each experiment, paid with a proportional reduction to all household consumption (i.e., a consumption tax). In formulating that WTP, households take expectation over their initial type, initial marital status, and the various life cycle risks. Consequently, households may derive insurance value from reforms that hedge against the realization of both initial conditions and adulthood risks. This commonly-used, consumption-based measure of WTP is particularly attractive because extracting it does not distort the non-consumption decisions of simulated households, allowing me to focus on how those behaviors are affected by the alternative reforms themselves. It is monotonically linked to the proportional increase in expected utility caused by the reform. Letting $\text{EV}_{\text{baseline}}$ and $\text{EV}_{\text{reform}}$ denote expected lifetime utility under the baseline model and under the reform respectively, WTP can be calculated in percentage terms as:

$$\text{WTP} = 100 \times \left(1 - \left(\frac{\text{EV}_{\text{baseline}}}{\text{EV}_{\text{reform}}}\right)^{\frac{1}{1-\gamma}}\right) \tag{13}$$

The policy experiments that I consider approximate historically enacted and recently proposed reforms to the DI system. For the first reform, I proportionately increase the probability of allowance for DI applicants (i.e., I increase DI leniency). This approximates an adjustment to policies guiding how DI examiners assess cases, holding fixed applicant behavior.[12] For the second reform, I proportionately increase the generosity of payments to DI beneficiaries.[13] I focus on the effects of a

[12]The leniency of the DI allowance process has effectively varied over time due to a period of retrenchment (1979 to 1984) and a subsequent period of liberalization (1984 to 1998), in addition to judge retrainings and SSA office closings in the 2010's (Autor and Duggan, 2003; Schwartz, 2018; Deshpande and Li, 2019). The reform I implement is equivalent to reducing the time examiners spend on each case in the following simple model of the application process: DI applicants send exponentially distributed signals of good health, truncated based on their health status. If an examiner sees a signal of positive health while reviewing an applicant's case, the application is rejected. Otherwise, the applicant is allowed DI benefits.

[13]The generosity of DI payments, relative to potential earnings, have been increasing over time for most workers. This is because wage growth has been concentrated at the top of the income distribution, and the DI benefit formula is indexed to *mean* wage growth.

10 percent increase in allowance probabilities for the first reform, and a 10 percent increase in cash benefits for the second reform. The third reform is to implement a temporary DI program, which provides DI applicants with DI-like income support for the year in which they apply. This reform substantially reduces the implicit costs of applying for DI benefits, but in doing so undermines their function as a screening device. The enacted temporary DI program is intended to mimic existing and recently-implemented state-level temporary DI programs.[14] These programs in practice have certification requirements intended to limit take-up to individuals who suffer from a condition, but I provide benefits to *all* applicants in order to err on the side of conservatively over-estimating work disincentives. Lastly, I compare the welfare implications of these reforms to a benchmark of increasing food stamp benefits. Like the reform to DI benefit generosity, I focus on a 10 percent increase in the cash value of benefits provided by food stamps.

A. Welfare implications of revenue-neutral reforms

I first consider the overall effects of these reforms if they are funded through adjustments to the main conventional revenue channel available to the DI system: the Social Security tax on employment earnings.[15] The budgetary costs of each reform primarily come from net transfers to households through safety net programs and changes in income tax receipts due to labor supply adjustments. However, they also include administrative overhead reflecting the size of DI beneficiary rolls. To approximate that overhead, I suppose that the government incurs a cost of $468 for each CDR that it administers, borrowing the value calculated by Low and Pistaferri (2015). As tax adjustments distort work and other household choices, solving for the tax adjustment that offsets the budgetary costs of each reform is an iterative process.

Panel A of Table 10 summarizes the welfare impacts of these revenue-neutral reforms. Column 1 reports the WTP for each reform, indicating for example that a proportional 10 percent increase in DI leniency leads to a *negative* WTP amounting to 0.026 percent of lifetime consumption. A proportional increase in DI benefit generosity is also slightly welfare decreasing, whereas implementing a temporary DI program causes more substantial welfare losses amounting to 2.836 percent of lifetime consumption. By contrast, a proportional increase in food stamp benefits improves welfare, with gains amounting to 0.329 percent of expected lifetime consumption. Note that the magnitude

[14]Temporary DI policies have existed in California, Rhode Island, New Jersey, and New York since the 1940s, and Hawaii since 1969. These programs provide benefits for 13 to 52 weeks, generally replacing roughly two-thirds of the worker's recent wages, conditional on a medical certification. The New York program, though, has an aggressive ceiling on weekly benefits ($170 in 2020).

[15]This is a regressive tax of 6.20 percent on the first $62,700 of wage income. I do not adjust the symmetric wage tax that is also levied on employers, which would affect potential wages (and consequently federal income taxes) depending on its pass-through to employees.

of these welfare effects are partly driven by the reforms' fiscal scopes, which differ substantially. Therefore, it is not clear how households comparatively prefer (or in this case oppose) marginal spending on each of the DI reforms based on these WTP estimates alone. Those preferences are examined in the exercise of Section VI.C.

The remaining columns of Table 10 reveal two observations that can help explain why the expansions to DI decrease welfare. First, while relaxing screening costs benefits worse-insured individuals, it also enables broader take-up which increases fiscal costs: Columns 2 and 3 show that the temporary DI program is funded by a 10.83 percentage point increase in the Social Security tax rate (a 175 percent increase) and causes a 18.2 percentage point decline in employment. Second, selection into benefit receipt can lead to gains that often accrue to married workers, who already have more robust self-insurance capacity: Columns 4 through 9 show that increasing DI leniency or DI benefit generosity both tend to garner comparable if not *higher* DI take-up among married households than unmarried households, fixing household type.

Table 10—: The welfare consequences of revenue-neutral reforms to disability insurance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Change in DI Rolls (p.p.) by Type and Marital Status					
		Change in (p.p.)		Married			Single		
Reform	WTP	SS Tax	Work	Low	Mid	High	Low	Mid	High
Panel A: Effect of Baseline Reforms									
Temp. DI	-2.836	10.83	-18.20	1.19	1.33	0.86	1.32	1.27	1.12
DI Leniency	-0.026	0.06	-0.18	0.32	0.40	0.16	0.49	0.35	0.14
DI Generosity	-0.043	0.08	-0.18	0.22	0.13	0.10	0.18	0.14	0.08
Food Stamp Generosity	0.329	0.38	-0.73	0.04	0.03	0.04	0.14	0.14	0.04
Asset Test Alone	0.210	-0.25	0.48	-2.50	-1.73	-0.94	-0.71	-0.63	-0.38
Panel B: Marginal Effect, With Asset-Tested DI									
Temp. DI	-0.385	7.34	-15.06	0.99	0.82	0.35	1.05	0.95	0.76
DI Leniency	0.004	0.03	-0.10	0.09	0.10	0.02	0.42	0.27	0.06
DI Generosity	-0.002	0.04	-0.09	0.12	0.06	0.04	0.19	0.10	0.05
Food Stamp Generosity	0.336	0.37	-0.73	0.01	0.01	0.01	0.13	0.12	0.03

Note: Willingness to pay is reported as a percent of total consumption. Column 2 reports the percent increase (or decrease) in the Social Security tax on earnings necessary for the reform to maintain revenue neutrality. Panel B reports the marginal effects of enacting each alternative reform after first imposing the asset test on DI. For each row in Panel B, the total effect of the reform and asset test can be calculated by adding the effects of the asset test alone (from Panel A). DI rolls include long-term beneficiaries, excluding those receiving benefits from temporary DI (if applicable).

The positive effects of expanding food stamps suggest that expansionary reforms to DI may benefit households if they are paired with a screening device that controls their fiscal costs and discourages selection into DI by strongly self-insured households. To explore this possibility, I next consider the welfare effects of imposing an asset test alongside each reform, which limits DI eligibility to households possessing no more than $20,000 in savings.[16] On its own, the asset test disproportionately discourages DI receipt by married households, permits a modest tax cut, and ultimately improves welfare (Panel A of Table 10). Furthermore, the asset test improves the marginal welfare implications of the expansionary DI reforms. This is clear from Panel B of Table 10, which summarizes the marginal impacts from each alternative reform after the asset test has been imposed, mirroring their baseline effects in Panel A. Focusing first on the increase in DI leniency and benefit generosity, the asset test reduces take-up responses by married households to both of these reforms, while leaving take-up responses by unmarried households comparatively unaffected (Columns 4 through 9). The decline in take-up response among married households is particularly sharp for the DI leniency reform, allowing that reform to provide marginal welfare gains worth 0.004 percent of total consumption, on top of the gains already generated by imposing the asset test. By contrast, the reduction in take-up response among married households is less sharp for the DI benefit generosity reform, which still causes welfare losses amounting to 0.002 percent of total consumption.

Turning to the temporary DI program, its marginal effects on households suggest that the asset test does not act as an adequate substitute for the lengthy application process as a screening device. Using the temporary DI program to eliminate the implicit burden imposed by that process still causes substantial work disincentives and fiscal costs after the asset test has been imposed. The welfare losses imposed by the temporary DI program (0.385 percent of total consumption) are smaller than the losses imposed in the absence of the asset test, but the program nevertheless remains welfare-decreasing.

B. The household determinants of willingness to pay

I next assess how the benefits of each reform are crowded out by self-insurance mechanisms or driven by health risks by shutting down these aspects of the model. These exercises serve two separate purposes. Crowd-out reveals how household self-insurance behaviors substitute for DI receipt or supplement it for inframarginal beneficiaries, in either case undermining the insurance value of DI.[17] On the other hand, how a reform's gains are derived from health risks reveals its alignment with the statutory purpose of the DI system. For the purpose of these exercises, I isolate

[16]For healthy, moderately work-limited, and severely work-limited households, the share of DI beneficiaries in the baseline model with savings above the asset test threshold are 64 percent, 68 percent, and 69 percent respectively.

[17]In the event that marginal DI recipients only utilize DI when they can supplement benefit receipt with self-insurance behaviors, they can alternatively cause *crowd-in*.

the benefits of each reform by measuring willingness to pay while holding tax rates fixed (i.e., without maintaining revenue neutrality) and taking as given the presence (or absence) of the asset test. The results are presented in Table 11, where each column reports the percent increase (or decrease) in WTP that occurs when a particular risk or insurance mechanism is removed from the model.

Marriage overall crowds out the value of DI expansions substantially, so that failing to account for it would cause one to overestimate the gains from increasing DI leniency or benefit generosity by 89 percent and 52 percent respectively, and the gains from the temporary DI program by 28 percent (Panel A, Column 1).[18] The ways in which marriage broadly enables workers to accumulate higher savings is key to these effects, as marriage crowds out the benefits from all three DI reforms to an even greater extent after an asset test has been imposed (Panel B, Column 1). However, spousal labor supply only drives at most a modest share of these effects. On its own, omitting spousal labor supply slightly *increases* gains from DI expansions without the asset test (Panel A, Column 2), and it reduces them by at most 19 percent with the asset test (Panel B, Column 2). Other aspects of marriage that are unrelated to spousal labor supply but which implicitly insure couples against risks evidently play an important role in reducing the gains from expanding DI.[19] By contrast, the welfare gains from increasing food stamp benefit generosity are not meaningfully crowded out by either marriage broadly or spousal labor supply specifically.

To judge whether and to what extent the gains from each DI reform arise from health risks, I assess how they change when all of the costs of disability to households are shut down, so that disability status only affects DI allowance probabilities (Column 3). If those gains are exclusively derived from the consequences of disability rather than life cycle reallocation or non-health risks, then they should vanish entirely. With or without an asset test, 26 to 38 percent of the gains from increasing DI leniency or DI benefit generosity persist, implying that they depend heavily but not entirely on health risks. By contrast, the temporary DI program provides value more deeply tied to realizations of non-health states, maintaining more than 90 percent of its original welfare gains. In this way, the temporary DI program behaves more like the benchmark reform to food stamps than it does an expansion to the DI system.

[18]Note that this does not contradict the previous result that initially married households have *higher* WTP than initially unmarried households for expansions to DI leniency or benefit generosity. Ex-ante WTP of unmarried (and married) households is depressed by the expectation that they may be married at the time that negative shocks arise in the future.

[19]Married reference persons enjoy more favorable treatment by the tax system and the general social safety net, a lower utility cost of work which allows them to more easily accumulate savings, and higher DI allowance probabilities which reduce the implicit costs of acquiring DI. Marriage also has costs which can adversely affect self-insurance: a larger household size implies a higher price of equivalized consumption, and married reference persons have slightly higher health-dependent fixed costs of work. However, the self-insurance benefits of marriage are large compared to these costs.

Table 11—: Sensitivity of reform benefits to risks and self-insurance mechanisms

	(1) No Marriage	(2) No Spousal Work	(3) No Health Costs
Panel A: Effect of Baseline Reforms			
Temp. DI	28%	-3%	-7%
DI Leniency	89%	-9%	-62%
DI Generosity	52%	-2%	-69%
Food Stamp Generosity	2%	33%	-8%
Panel B: Marginal Effect, With Asset-Tested DI			
Temp. DI	47%	2%	-7%
DI Leniency	119%	4%	-71%
DI Generosity	92%	19%	-74%
Food Stamp Generosity	2%	33%	-8%

Note: All columns report the percent increase (or decrease) in WTP that occurs when a particular risk or insurance mechanism is removed from the model. Marginal WTPs are compared in Panels B and C, after implementing the given screening measures.

C. *The insurance value of DI reforms*

Relative household preferences over spending on the alternative reforms—and the role that insurance plays in driving them—can be seen by comparing ex-ante and ex-post estimates of the reforms' marginal value of public funds (MVPFs). The MVPF is a standardized measure of the welfare impacts of a policy, dating to Mayshar (1990) and Slemrod and Yitzhaki (1996). It is calculated by dividing the WTP for a policy change by its fiscal costs, and whether (or how) insurance value is incorporated into it critically depends on how WTP is measured. To compare against ex-ante estimates of WTP underlying Tables 10 and 11, I use the model to define a natural ex-post WTP measure:

$$
\begin{aligned}
& WTP^{\text{ex-post}} \equiv \sum_i \sum_t \frac{WTP_{it}^{\text{ex-post}}}{R^t} \\
(14) \qquad & WTP_{it}^{\text{ex-post}} \text{s.t. } U(C_{it}^{\text{reform}} - WTP_{it}^{\text{ex-post}}, L_{it}^{1,\text{reform}}, L_{it}^{2,\text{reform}}; \boldsymbol{H}_t^1, \boldsymbol{H}_t^2) = U_{it}^{\text{baseline}} \\
& U_{it}^{\text{baseline}} \equiv U(C_{it}^{\text{baseline}}, L_{it}^{1,\text{baseline}}, L_{it}^{2,\text{baseline}}; \boldsymbol{H}_t^1, \boldsymbol{H}_t^2)
\end{aligned}
$$

where reform and baseline household decisions (e.g., C_{it}^{reform} and C_{it}^{baseline}) are simulated out of life cycles under the alternative policy regimes. The ex-post WTP aggregates the series of household-by-year consumption adjustments which leave all post-reform utility flows at their pre-reform values. Among "inframarginal" households for whom labor supply decisions do not respond to the reform, the ex-post WTP is exactly the present discounted dollar value of the transfers that they receive from the reform (i.e., the reform's "mechanical cost").[20] In this way, the ex-post WTP is closely related to the reduced-form WTP estimators commonly used to construct MVPFs in the public finance literature, including for the purpose of evaluating DI reform (Hendren, 2016; Hendren and Sprung-Keyser, 2020; Finkelstein and Hendren, 2020).[21]

Table 12 reports the ex-ante and ex-post welfare estimates for each reform while holding tax rates fixed and (in Panel B) taking as given the asset test, and it reveals two important results. First, ex-ante MVPF estimates are substantially larger than ex-post MVPFs, indicating that much of the ex-ante value of these reforms comes from reallocating across potential states of the world. Increasing food stamp benefits has an ex-ante MVPF slightly above 1.8. Ex-ante MVPFs for all expansionary DI reforms are below 1 without an asset test (Panel A) and above 1 after implementing the asset test (Panel B). However, ex-ante MVPFs for the temporary DI program and the increase in benefit generosity are both very close to 1, implying that the result from Section VI.A that these reforms *decrease* welfare arises due to the deadweight loss associated with funding them through a wage tax. By contrast, all reforms (including increasing food stamp benefits) have ex-post MVPFs that are well below 1 and which only tend to decrease further with an asset test. As a result, a policymaker guided by ex-post MVPFs would erroneously conclude that households are unwilling to pay the costs for any of these reforms.

Second, Table 12 reveals that insurance value plays a role in the preferential ranking of alternative reforms. After imposing the asset test, increasing DI leniency provides the highest ex-ante MVPF among expansions to the DI system, but the lowest ex-post MVPF. Policymakers guided by ex-post MVPFs would instead conclude that households prefer implementing a temporary DI program or an increase in benefit generosity. The intuition behind the divergent policy recommendations is straightforward: fixing work, ex-ante welfare analysis can favor a policy change that reallocates consumption into states of the world where its marginal utility is high, even if that reallocation causes deadweight loss. On the other hand, ex-post welfare analysis will favor whichever reform increases aggregate consumption (by providing transfers to households) with the least per-dollar deadweight loss, without regard for when consumption occurs.

[20]Neither the structural model nor the reduced-form MVPF estimates reported for DI by Hendren and Sprung-Keyser (2020) allow for spillovers, which would also enter ex-post WTP.

[21]Note, however, that this ex-post WTP measure differs from the reduced-form WTP constructed for DI reforms by Hendren and Sprung-Keyser (2020). This measure also accounts for the welfare effects on the "marginal" beneficiaries who respond behaviorally to reform, who are assumed by Hendren and Sprung-Keyser (2020) to be indifferent. However including "marginal" beneficaries does not necessarily increase aggregate ex-post WTP; their ex-post WTP can be *negative* even for a reform that strictly expands the choice set, if the behavioral response involves trading losses in states where the marginal utility of consumption is low for gains in states where it is high. Ex-post MVPFs reported here also differ from the MVPFs of Hendren and Sprung-Keyser (2020) through the denominator. Fiscal costs accounted for in the model include spillovers onto general safety net programs and broader work disincentives (e.g., due to changes in precautionary savings motives), which are not accounted for in estimates reported by Hendren and Sprung-Keyser (2020).

D. Robustness

Using an individual-level model, previous work by Low and Pistaferri (2015) has found that increasing DI leniency or benefit generosity delivers substantial welfare gains, even in the absence of an asset test. The analyses of the fitted model from this paper imply a different conclusion, but that difference importantly does not stem from the key economic forces that are the focus of this paper. Instead, it can be explained by differences in the scope of the models: as mentioned in Section II, I exclude reference persons who collect DI through SSI alone or who never work without a work limitation. Under this restriction, individuals in the fitted model are far less likely to suffer severe disability than are individuals in the model of Low and Pistaferri (2015). For instance, reference persons are 85 percent more likely to be suffering a severe work limitation in any year if this restriction is lifted. Imposing this additional restriction takes a more conservative approach to studying expansionary DI reforms, to the extent that the excluded households represent an unmodeled state of the world against which DI provides some insurance, even in the presence of an asset test.

To illustrate the impact of this restriction, I consider how some key welfare implications of the model change when it is relaxed. On the less restricted PSID sample, I re-classify household types and substitute the resulting (higher) health risks into the fitted model. Table A.2 of the Appendix summarizes the fiscal costs, welfare implications, and MVPFs of each reform in this exercise, without maintaining revenue neutrality, exactly as Table 12 does for the fitted model. Table A.2 agrees with the findings of Low and Pistaferri (2015); households benefit from simple expansionary reforms to DI even without implementing an asset test, as MVPFs for an increase in the leniency or generosity of DI are both larger than 1. Note that in this alternative model specification, other important insights from the fitted model hold. Households benefit to a greater extent on a per-dollar basis when expansions to DI are implemented with the asset test, and ex-post WTP remains comparatively low (and differently ranked), implying that households still derive substantial redistributive value from DI reform.

VII. Conclusion

This paper develops a life cycle model of work and DI receipt in households that accounts for self-insurance, and uses it to study the welfare implications of DI reform. Individuals in the model begin the adult phase of the life cycle with no work limitations, but face different early life draws that affect their potential wages and risk profiles in adulthood. While work-limiting disability

Table 12—: The ex-ante and ex-post marginal value of public funds for alternative DI reforms

	(1)	(2)		(3)	
	Total	Welfare Measures			
	Fiscal	Ex-Ante		Ex-Post	
	Costs ($)	WTP ($)	MVPF	WTP ($)	MVPF
Panel A: Effect of Baseline Reforms					
Temp. DI	2,022.49	1,742.97	0.86	598.85	0.30
DI Leniency	14.09	10.49	0.74	2.13	0.15
DI Generosity	20.37	13.70	0.67	5.81	0.29
Food Stamp Generosity	93.77	171.98	1.83	53.75	0.57
Panel B: Marginal Effect, With Asset-Tested DI					
Temp. DI	1,406.98	1,452.95	1.03	295.67	0.21
DI Leniency	6.34	8.04	1.27	0.01	0.00
DI Generosity	9.19	9.77	1.06	1.80	0.20
Food Stamp Generosity	93.01	171.69	1.85	54.01	0.58

Note: Willingness to pay and fiscal costs are reported in 1996 dollars per capita per year. Marginal WTP and marginal costs are reported in in Panels B and C, after implementing the given screening measures.

does not have much of an effect on potential earnings in the fitted model, it substantially increase the disutility of work and the marginal utility of consumption. Households therefore have strong motives to support consumption while avoiding work in poor states of health. Consistent with reality, the DI system has a lengthy application process and faces negative selection, with take-up being higher among comparatively well-insured married workers who also benefit from spousal labor supply and pooled family savings.

Using the model to simulate out alternate reforms to the DI system reveals that expansionary DI reforms are not necessarily welfare-improving on their own, and the lengthy application process plays an important role in limiting the work disincentives and fiscal costs of the DI system. An asset test does not serve as an adequate substitute when it comes to screening DI beneficiaries, from either the perspective of controlling fiscal costs or improving household welfare. However, an asset test limits selection into benefit receipt by households with more robust self-insurance capacity, who tend to value benefit receipt less. As a result, it can substantially improve the welfare implications of other expansionary DI reforms. For instance, households benefit when DI is asset-tested and the leniency of the DI allowance process is increased.

Studying these reforms more closely provides further insight into how the welfare gains from DI arise. The value of DI reform is partially crowded out by the self-insurance broadly provided by marriage, especially when eligibility for DI is asset-tested. In that setting, failing to account for spousal labor supply alone would cause one to overestimate welfare benefits from DI expansions by at most 19 percent, but failing to account for other aspects of marriage can cause one to overestimate them by a factor of two or more. Furthermore, reforms differ in their alignment with the statutory purpose of DI: increasing the leniency or benefit generosity of DI yields value that is mostly connected to costs of poor health. If viewed as an extension of DI though, a very lenient temporary DI program broadly expands the system's scope, providing welfare gains that persist almost entirely in the absence of health costs.

By comparing the ex-ante and ex-post welfare implications of DI reforms, it is clear that accounting for the insurance and redistributive value that DI expansions provide—regardless of the risks from which it is derived—is important for welfare analysis. If one were to use welfare measures that do not account for the value of reallocation across potential states, one would erroneously conclude that households are unwilling to pay for any expansionary DI reforms. Moreover, one would misjudge households' relative preferences over marginal spending on alternative DI reforms. Whereas households in an asset-tested DI system benefit most from marginal spending on a more lenient allowance process, a policymaker guided by ex-post MVPFs would find that they instead prefer spending on DI benefit generosity or a temporary DI program.

When considering the interpretation of these results, it is important to keep in mind three caveats. First, the model abstracts away from Medicare, Medicaid, private health insurance, and their relationships with work and DI. Private health insurance in the United States is closely tied to employment and marriage, and health insurance provision through DI is another incentive for take-up of the program. Health insurance is also another way in which the application process of DI may be disproportionately costly for unmarried workers; Unmarried workers can neither rely on their own employer-provided health insurance during the application process, nor on health insurance coverage provided by a spouse. In that sense, the model should underestimate the value of DI benefits, but also underestimate the opportunity cost of applying for unmarried workers.

Second, I do not fully model the application and appeals process of DI. Individuals can, in reality, choose to initiate an application and then choose whether to continue the process based on the signals they receive about their allowance probability as review and appeals proceed. The static, year-long application decision in my model likely overestimates application costs for some workers, though it roughly matches the average wait time for benefits.

Lastly, this study focuses on the value of DI for male workers with low educational attainment who begin adulthood in good health. While less educated working males are a subgroup of the U.S. labor force that is especially prone to DI receipt, and perhaps the group in which the welfare effects of DI may be most concentrated, the restriction does limit the scope of my welfare analysis.

On the one hand, I may understate the disincentive costs of expansionary DI reforms by excluding higher-educated workers who are plausibly better self-insured against risks. On the other hand, I may also understate willingness to pay for these reforms, as they may insure against childhood risks that drive education decisions or cause poor health in early adulthood. How the incentive-insurance trade-off underlying the DI system tends to plays out for female workers is also an understudied topic in the literature that warrants further attention, but is one which this paper does not address.

REFERENCES

Adda, J., J. Banks, and H.-M. von Gaudecker (2009): "The Impact of Income Shocks on Health: Evidence from Cohort Data," *Journal of the European Economic Association*, 7, 1361–1399. 16

Altonji, J. G. and L. M. Segal (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics*, 14, 353–366, publisher: [American Statistical Association, Taylor & Francis, Ltd.]. 22, 70

Andreski, P., G. Li, M. Z. Samancioglu, and R. Schoeni (2014): "Estimates of Annual Consumption Expenditures and Its Major Components in the PSID in Comparison to the CE," *American Economic Review*, 104, 132–135. 52

Autor, D., A. Kostøl, M. Mogstad, and B. Setzler (2019): "Disability Benefits, Consumption Insurance, and Household Labor Supply," *American Economic Review*, 109, 2613–2654. 2, 12, 17, 22, 70

Autor, D. H. and M. G. Duggan (2003): "The Rise in the Disability Rolls and the Decline in Unemployment," *The Quarterly Journal of Economics*, 118, 157–206. 35

Autor, D. H., N. Maestas, K. J. Mullen, and A. Strand (2017): "Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants," Working Paper 20840, National Bureau of Economic Research. 6

Benítez-Silva, H., M. Buchinsky, H. M. Chan, S. Cheidvasser, and J. Rust (2004): "How Large Is the Bias in Self-Reported Disability?" *Journal of Applied Econometrics*, 19, 649–670. 53

Blundell, R., M. Costa-Dias, C. Meghir, and J. Shaw (2016a): "Female Labor Supply, Human Capital, and Welfare Reform," *Econometrica*, 84, 1705–1753. 22, 34, 70

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2016b): "Consumption Inequality and Family Labor Supply," *American Economic Review*, 106, 387–435. 34

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "Discretizing Unobserved Heterogeneity," Becker Friedman Institute for Economics Working Paper No. 2019-16, University of Chicago. 18

Bound, J. and R. V. Burkhauser (1999): "Economic analysis of transfer programs targeted on people with disabilities," in *Handbook of Labor Economics*, Elsevier, vol. 3, 3417–3528. 33, 53

Bound, J. and T. Waidmann (2007): "Estimating the Health Effects of Retirement," Tech. Rep. 2007-168, University of Michigan. 16

Callaway, B. and P. H. C. Sant'Anna (2020): "Difference-in-Differences with multiple time periods," *Journal of Econometrics*. 54

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. 17, 27, 65

Ciupke, K. (2016): *psoptim: Particle Swarm Optimization*, r package version 1.0. 70

de Chaisemartin, C. and X. D'Haultfœuille (2020): "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110, 2964–2996. 54

Deshpande, M. and Y. Li (2019): "Who Is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Economic Policy*, 11, 213–248. 35

Deshpande, M. and L. Lockwood (2020): "Beyond Health: Non-Health Risk and the Value of Disability Insurance," *Working Paper*. 5, 53

Feenberg, D. and E. Coutts (1993): "An Introduction to the TAXSIM Model," *Journal of Policy Analysis and Management*, 12, 189–194, publisher: [Wiley, Association for Public Policy Analysis and Management]. 52

Finkelstein, A. and N. Hendren (2020): "Welfare Analysis Meets Causal Inference," *Journal of Economic Perspectives*, 34, 146–167. 5, 41

French, E. and J. Song (2014): "The Effect of Disability Insurance Receipt on Labor Supply," *American Economic Journal: Economic Policy*, 6, 291–337. 6, 65

Garner, T., G. Janini, L. Paszkiewicz, and M. Vendemia (2006): "The CE and the PCE: a comparison," *Monthly Labor Review*, 129, 20. 52

Golosov, M. and A. Tsyvinski (2006): "Designing Optimal Disability Insurance: A Case for Asset Testing," *Journal of Political Economy*, 114, 257–279. 4

Gourieroux, C., A. Monfort, and E. Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8, S85–S118. 26, 71

Haveman, R. and B. Wolfe (2000): "Chapter 18 The economics of disability and disability policy," in *Handbook of Health Economics*, Elsevier, vol. 1, 995–1051. 33

Heckman, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–161, publisher: [Wiley, Econometric Society]. 20

Hendren, N. (2016): "The Policy Elasticity," *Tax Policy and the Economy*, 30, 51–89, publisher: The University of Chicago Press. 41

——— (2021): "Measuring Ex Ante Welfare in Insurance Markets," *The Review of Economic Studies*, 88, 1193–1223. 5

Hendren, N. and B. Sprung-Keyser (2020): "A Unified Welfare Analysis of Government Policies," *Quarterly Journal of Economics*, 135, 1209–1318. 5, 41

Hoynes, H. W. and D. W. Schanzenbach (2009): "Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program," *American Economic Journal: Applied Economics*, 1, 109–139. 67

Keane, M. P. (2011): "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 49, 961–1075. 34

Kim, B. and C. J. Ruhm (2012): "Inheritances, health and death," *Health Economics*, 21, 127–144. 16

Kreider, B. (1999): "Latent Work Disability and Reporting Bias," *The Journal of Human Resources*, 34, 734–769. 53

Low, H. and L. Pistaferri (2015): "Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off," *American Economic Review*, 105, 2986–3029. 2, 7, 11, 17, 19, 20, 33, 36, 42, 52, 53, 65, 71, 72, 75

Mayshar, J. (1990): "On measures of excess burden and their application," *Journal of Public Economics*, 43, 263–289. 40

McFadden, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57, 995–1026. 23, 24, 25, 30, 31, 33

Meyer, B. D. and W. K. C. Mok (2019): "Disability, earnings, income and consumption," *Journal of Public Economics*, 171, 51–69. 53

Pfeffer, F. T., R. F. Schoeni, A. Kennickell, and P. Andreski (2016): "Measuring wealth and wealth inequality: Comparing two U.S. surveys," *Journal of Economic and Social Measurement*, 41, 103–120. 32, 52

Schwandt, H. (2018): "Wealth Shocks and Health Outcomes: Evidence from Stock Market Fluctuations," *American Economic Journal: Applied Economics*, 10, 349–377. 16

Schwartz, N. D. (2018): "Disability Applications Plunge as the Economy Strengthens (Published 2018)," *The New York Times*. 35

Slemrod, J. and S. Yitzhaki (1996): "The Costs of Taxation and the Marginal Efficiency Cost of Funds," *Staff Papers (International Monetary Fund)*, 43, 172–198, publisher: Palgrave Macmillan Journals. 40

Social Security Administration (2019): "Annual Statistical Report on the Social Security Disability Insurance Program, 2018," Tech. rep., Social Security Administration. 2, 53

Stern, S. (1989): "Measuring the Effect of Disability on Labor Force Participation," *The Journal of Human Resources*, 24, 361–395. 16

Table A.1—: Validating k-means: share of households, by type classification

Model	Household Type K-Means: Low	Mid	High
Low	0.08	0.02	0.06
Mid	0.00	0.21	0.18
High	0.00	0.01	0.45

Note: Columns divide sample households according to the type assigned to them by k-means clustering. Rows divide sample households according to the type that is most likely to generate the average characteristics used in k-means classification (log wages and employment rate when healthy), based on the fitted model. Households for which the two classifications agree lie on the main diagonal (74 percent total).

Table A.2—: Ex-ante and ex-post marginal value of public funds, with higher health risks

	(1)	(2)		(3)	
	Total	Welfare Measures			
	Fiscal	Ex-Ante		Ex-Post	
	Costs ($)	WTP ($)	MVPF	WTP ($)	MVPF
Panel A: Effect of Baseline Reforms					
Temp. DI	1,966.70	1,735.94	0.88	584.28	0.30
DI Leniency	20.78	29.29	1.41	3.17	0.15
DI Generosity	39.48	40.26	1.02	13.81	0.35
Food Stamp Generosity	93.25	178.25	1.91	54.71	0.59
Panel B: Marginal Effect, With Asset-Tested DI					
Temp. DI	1,388.59	1,461.09	1.05	295.46	0.21
DI Leniency	13.62	25.58	1.88	1.26	0.09
DI Generosity	26.62	33.94	1.27	7.73	0.29
Food Stamp Generosity	93.74	178.47	1.90	55.18	0.59

Note: This table reports welfare estimates after substituting higher risks of work-limiting disability into the fitted model, as described in Section VI.D. See Table 12 for further notes on the construction of WTP and fiscal costs.

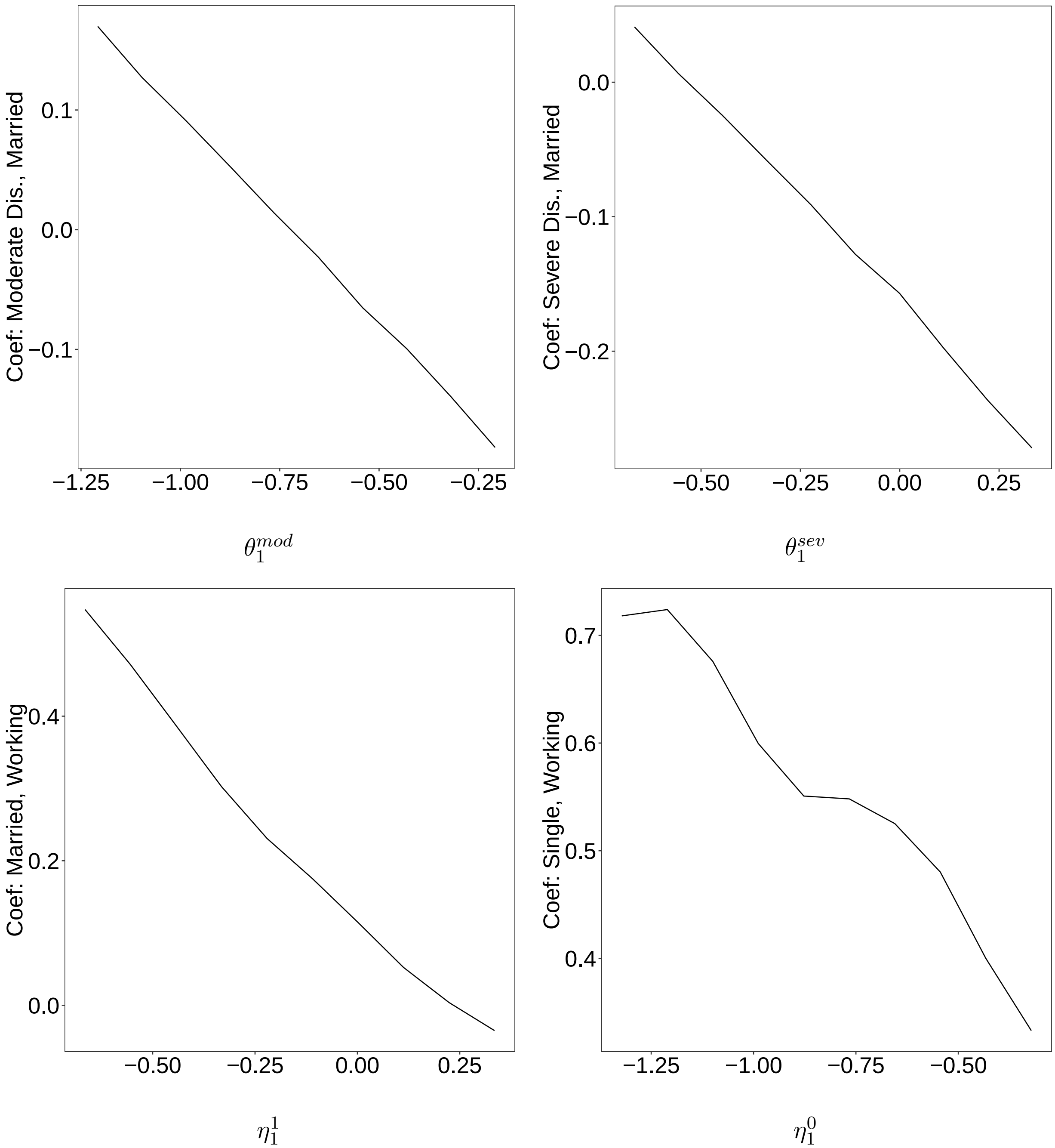


Figure A.1. : Local identification of select parameters by log consumption regression

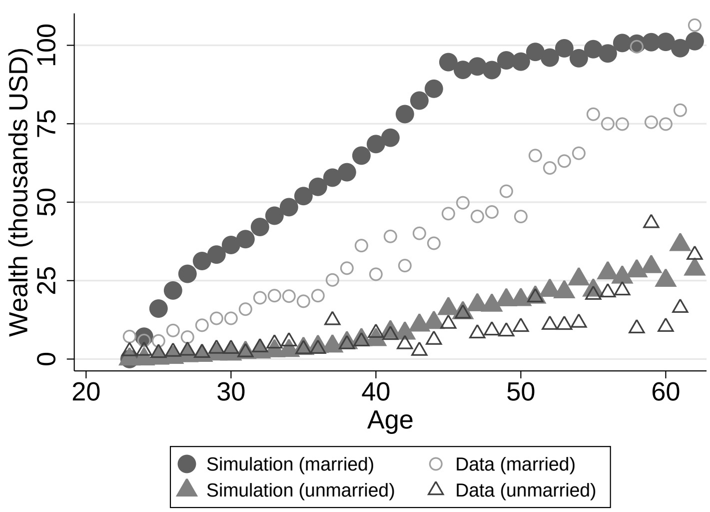


Figure A.2. : Simulated and empirical median accumulation of wealth

Note: This figure plots median wealth in 1996 US dollars.

Definitions of Key PSID Measures

B1. Income

In each wave of the PSID, income measures are collected separately on reference persons, their spouses, and other household members in the calendar year prior to interview. Those individuals report their taxable income (work and business returns) and their transfer receipts from a broad range of public and private sources, including from SSI, AFDC/TANF, UI, workers' compensation, public and private pensions, help from friends and relatives, child support and alimony, and any other sources of public assistance. All of these measures represent pre-tax income; to impute post-tax income for households in the data, I use NBER TAXSIM Version 27 (Feenberg and Coutts, 1993).

B2. Consumption

Limited consumption information is available for all waves of the PSID. I follow Low and Pistaferri (2015) in using more comprehensive measures of household consumption which are only first collected in the 1999 survey wave. This includes information on food, health care, housing and rent, utilities, transportation, education, and child care spending over the calendar year prior to interview.[22] Of these categories, only food spending and rent are available in years prior to 1999. While additional categories are introduced in later years, I restrict to the previous measures for consistency over time.[23] Average total consumption based on these measures compares closely to the more detailed consumption information in the Consumer Expenditure Survey, and underestimates aggregated Personal Consumption Expenditures data of the National Income and Product Accounts by about 35% (Andreski, Li, Samancioglu, and Schoeni, 2014; Garner, Janini, Paszkiewicz, and Vendemia, 2006).[24]

B3. Wealth

Wealth information is also collected at the time of survey from 1999 onward in nine categories: business assets, transaction accounts (including savings), home equity, equity in vehicles, stocks, other real estate equity, retirement accounts, other assets, and other debt. Compared to the more detailed information collected by the Survey of Consumer Finances (SCF), reported wealth tends to be slightly lower in the PSID (Pfeffer et al., 2016). Differences between the SCF and PSID are concentrated in the right tail of wealth, and among households with business assets. Given that the households of interest for this paper tend to have lower wealth and are not self-employed, PSID measures should serve reasonably well.

B4. Disability

The PSID contains self-reported measures of work-limiting disability for reference persons and their partners. Heads and their partners are first asked (Q1): *Do you have any physical or nervous condition that limits the type of work or the amount of work you can do?* If responding with "Yes", the individual is asked (Q2): *Does this condition keep you from doing some types of work*, to which respondents may reply "Yes", "No," or "Can do nothing." To those who reply to the previous question with "Yes" or "No", the final question is posed (Q3): *For work you can do, how much*

[22] Implied rent for homeowners is imputed as 6% of the home's reported value.
[23] The new measures introduced in 2005 include home repairs, household furnishing, clothing, vacations, and recreation.
[24] The CES underestimates the NIPA PCE by 25% in more recent preliminary tables published by the BLS.

does it limit the amount of work you can do? Individuals may respond with " A lot," "Somewhat," "Just a little," or "Not at all."

I code these responses into three exhaustive and mutually exclusive disability categories following Low and Pistaferri (2015), Meyer and Mok (2019), and Deshpande and Lockwood (2020). I classify those who report "No" to Q1 or "Not at all" to Q3 as having no work limitation ($H_{it} = 0$). I classify those who respond "Yes" to Q1 and either "Somewhat" or "Just a little" to Q3 as moderately limited ($H_{it} = 1$). I classify those who report "Yes" to Q1 and either "Can do nothing" in Q2 or "A lot" to Q3 as severely limited ($H_{it} = 2$).

These self-reported measures are commonly used in the literature, and they are arguably preferable over objective diagnostic criteria for studying worker utilization of DI. Many work-limiting disabilities that result in DI receipt stem from conditions that may be more reliably captured with subjective criteria rather than objective criteria: almost 63% of all disabled workers on DI rolls in 2018 were awarded benefits for a mental or musculoskeletal disorder (e.g., back pain) (Social Security Administration, 2019). Nevertheless, subjective measures of health suffer from two potential weaknesses.

First, variation in subjective self-assessments of work limitations may not reflect variation in actual health. However, similar subjective measures have been found to be strongly correlated with objective health criteria (see the review of Bound and Burkhauser, 1999, Section 2.1). Furthermore, both Low and Pistaferri (2015) and Meyer and Mok (2019) show that the strong correlation holds for this particular subjective measure in the PSID. To briefly summarize that result, Table B.1 reports correlations between current subjective work limitations and limitations related to activities of daily living (ADLs) or serious health events. In my analytic sample drawn from the PSID (Panel A), severely work-limited individuals are far more likely to report both ADL-related limitations and serious health events than are moderately work-limited individuals, who in turn are more likely to report them than individuals with no work limitation at all. Importantly, "healthy" individuals (based on work limitation status) rarely report ADL-related limitations or serious health events.

While the exclusion of SSI beneficiaries (who do not have the work history to qualify for SSDI) and individuals never observed as healthy workers is motivated by modeling decisions, note that work limitation is a less precise indicator of objective health for the excluded individuals (Panel B). Compared to the analytic sample of Panel A, the excluded reference persons (reported in Panel B) often suffer from poor objective health even when they report no work limitations. Among "healthy" individuals, 17 percent of the excluded non-sample individuals report ADL-related limitations (versus 4 percent of the analytic sample) and 11 percent report suffering a serious health event (versus 7 percent of the analytic sample).

The second concern with subjective self-assessments is that they may be partly driven by non-health factors, or may be rationalized ex-post to justify (non-)work and DI receipt. On the one hand, Benítez-Silva, Buchinsky, Chan, Cheidvasser, and Rust (2004) show that subjective measures of health are unbiased predictors of disability determinations made by the SSA among those who apply for DI. On the other hand though, Kreider (1999) find that disability is over-reported among the unemployed, compared to the employed who have similar propensity to work based on observables. These concerns are partly addressed, again, by the restrictions that exclude the households of Panel B of Table B.1 from the analytic sample. One consequence of this restriction is that all individuals in my analytic sample have, by revealed preferences, demonstrated some level of willingness to work.

Table B.1—: Health characteristics by self-reported subjective work limitation status

	(1)	(2)	(3)
	Work Limitation Severity		
	None	Moderate	Severe
Panel A: Analytic Sample			
Has some daily limitation	0.04	0.40	0.69
Serious health event	0.07	0.28	0.44
Panel B: Persons Excluded due to SSI or Non-Work			
Has some daily limitation	0.17	0.66	0.84
Serious health event	0.11	0.34	0.38

Note: The first panel reports average health measures conditional on current work limitation status for reference persons in the analytic sample. The second panel does the same for households that are excluded from the analytic sample on the basis of ever receiving SSI (and not SSDI) or labor force non-participation. Within each panel, the first row shows the share of the sample that reports some limitation related to activities of daily living: walking or climbing stairs (1986, 2017), bending or lifting (1986, 2017), driving (1986), seeing (1986), preparing meals (2003-2017), shopping (2003-2017), managing money (2003-2017), using a phone (2003-2017), doing heavy housework (2003-2017), or doing light housework (2003-2017). The second row shows the share that suffer form a major clinical health event (all recorded only in survey waves 1999-2017): cancer, a heart attack, heart disease, lung disease, or a stroke.

Event studies of disability onset

To estimate the causal effects of disability onset on work, income, DI receipt, and DI applications, I rely on an event study design that compares trends over time among work-limited households to trends over time among comparable control households. This design amounts to a form of the inverse-probability-weighted estimator due to Callaway and Sant'Anna (2020) and is similar in spirit to the alternative estimator of de Chaisemartin and D'Haultfœuille (2020), and the current section describes it in detail.

C1. Event Study Design

To provide a formal description of the event study design, let D denote the first year in which a household reports suffering a work limitation ("disability onset"), and let $Y_t(\ell)$ denote the potential outcome Y of a household in calendar year t if it were to face disability onset ℓ years in the past. The parameters of interest are the average treatment effects on the treated (ATTs) associated with work-limiting disability, in each year following disability onset:

$$E[Y_t(\ell) - Y_t(-\infty)|D = t - \ell], \quad \ell \in \{-6, \ldots, 6\} \tag{C1}$$

For each of these year-ℓ ATTs, an estimator is implemented by a simple difference-in-difference regression on a subsample of data from years ℓ and ℓ_0, containing work-limited households and their control groups:

$$Y_{id} = \alpha_i + \gamma_\ell \mathbb{1}\{d = \ell\} + \delta_\ell \mathbb{1}\{d = \ell, i \text{ in work-limited group}\} + \epsilon_{id} \tag{C2}$$

where the time index d is centered around the year of disability onset ($d = 0$) and where δ_ℓ, the difference-in-difference coefficient, is the estimator for the year-ℓ ATT. The year ℓ_0 is the reference period from which within-household differences over time are taken, which I set as the last year before disability onset.[25]

Estimating Equation C2 first requires constructing the control groups. For each year-ℓ subsample, the control group consists of contemporaneous peer households that have not yet faced (or never face) their own work-limiting condition. These peer households match exactly to one (or more) work-limited household on reference person characteristics: birth cohort, age and marital status in the reference year, race, and access to a state-level temporary DI program. For the control households, d denotes time relative to disability onset for the households to which they were matched.[26]

Recovering ATTs from Equation C2 also requires reweighing the data, to allow for treatment effect heterogeneity as a function of observables. Households that suffer a work-limiting disability at a young age will tend to have more not-yet-disabled peer households to serve as control units than households that suffer a work-limiting disability at a later age. To address this and other imbalances across work limitation status, I reweigh the control group so that the distribution of its observable characteristics are identical to the distribution of observables for work-limited households. With weights applied, estimating δ_ℓ of Equation C2 is numerically equivalent to separately estimating Equation C2 for each distinct group of matched work-limited and control households, and integrating over the group-specific estimates of δ_ℓ according to the density of work-limited households.

A key assumption is necessary for the difference-in-difference estimators characterized by Equation C2 to recover ATTs: there must be *common trends* over time between work-limited and otherwise comparable non-work-limited peer households in the absence of work limitations. This will hold if, for lower-educated workers, work-limiting disabilities and their timings are unanticipated and randomly assigned conditional on other observable characteristics. Anticipation of future work-limiting disabilities must be ruled out, as it would contaminate treatment effect estimates through the behavior of households in the control group. As the basis for testing the common trends assumption, I estimate δ_ℓ for $\ell < 0$. Absent anticipation of disability onset, δ_ℓ should be close to zero for all $\ell < 0$.

C2. Data and sample restrictions

In implementing the design, I draw on data from the PSID (described in Section II) and from the Health and Retirement Study (HRS). The HRS follows a longitudinal sample of older US residents (age 50 and above) and their households, collecting detailed information on health and non-work income. Importantly, the HRS has panel data on DI *applications* that is unavailable in the PSID. To better accommodate the nature of the HRS data, I impose additional restrictions on in-sample individuals of the HRS, on top of the restrictions imposed on the PSID in Section II. Namely, I drop individuals who report applying for or receiving DI before the health event captured in the survey data. I use all available waves of the RAND HRS file (biannually from 1992 to 2018). In order to ensure that the sample restrictions do not condition on future labor force participation, I restrict to reference persons in both surveys who report working at least once before the health event.

I separate partnered households from unpartnered households based on survey reports just prior to disability onset, with a definition that differs slightly across the two data sources. In the PSID, partnered households contain a reference person that is married or has been cohabiting with a

[25] For households suffering disability onset on or before 1996 (when the survey is taken annually), $\ell_0 = -1$. For households suffering disability onset after 1996 (when the survey is taken biannually), $\ell_0 = -2$.

[26] A non-work-limited household may match to multiple households with the same observable characteristics that face disability in different years. In that case, the non-disabled household will appear as a control unit multiple times, but the reference year for each duplicate will be pinned to a different age of disability onset.

partner for at least one year. I do not impose any duration requirement on cohabiting partners in the HRS, which has shorter panels of older reference persons. As most of these households contain a married couple, I refer to partnered reference persons interchangeably as "married" reference persons, contrasting them with the remaining "unmarried" or unpartnered reference persons. Table C.1 reports average values for key measures of interest in my analytic sample of PSID households, by marital status of the reference person. Compared to unpartnered reference persons, partnered reference persons are older, more white, wealthier, more likely to have a high school degree than have no degree, and less likely to suffer a work-limiting disability. Partnered households tend to contain more adults and more children than unmarried households. Similar statistics are reported for the analytic sample of the HRS in Table C.2.

Table C.1—: PSID analytic sample summary statistics

	(1) Single	(2) Partnered	(3) Partnered - Single
Panel A: Cross-Sectional Statistics			
Age	36.77***	39.74***	2.96
	(0.36)	(0.19)	(0.40)
High School Grad	0.72***	0.71***	0.00
	(0.02)	(0.01)	(0.02)
White	0.45***	0.61***	0.17
	(0.02)	(0.01)	(0.02)
Number of Adults	1.23***	2.29***	1.06
	(0.02)	(0.01)	(0.02)
Number of Kids	0.27***	1.31***	1.04
	(0.02)	(0.02)	(0.03)
Moderate Dis.	0.05***	0.06***	0.00
	(0.01)	(0.00)	(0.01)
Severe Dis.	0.04***	0.03***	-0.02
	(0.01)	(0.00)	(0.01)
Household Net Worth	23.52***	61.96***	38.44
	(2.55)	(2.72)	(3.61)
Years Worked	18.49***	21.86***	3.37
	(0.36)	(0.19)	(0.40)
Years Worked Full-Time	16.58***	20.57***	3.99
	(0.37)	(0.20)	(0.41)
Panel B: Panel Statistics			
Households per Wave	183	864	681
Waves per Household	4.50	7.69	3.18

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the household level. Wealth is reported in thousands of 1996 US dollars, and excludes businesses assets. Partners include married spouses and non-marital partners.

C3. Event Study Results

I now use the event study design to examine the effects of a work-limiting disability on household work, income, DI utilization, consumption, and savings. Table C.3 summarizes the effects of a *severe* work limitation on work, income, consumption, and savings, estimated using the PSID. Table C.4 summarizes effect estimates from the HRS, focusing on effects on DI applications, appeals, and allowance. As the HRS does not collect the same information about severity, Table C.4 reports effects of *any* work limitation. Columns 1 and 2 of both tables report, for scale, average outcomes in the six years leading up to disability onset for unpartnered and partnered workers respectively. Columns 3 and 4 report the average effects of disability for unpartnered and partnered workers. These coefficients are constructed by averaging the effects of disability on households (the coefficients $\hat{\delta}_\ell$) over the first seven years after disability onset.[27] Column 5 reports the average difference in effects between married and unmarried workers. While differences in Column 5 are generally not statistically significant, these results nevertheless reveal response patterns consistent with negative selection into DI on household self-insurance that are useful for thinking about a model of DI and the household.

Effects on DI receipt and work.

Panel A of Table C.3 summarizes the effects on DI receipt, the same outcome plotted in Figure 1, and work hours for reference persons and partners. After disability onset, married reference persons are roughly 50 percent more likely than unmarried reference persons to take up DI benefits (23 percent versus 16 percent), and they correspondingly reduce their post-disability work to a greater extent than the unmarried (reducing by 790 annual hours versus 614 annual hours). While unmarried households do not tend to have working partners prior to disability, they experience a decline in *future* partner work similar of similar magnitude to the decline experienced by married households.

Effects on income.

Panel B of Table C.3 summarizes the effects of a severe work-limiting disability on measures of equivalized income.[28] Married reference persons tend to experience a smaller decline in annual equivalized work income compared to unmarried reference persons ($4,820 versus $5,300).

The impact of lost earned income of the reference person is attenuated in three ways. First, reference persons receive income from SSDI, which the PSID does not separate form other social security income. Consistent with their higher DI take-up, married households receive an additional $1,440 in annual equivalized dollars, while unmarried households receive $1,160. By contrast, married households receive less in equivalized terms from broader safety net programs such as unemployment insurance, worker's compensation, food stamps, and housing support ($880 versus $1,200).

Second, households can rely on income generated by other household members. Adjustments in the work of other household members due to disability do not tend to drive large changes in income, and are not consistent with added worker effects as a self-insurance mechanism against disability. However, partners in married households worked on average 1,345 hours per year, or roughly 25 hours per week. Consequently, with a pre-disability average of $8,230 in equivalized dollars, work income from the spouse and other household members is an important continued source of income for married households.

Third and finally, lost work income can be attenuated by changes in household tax obligations. Taking the effects on gross income together, married and unmarried households both lose about $2,900 in equivalized income per year due to disability onset, on average. However, factoring in simulated income taxes and credits paints a more positive picture for married households. Married households face a $2,350 decline in equivalized post-tax income thanks in part to spousal work

[27] These coefficients are average treatment effects on the treated, averaging over person-years. More households are observed in years closer to disability onset, so the year of disability onset receives greater weight.

[28] Equivalized measures are divided by the OECD-modified scale, which approximates how household income, spending, and wealth can map into individual-level consumption. It assigns a value of 1 for the reference person, 0.5 for each additional adult, and 0.3 for each child.

income, which receives favorable tax treatment in the absence of labor supply from the reference person. By contrast, unmarried households tend to have little other sources of work income and consequently face a *larger* decline of $3,480.

Effects on spending and wealth.

Panel C of Table C.3 summarizes effects of a severe disability onset on household log spending and log wealth. While these results are particularly noisy because data on spending and wealth are available only after 1999, the point estimates suggest that unmarried households experience a large proportional decline in equivalized spending due to disability onset (18 log points). By contrast, they suggest that married households tend to completely insulate their spending against the decline in income that they experience. Similarly, unmarried households experience a large proportional decline in equivalized wealth (80 log points). Thanks in part to their higher pre-disability savings, the consumption smoothing that married households achieve does not require them to proportionally spend down their wealth to any appreciable extent.

Effects on survey attrition.

Panel D of Table C.3 summarizes effects of a severe disability onset on attrition from the PSID panel that can be attributable to three potential causes: leaving the sample household, death, or other unexplained non-response ("vanishing" from the sample). Neither partnered nor unpartnered reference persons die or leave their sample households due to disability onset. However, unpartnered reference persons are 10 percentage points more likely to vanish from the sample for other unexplained reasons as a result of disability onset. Why those individuals vanished from the sample has implications for how well-insured unpartnered households are against disability risk: if they vanish because disability drives them into homelessness, then unpartnered households are arguably even more poorly-insured against disability risk than the effects on survey respondents would imply. Alternatively, if unpartnered households vanish because they merge with other related households that can help support them (in a manner which the PSID fails to track), then this would imply that the effects on survey respondents under-estimate their average self-insurance against disability risk.

DI applications, appeals, and allowance after disability onset.

The results from the PSID suggest that married households benefit from comparatively high DI take-up, high savings, and stable spousal labor supply after severe disability onset, so that they are able to effectively smooth consumption. Unmarried households, by contrast, rely on other transfer programs, experience a larger decline in post-tax income, and experience a decline in consumption due to disability.

One important question is whether the differences in DI take-up are driven by selection into the application process. However, the PSID lacks information about DI appeals or applications. Thus, turning to the older-age participants of the HRS, Table C.4 summarizes effects of a work-limiting disability on applications, appeals, or allowance decisions. As the HRS does not record the same measures of subjective severity that appear in the PSID, the events in Table C.4 represent work-limiting disabilities of *any* severity. It is first worth noting that even after restricting to adults over the age of 50, it remains true that married workers are slightly more likely to receive DI benefits after disability than are unmarried workers. However, the differences in DI take-up appear to be driven by allowance probabilities conditional on applying; married workers are *less likely* to both submit an application for DI after disability onset and to report reapplying or appealing after a rejection. One potential reason that unmarried workers are less successful in the DI application process is that they are quicker to drop their applications and appeals. Married workers who suffer a work-limiting condition and ultimately receive DI benefits report waiting an average of 273 days between application and allowance, whereas unmarried workers wait only 178 days on average.

This hypothesis is also supported by evidence surrounding the onset of back pain in the HRS, which is one of the notoriously difficult-to-verify health conditions which has been responsible for a growing share of DI claims. Table C.5 shows that unmarried respondents are *less likely* than their peers to receive DI benefits after the onset of back pain, whereas married take-up of DI increases by 6 percentage points. Claims related to these conditions are particularly contentious; respondents who suffer back problems report waiting 325 days between application and benefit allowance. [29]

Accounting for observable differences correlated with household structure.

Married households tend to differ from unmarried households on several other margins that could relate to their responses to disability onset. As shown in Table C.1, they tend to be more white, tend to belong to earlier birth cohorts, and they tend to be older at the time of disability onset. This raises the question: can these observable differences alone explain the differences in responses to disability? To answer this question, I repeat the event study exercise after reweighing households so that neither the distribution of those household characteristics nor exposure to DI-related policies vary between married and unmarried households.[30] Table C.6 reports the results, showing that fixing these observable characteristics attenuates some—but not all—of the differential responses by marital status.

Testing common trends.

As a test of the common trends assumption, Table C.7 reports the average differential pre-trend for severely work-limited households in the PSID, constructed by averaging the pre-period event study coefficients. Column 1 reports this value for unmarried households, and Column 2 reports it for married households. I overall find little evidence of violations of common trends. The pre-trend test is rejected ($p < 0.05$) for the transfer income, post-tax income, and wealth of unmarried households, and is weakly rejected ($p < 0.10$) for the post-tax income of married households. It is not rejected, at any conventional significance level, for any key measures related to DI receipt, work, earned income, or spending.

The income pre-trends suggest that unmarried households (and married households, to a lesser extent) are already losing income compared to their peers prior to disability onset. However, this result can be naturally explained by mismeasurement in the timing of disability onset. A work-limiting disability first reported in year t could have in truth began at any point after their $t-1$ interview, including sometime during year $t-1$ (and consequently affect income collected over year $t-1$). If we were to take the pre-trend seriously and account for it though, it would suggest that both unmarried and married households do not experience a change in household income due to disability, though the effects on work, DI take-up, and consumption would persist.

[29] Only one unmarried respondent who suffers back problems ever reports later receiving DI benefits, which precludes a statistical comparison of the wait times for married and unmarried respondents.

[30] The policies on which I balance households include exposure to 1996 welfare reform and state-level temporary DI programs.

Table C.2—: HRS sample summary statistics

	(1) Single	(2) Partnered	(3) Partnered - Single
Panel A: Cross-Sectional Demographics			
Age	58.21***	57.67***	-0.53
	(0.07)	(0.04)	(0.08)
High School Grad	0.80***	0.76***	-0.04
	(0.01)	(0.01)	(0.01)
White	0.64***	0.75***	0.11
	(0.01)	(0.01)	(0.01)
Number of Adults	1.14***	2.01***	0.86
	(0.01)	(0.00)	(0.01)
Number of Kids	2.31***	3.33***	1.02
	(0.06)	(0.03)	(0.06)
Panel B: Cross-Sectional Health Measures			
Has Back Problem	0.18***	0.22***	0.05
	(0.01)	(0.00)	(0.01)
Has Hypertension	0.37***	0.41***	0.04
	(0.01)	(0.01)	(0.01)
Has Diabetes	0.13***	0.15***	0.02
	(0.01)	(0.00)	(0.01)
Has Cancer	0.04***	0.05***	0.00
	(0.00)	(0.00)	(0.01)
Has Lung Disease	0.05***	0.05***	0.00
	(0.00)	(0.00)	(0.01)
Has Heart Problem	0.09***	0.13***	0.04
	(0.01)	(0.00)	(0.01)
Has Stroke	0.02***	0.03***	0.01
	(0.00)	(0.00)	(0.00)
Has Psych. Problem	0.09***	0.07***	-0.02
	(0.01)	(0.00)	(0.01)
Has Arthritis	0.27***	0.33***	0.06
	(0.01)	(0.01)	(0.01)
Has Work Limitation	0.14***	0.13***	-0.01
	(0.01)	(0.00)	(0.01)
Panel C: Panel Statistics			
Households per Wave	374	1,591	1,217
Waves per Household	3.01	3.84	0.83

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the household level.

Table C.3—: Event study results for a severe work limitation of the reference person in the PSID

	Pre-Disability Means		Effects of Work-Limiting Disability		
	(1)	(2)	(3)	(4)	(5)
	Single	Partnered	Single	Partnered	(4) - (3)
Panel A: Disability Insurance Receipt and Work					
DI Receipt	0.00	0.00	0.16***	0.23***	0.08
	(0.00)	(0.00)	(0.05)	(0.05)	(0.07)
Ref. Hours Worked	1,896.68***	2,075.17***	-613.70***	-790.34***	-176.64
	(96.42)	(85.48)	(174.02)	(137.90)	(222.04)
Partner Hours Worked	41.24**	1,345.11***	-204.39***	-163.33	41.06
	(18.49)	(107.09)	(70.14)	(107.64)	(128.48)
Panel B: Household Income					
Ref. Work Inc.	20.22***	14.09***	-5.30**	-4.82***	0.48
	(1.90)	(0.92)	(2.17)	(0.94)	(2.37)
Soc. Sec. Inc.	0.30	0.40**	1.18***	1.33***	0.15
	(0.19)	(0.19)	(0.40)	(0.37)	(0.54)
Other Transfers	1.40*	0.99***	1.20**	0.88**	-0.32
	(0.78)	(0.29)	(0.51)	(0.37)	(0.63)
Work Inc. (Others)	1.11***	8.23***	-0.04	-0.29	-0.24
	(0.39)	(1.01)	(0.96)	(0.76)	(1.22)
Gross Income	23.04***	23.71***	-2.89	-2.91***	-0.02
	(1.77)	(1.46)	(2.04)	(1.07)	(2.30)
Post-Tax Income	19.07***	19.16***	-3.48	-2.35**	1.13
	(1.78)	(1.18)	(2.48)	(1.19)	(2.75)
Panel C: Spending and Wealth					
Log Spending	9.26***	9.26***	-0.18	0.04	0.22
	(0.16)	(0.11)	(0.20)	(0.14)	(0.24)
Log Wealth	10.11***	10.29***	-0.80*	-0.01	0.79
	(0.49)	(0.35)	(0.46)	(0.34)	(0.57)
Panel D: Attrition					
Leaves Sample Household	0.00	0.00	0.01	0.00	-0.01
			(0.01)	(0.01)	(0.01)
Dies	0.00	0.00	0.02	0.01	-0.01
			(0.01)	(0.01)	(0.01)
Vanishes from Sample	0.00	0.00	0.10***	-0.01	-0.11***
			(0.03)	(0.02)	(0.04)

Note: Point estimates in columns (3) and (4) are averages of the coefficients $\hat{\delta}_\ell$ of Equation C2, weighted proportionately to the distribution of household-year observations across the post-disability years ℓ. Dollar values are reported in thousands, and are adjusted for the size of the household using the OECD-modified equivalence scale. Post-tax income is imputed using NBER TAXSIM Version 27 Wealth excludes any assets associated with an owned business. The log spending (wealth) event study restricts to households on or after 1999 with positive household spending (wealth). Confidence stars indicate: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the household level.

Table C.4—: Consequences of work limitation for the reference person in the HRS

	(1)	(2)	(3)	(4)	(5)
	Panel A: Event Study Results				
	Pre-Disability Means		Effects of Work-Limiting Disability		
	Single	Partnered	Single	Partnered	Partnered - Single
Applied for DI	0.16***	0.15***	0.36***	0.26***	-0.10*
	(0.0233)	(0.0143)	(0.0523)	(0.0250)	(0.0579)
Reapplied/Appealed	0.03***	0.03***	0.08***	0.06***	-0.02
	(0.0124)	(0.0081)	(0.0282)	(0.0139)	(0.0314)
Receiving DI	0.06***	0.07***	0.14***	0.16***	0.02
	(0.0162)	(0.0092)	(0.0370)	(0.0196)	(0.0419)
	Panel B: Statistics on Work-Limited Households				
	Single	Partnered	Partnered - Single		
Days to secure DI allowance	178***	273***	96		
	(53)	(44)	(69)		

Note: Point estimates in columns (3) and (4) of the main panel are averages of the coefficients $\hat{\delta}_\ell$ of Equation C2, weighted proportionately to the distribution of household-year observations across the post-disability years ℓ. Confidence stars indicate: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the household level. Days to secure DI allowance are measured from the date of that the reference person began an application.

Table C.5—: Consequences of back problems for the reference person in the HRS

	(1)	(2)	(3)	(4)	(5)
	Panel A: Event Study Results				
	Pre-Disability Means		Back Problem		
	Single	Partnered	Single	Partnered	Partnered - Single
Applied for DI	0.04*	0.04***	0.01	0.06***	0.05
	(0.0223)	(0.0092)	(0.0273)	(0.0191)	(0.0333)
Reapplied/Appealed	0.01	0.01**	0.01	0.01	0.00
	(0.0067)	(0.0035)	(0.0154)	(0.0084)	(0.0175)
Receiving DI	0.00	0.03***	-0.02**	0.04**	0.06***
	(0.0022)	(0.0064)	(0.0094)	(0.0154)	(0.0181)

Note: Point estimates in columns (3) and (4) of the main panel are averages of the coefficients $\hat{\delta}_\ell$ of Equation C2, weighted proportionately to the distribution of household-year observations across the post-disability years ℓ. Confidence stars indicate: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the household level.

Table C.6—: Event study results for a severe work limitation of the reference person in the PSID, fixing observable differences across single and partnered workers

	Pre-Disability Means		Effects of Work-Limiting Disability		
	(1) Single	(2) Partnered	(3) Single	(4) Partnered	(5) (4) - (3)
Panel A: Disability Insurance Receipt and Work					
DI Receipt	0.00	0.00	0.18*	0.21**	0.04
	(0.00)	(0.00)	(0.10)	(0.10)	(0.15)
Ref. Hours Worked	1,896.68***	2,075.17***	-629.61*	-1,785.70***	-1,156.10***
	(96.42)	(85.48)	(325.83)	(205.03)	(384.97)
Partner Hours Worked	41.24**	1,345.11***	-408.86**	-439.06*	-30.21
	(18.49)	(107.09)	(175.33)	(251.72)	(306.76)
Panel B: Household Income					
Ref. Work Inc.	20.22***	14.09***	-3.60	-7.97***	-4.37
	(1.90)	(0.92)	(4.69)	(1.69)	(4.99)
Soc. Sec. Inc.	0.30	0.40**	1.04	1.42**	0.38
	(0.19)	(0.19)	(0.74)	(0.71)	(1.03)
Other Transfers	1.40*	0.99***	0.98	1.31**	0.33
	(0.78)	(0.29)	(0.78)	(0.57)	(0.96)
Work Inc. (Others)	1.11***	8.23***	2.81	-2.38***	-5.18
	(0.39)	(1.01)	(3.71)	(0.89)	(3.81)
Gross Income	23.04***	23.71***	1.07	-7.57***	-8.64**
	(1.77)	(1.46)	(3.90)	(1.94)	(4.35)
Post-Tax Income	19.07***	19.16***	-1.05	-7.43***	-6.38
	(1.78)	(1.18)	(4.14)	(1.78)	(4.51)
Panel C: Spending and Wealth					
Log Spending	9.37***	8.79***	-0.15	-0.08	0.07
	(0.11)	(0.25)	(0.23)	(0.30)	(0.38)
Log Wealth	9.46***	9.05***	-0.20	1.22*	1.42*
	(0.66)	(0.74)	(0.54)	(0.61)	(0.82)
Panel D: Attrition					
Leaves Sample Household	0.00	0.00	0.01	0.00	-0.01
			(0.01)	(0.01)	(0.01)
Dies	0.00	0.00	0.02	0.01	-0.01
			(0.01)	(0.01)	(0.01)
Vanishes from Sample	0.00	0.00	0.10***	-0.01	-0.11***
			(0.03)	(0.02)	(0.04)

Note: This table reports event study coefficients after reweighing single households to match partnered households on race, exposure to DI-related policies, and birth decade. See Table C.3 for other notes.

Table C.7—: Testing Pre-trends for Event Studies of a Severe Disability Onset

	Pre-Trend	
	(1) Single	(2) Partnered
Panel A: Disability Insurance Receipt and Work		
DI Receipt	0.00	-0.01
	(0.00)	(0.01)
Ref. Hours Worked	-50.90	38.08
	(84.77)	(129.44)
Partner Hours Worked	47.80	60.17
	(70.74)	(107.52)
Panel B: Household Income		
Ref. Work Inc.	-0.94	-0.44
	(1.34)	(1.12)
Soc. Sec. Inc.	0.14	0.20
	(0.10)	(0.19)
Other Transfers	-1.32**	-0.08
	(0.57)	(0.44)
Work Inc. (Others)	-0.08	-0.76
	(0.58)	(0.79)
Gross Income	-2.20	-1.09
	(1.37)	(1.32)
Post-Tax Income	-3.22**	-2.21*
	(1.50)	(1.32)
Panel C: Spending and Wealth		
Log Spending	0.16	0.02
	(0.21)	(0.14)
Log Wealth	1.46**	-0.20
	(0.64)	(0.21)
Panel D: Attrition		
Leaves Sample Household	0.00	0.00
	(0.00)	(0.00)
Dies	0.00	0.00
	(0.00)	(0.00)
Vanishes from Sample	0.00	0.00
	(0.00)	(0.00)

Note: The pre-trend test averages the negative of the pre-period event study coefficients in Equation C2. Years are weighted according to the distribution of household-year observations. A positive value means that, on average, the outcome is increasing over the pre-period for work-limited households relative to their peers. Dollar values are reported in thousands, and are adjusted for the size of the household using the OECD-modified equivalence scale. Wealth excludes any assets associated with an owned business. Confidence stars indicate: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Modeling the Tax and Transfer System

The tax and transfer system in the life cycle model of the household incorporates federal income taxes, child tax credits, FICA taxes, the EITC, unemployment insurance, food stamps, Social Security retirement, SSDI, and SSI. To the extent it is feasible, each of these features follows its real program rules as of 1996.

D1. Disability Insurance Program

Benefit payments through DI follow the true program rules for SSDI and SSI described in Section I. Individuals must explicitly choose to apply for DI. Applicants are admitted to DI with a probability depending on current health and age, consistent with the Vocational Grid used by examiners (see Chen and van der Klaauw 2008). While applying for or receiving benefits, individuals are subject to work constraints. Beneficiaries may leave DI in two ways. The first is voluntary, as individuals choose to return to work. The second is by being involuntarily removed from DI by reassessment. Reassessments occur with a probability based on one's health status. The penalty for being removed from DI is that one cannot work and receives no DI benefits in the period of removal.

The model of DI abstracts from reality in a handful of ways. First, DI applicants and beneficiaries in my model are restricted from performing *any* work. In reality, workers may earn up to Substantial Gainful Activity (SGA), which was $700 per month in 2019. In practice, few DI beneficiaries or applicants engage in work. I rationalize this constraint as the result of strategic behavior (e.g., working being viewed by examiners as a positive signal of health).

Second, I follow Low and Pistaferri (2015) in making some concessions which greatly simplify the state space of the model. This includes simplifying DI payments, which are made based on the worker's permanent wage rather than their true earnings history. The same simplification applies to Social Security retirement benefits in the model. It also includes simplifying DI eligibility; a reference person is eligible to apply for DI if he has worked for at least a year since his previous application. Ex-post, applicant reference persons in the model tend to satisfy the true, more extensive work history requirements for DI regardless. As in Low and Pistaferri (2015), SSI enters as an interaction between general welfare programs and DI receipt. SSI beneficiaries receive $5,100 per year, reduced by spousal earnings.

Third, I allow only the (male) reference persons in this model to apply for DI, not their (typically female) partners. This is mainly to avoid concerns arising from how I approximate eligibility with recent work history (since labor force attachment among these partners is weaker). Though overall women take up DI almost as much as men, these partners in the PSID are less than half as likely to take up DI as reference persons, and joint take-up of DI among both reference persons and their partners is rare.

Fourth, I approximate the uncertain and potentially lengthy application process. Eligible workers must choose to apply, which takes a full year (during which the applicant may not work). At the end of the application year, an allowance decision is realized. Beneficiaries begin receiving payments in the following year, whereas disallowed applicants may return to work (before reapplying in the future, if they choose to do so). This is intended to approximate the five month waiting period, alongside any necessary appeals to reach a final decision.[31] In this model, choosing to apply for DI represents applying *and* committing to the potentially lengthy appeals process.

Lastly, I abstract away from the health insurance aspect of work and DI receipt. In reality, health insurance is closely tied to both employment and marriage. DI receipt is an alternative mechanism

[31] About 30% of applicants are allowed onto DI within a year of applying. Another 30% are allowed within 1 to 3 years of applying, with a right tail of applicants who remain in the process even longer (French and Song, 2014).

by which individuals may acquire public health insurance through Medicare or Medicaid. Health insurance coverage may also be an important consideration for work-limited individuals, especially prior to the passage of the Affordable Care Act and the protections it provides nationally for pre-existing conditions. Limited data on health insurance coverage in the PSID precludes an adequate accounting of access to health insurance coverage for workers. Examining the role of health insurance in the welfare value of DI is consequently beyond the scope of this study, as is the case for most work in the DI literature.

D2. Joint Tax and Transfer System

Household income is subject to key features of the tax and transfer system: the EITC, federal income taxes, food stamps, unemployment insurance and Social Security retirement. Parameters of the tax and transfer system are set to their 1996 values.

Transfers through the EITC are determined by household earned income, the maximum credit amount, phase-in rate, and phase-out rate which are all a function of the number of children in the household (consistent with real program rules). The earned income level at which EITC phase-out begins is a function of both marital status and the number of children in the household.[32]

Federal taxes imposed on pre-tax income also follow the true progressive tax schedules from 1996.[33] Households take non-refundable standard deductions ($6,700 for married or $4,000 for a single filers), personal exemptions ($2,550 per person), and child tax credits ($400 per child).[34] Social security payments are subject to federal income taxes, with deductions depending on the household's earned income.

The above description of the tax system assumes that married couples file jointly. The model restricts to less educated and relatively lower-earning households which would tend to benefit from joint filing (filing separately would disqualify households for the EITC). Empirically, the large majority of married households file jointly across the income distribution. Through the jointness of the tax system and the EITC, a worker's net tax obligation is weakly decreasing in the labor force participation of their partner.

Food stamps enter into the model as a key general safety net program, which households may utilize during both the working and retirement phases of the life cycle. Eligibility and cash benefits are modeled according to real program rules. Eligibility requires that gross household income falls below 130% of the Federal Poverty Level (FPL), and income net of deductions falls below 100% FPL. Deductions include a standard $1,608 deduction, a 20% deduction for earned income, and a $2,100 deduction per dependent child. Food stamp transfers to the household are increasing in the size of the household and reduced by 30% of household income net of deductions, including DI or retirement benefits or employment earnings of a spouse.

Food stamps depart in the model depart from the true program in three ways. First, take-up in the model is automatic among eligible households and households automatically satisfy any work or work training requirements. Second, means tests in the model are performed only on income and not on wealth.[35] Third, food stamps in the model provides a cash transfer, rather than in-kind

[32] Maximum credit amounts were $347, $2,312, or$3,816 for 0, 1, or more than 1 child. Phase-in rates are 7.65%, 34%, or 40% for the same. Phase-out rates were 7.65%, 15.98%, or 21.06%. Phase-out begins at $5,700 for households with no children or $12,500 for households with children.

[33] For married households, the 1996 tax schedule consisted of marginal tax rates of 15%, 28%, 31%, 36%, and 39.6% beginning at $0, $40,100, $96,900, $147,700, and $263,750 respectively.

[34] I assume that unmarried reference persons file with single status rather than head of household status, and do not claim children for the purposes of the EITC. The reason for this is that most unmarried reference persons empirically have no children and would not qualify. Among the 28% of household-years in which unmarried reference persons are observed with children, 80% of households contain at least one other adult which renders unclear who may claim the child for head of household status.

[35] In reality, food stamps includes restrictions on liquid wealth, but not on durables, housing, or pension wealth. The model only includes a single liquid asset.

transfers.[36]

Unemployment insurance (UI) is available during the working phase of life. It provides a cash transfer in the event of exogenous job destruction.[37] Unemployment insurance rules in practice vary from state to state. Its role in this model is to dampen the consequences of job destruction. For this reason, I model unemployment insurance as a single cash payment equal to 25% of the previous year's earnings (up to a limit). This roughly approximates the average 47% replacement rate reported by the Department of Labor's Benefit Accuracy Measurement survey, with benefits typically collectible for a maximum of two quarters. With job destruction causing a fixed year-long involuntary unemployment spell, unemployment insurance has no moral hazard consequences in this model, except indirectly through interactions with long-term DI receipt.

During the retirement phase of the life cycle, individuals cannot work but instead collect social security retirement benefits. Like SSDI, retirement benefit amounts are determined by the true formula, but based on the individual's permanent wage rather than their realized past earnings. Though retirement benefits begin at the early Social Security retirement age (62), I allow individuals to collect their full retirement benefits without penalty. While spouses can in practice claim benefits based on the earnings of the reference person (or vice versa), I assume that all individuals claim benefits on the basis of their own individual earnings.

Details on the Solution of the Model

To solve the model, I numerically evaluate value functions in each year of life, beginning in the final year and working backwards. The value function in period t is the sum of a utility flow and a continuation value:

$$V_t(\boldsymbol{X}_t, S_t) \equiv \max_{\boldsymbol{C}} U(C, L^1, L^2; \boldsymbol{H}_t^1, \boldsymbol{H}_t^2) + \beta EV_{t+1}(\boldsymbol{C}, \boldsymbol{X}_t, S_t) \tag{E1}$$

where $\boldsymbol{X_t} \equiv (W_t^1, W_t^2, g, Job_t^1, Job_t^2, M_t, DI_t^1, \boldsymbol{H}_t^1, \boldsymbol{H}_t^2)$ denotes the household's non-savings state space and $\boldsymbol{C} \equiv (C, L^1, L^2, App)$ denotes household choices in period t.[38] The continuation value EV_{t+1} integrates over the risks in the model, and depends on $\boldsymbol{C}$ through its effect on the household's position in the state space in $t+1$. All choices can affect savings in $t+1$, while work and DI application choices also affect job opportunities and DI allowance in $t+1$. With no bequest motive and exogenous death at age 73, $EV_{73} \equiv 0$. This household optimization problem is subject to the constraints described in Section III.

Taking as given a solution for V_{t+1} over a discrete grid of on the period-$(t+1)$ state space $\mathcal{X}_{t+1} \times \mathcal{S}_{t+1}$, I solve numerically for $V_t(\boldsymbol{X}_t, S_t)$ at each value in a period-t grid $\mathcal{X}_t \times \mathcal{S}_t$. The grid for each year of life includes 10 points on the dimension of potential wages of the reference person, 5 points on the dimension of potential wages of the spouse, and 20 points on the dimension of savings. The grid includes all possible values of the remaining (truly discrete) dimensions of the state space. To solve numerically for V_t, I must first approximate EV_{t+1}. This requires evaluating V_{t+1} at off-grid values of savings in $t+1$ ($S_{t+1} \notin \mathcal{S}_{t+1}$), which I numerically approximate using linear interpolation.[39] Then, evaluating EV_{t+1} requires taking expectation over the various risks

[36] As with liquid vs. illiquid assets, this model does not differentiate between food and other consumption goods. Furthermore, see Hoynes and Schanzenbach (2009) for evidence that households treat food stamp transfers as equivalent to cash.

[37] voluntarily leaving work does not qualify one for UI, I assume this is observable in the model.

[38] Note that the state variables of a spouse follow the household as latent characteristics when the household is unmarried. This allows shocks to accumulate for the household's potential spouse, which will matter if the household gets married in the future.

[39] S_{t+1} may lie outside the convex hull of $\mathcal{S}_{t+1}$ if household savings are higher (lower) than the maximum (minimum) value in $\mathcal{S}_{t+1}$. I do not extrapolate above the highest point on the savings grid; the marginal value of additional savings above that point is therefore zero. Zero savings is always a point on $\mathcal{S}_t$, so extrapolation below the lower bound of the grid is impossible.

of the model:

$$EV_{t+1}(\boldsymbol{C}, \boldsymbol{X}, S) \equiv \sum_{\boldsymbol{X}_{t+1}} P(\boldsymbol{X}_{t+1}|\boldsymbol{C}, \boldsymbol{X}) V_{t+1}(\boldsymbol{X}_{t+1}, S_{t+1}) \tag{E2}$$

To numerically approximate the integral over the two continuous risk margins (wages of the two household members), I use Gaussian quadrature.

To describe the solution method in further detail, it is useful to separate retirement from the working years of life.

E1. Solution Algorithm in Retirement Years

Consider the final year of life (age 72). Households cannot work in retirement, and the household exogenously dies next year meaning that $EV_{t+1} \equiv 0$. The value function in this year of life depends only on the utility flow, so the value function depends only on the utility flow and the household optimally consumes all of its retirement benefits and savings ($C^*_{72}(\boldsymbol{X}_{72}, S_{72}) = S_{72} + \tau^{retire}(W^1_{72}, W^2_{72})$). Therefore, Equation E1 takes the form:

$$V_{72}(\boldsymbol{X}_{72}, S_{72}) = \frac{\left[C^*_{72}\exp(\theta'_1\boldsymbol{H}^1_{72} + \theta'_2\boldsymbol{H}^2_{72})\right]^{1-\gamma}}{1-\gamma}$$

For the preceding nine years of exogenous retirement t ($63 \leq t \leq 71$), households choose how much to consume versus leave for savings in the following year. Equation E1 consequently takes the form:

$$V_t(\boldsymbol{X}_t, S_t) \equiv \max_C \frac{\left[C\exp(\theta'_1\boldsymbol{H}^1_t + \theta'_2\boldsymbol{H}^2_t)\right]^{1-\gamma}}{1-\gamma} + \beta EV_{t+1}(\boldsymbol{C}, \boldsymbol{X}_t, S_t)$$

Note that in retirement, C is the only component of $\boldsymbol{C}$ chosen by households, and I solve for its optimal value C^*_t numerically using the golden search algorithm. Doing this, however, requires approximating the continuation value EV_{t+1} for arbitrary choices $\boldsymbol{C}$, This continuation value depends on $\boldsymbol{C}$ only through future savings, S_{t+1}:

$$S_{t+1}(\boldsymbol{C}, \boldsymbol{X}_t, S_t) = R[S_t + \tau^{retire}(W^1_t, W^2_t) - C]$$

The continuation value also depends on household type, health, and marital status, which govern realization of next-period health and marital states. Individual potential wages which govern retirement benefits, are carried over from their values in period t.[40] Given a solution for V_{t+1} over the grid $\mathcal{S}_{t+1}$, EV_{t+1} can be numerically solved by linear interpolation over that grid.

E2. Solution Algorithm in Working Years

Now consider the final year of working life (age 62). Households choose how much to consume *and* whether or not to work (if not on DI or fired after age 61); maximizing over the choice vector $\boldsymbol{C}$ now requires more than simply maximizing over consumption.[41] To solve for $V_{62}(\boldsymbol{X}_{62}, S_{62})$, I therefore first solve the conditional welfare values and conditional optimal consumption for fixed employment decisions. For a particular set of these choices $\boldsymbol{D} \equiv (L^1, L^2, App)$, denote the year-$t$ conditional value and policy functions $V^{\boldsymbol{D}}_t(\boldsymbol{X}_t, S_t)$ and $C^{\boldsymbol{D}}_t(\boldsymbol{X}_t, S_t)$ respectively.

[40] The grids $\mathcal{X}_t$ and $\mathcal{X}_{t+1}$ are chosen so that interpolation over potential wages is not necessary in retirement. That is, $\mathcal{X}_t = \mathcal{X}_{t+1}$ when t is a retirement year.

[41] The reference person may also choose to apply for DI, but they will receive social security benefits (which are identical in disability and retirement) next year regardless. They are therefore indifferent between applying for DI and simply remaining unemployed.

Given a particular set of discrete decisions $\boldsymbol{D}$, the conditional continuation values $EV_{63}^{\boldsymbol{D}}$ are numerically approximated in the same way the continuation value EV_t was approximated in retirement. The only difference is in how savings are determined. One must account for work decisions, unemployment insurance eligibility (encoded into L^1 and L^2 for brevity), and DI beneficiary status:

$$S_{t+1}(\boldsymbol{C}, \boldsymbol{X}_t, S_t) = R[S_t + \tau(W_t^1, W_t^2, L^1, L^2, DI_t) - C - F_{t,H_t^1}^{1,M} L^1 - F_{t,H_t^2}^{2} L^2], \qquad t \leq 62 \tag{E3}$$

With that, we can numerically solve $V_{62}^{\boldsymbol{D}}$ and $C_{62}^{\boldsymbol{D}}$ for each $\boldsymbol{D}$ using the golden search algorithm.[42] I then pick the decisions $\boldsymbol{D}$ that maximizes household welfare, which determines $\boldsymbol{C}_{62}^{*}(\boldsymbol{X}_{62}, S_{62})$ and $V_{62}(\boldsymbol{X}_{62}, S_{62})$.

Now consider the earlier working years of the life cycle ($23 \leq t < 62$). In these years, households make work and consumption decisions like they did at age 62. The reference person also faces non-trivial decisions of applying for DI (if eligible) or voluntarily leaving DI receipt at the end of period t (if receiving DI benefits in period t).[43] These choices increase the number of conditional value functions $V_t^{\boldsymbol{D}}$ and conditional consumption decisions $C_t^{\boldsymbol{D}}$ to solve and from which to pick in finding $\boldsymbol{C}_t^{*}(\boldsymbol{X}_t, S_t)$ and $V_t(\boldsymbol{X}_t, S_t)$. The savings equation mapping states and choices into S_{t+1} follows the expression in Equation E3.

The household now faces several additional sources of uncertainty that affect the calculation of EV_{t+1}. Households take expectations over future job loss, DI beneficiary status (applicants may be allowed onto DI and continued beneficiaries may be removed) and shocks to potential wages. Accounting for the discrete sources of uncertainty is straightforward, but continuous uncertainty in wages must be accommodated in how I estimate EV_{t+1}. I do this by Gaussian quadrature, for which Equation E2 is approximated as follows:

$$EV_{t+1}(\boldsymbol{C}, \boldsymbol{X}_t, S_t) \approx \sum_{\boldsymbol{X}_{t+1} \in \mathcal{X}_{t+1}} \tilde{P}(\boldsymbol{W}_{t+1}|\boldsymbol{W}_t) P(\boldsymbol{Z}_{t+1}|\boldsymbol{C}, \boldsymbol{Z}_t) V_{t+1}(\boldsymbol{X}_{t+1}, S_{t+1})$$

where $\boldsymbol{W}_t$ is the vector of potential wages in the household in year t and $\boldsymbol{Z}_t$ are the remaining discrete components of the state space (so that $\boldsymbol{X}_t \equiv (\boldsymbol{W}_t, \boldsymbol{Z}_t)$), and where $\tilde{P}(\boldsymbol{W}_{t+1}|\boldsymbol{W}_t) = \frac{P(\boldsymbol{W}_{t+1}|\boldsymbol{W}_t)}{\sum_{\boldsymbol{W} \in \mathcal{W}_{t+1}} P(\boldsymbol{W}|\boldsymbol{W}_t)}$. Note that $\sum_{\boldsymbol{W} \in \mathcal{W}_{t+1}} P(\boldsymbol{W}|\boldsymbol{W}_t) \neq 1$ because we are summing wage densities over only a finite set of future potential wages. As in each previous period of the life cycle,

In contrast to the case of $t \geq 62$ where $P(\boldsymbol{X}_{t+1}|\boldsymbol{C}, \boldsymbol{X}_t)$ depended on $\boldsymbol{C}$ only through the manner in which choices affected future savings, $P(\boldsymbol{X}_{t+1}|\boldsymbol{C}, \boldsymbol{X}_t)$ for $t < 62$ depends more broadly on links between $\boldsymbol{C}$ and $\boldsymbol{X}_{t+1}$. One can only be allowed onto DI in $t+1$ if $App_t = 1$. On the other hand, one can only be involuntarily removed from DI and left unable to work in $t+1$ if $DI_t = 1$ and $App_t = 0$ (that is, the beneficiary does not choose to submit an "application" to leave DI). Partner s can only be fired and collect unemployment insurance benefits in $t+1$ if $L_t^s = 1$.

After numerically approximating $V_t^{\boldsymbol{D}}$ and $C_t^{\boldsymbol{D}}$ for each $\boldsymbol{D}$, we can find the policy function $\boldsymbol{C}_t^{*}(\boldsymbol{X}_t, S_t)$ and the value function $V_t(\boldsymbol{X}_t, S_t)$ by picking the discrete choices that maximize household utility. This characterizes the procedure for solving V_t in each year of life until the first age of life in the model, $t = 23$.

[42] When I consider counterfactuals that include adding means tests to DI, I use a modified search algorithm that checks whether the household is better off spending down its savings to qualify for DI.

[43] The choice of voluntarily leaving DI so that one can choose to work in $t+1$ is encoded into App_t if receiving DI benefits in period t.

F1. Simulating households from the model

Having solved for the policy functions and value functions over the life cycle (with a given set of parameter values), one can use the solved model to simulate out behaviors of households. This is an important step in the indirect inference procedure that is used to fit the model parameters to data. At the start of the life cycle, the household has zero savings and draws a type g and an initial marital status. In each period to follow, the process by which I simulate household behaviors and well-being is the same. Recall from Appendix E that each $V_t^{\boldsymbol{D}}$ and each $C_t^{\boldsymbol{D}}$ was solved over the grid of values $\mathcal{X}_t \times \mathcal{S}_t$. Given the simulated state values $\boldsymbol{X}_t$ and S_t, I linearly interpolate over the three continuous dimensions of the state space (potential wages and savings) to numerically approximate $V_t^{\boldsymbol{D}}(\boldsymbol{X}_t, S_t)$ based on the values at the grid points at which $V_t^{\boldsymbol{D}}$ has been solved. I numerically approximate $C_t^{\boldsymbol{D}}(\boldsymbol{X}_t, S_t)$ in the same way, except that I first transform consumption into log terms before interpolation.

Of the choices available to the household, which are determined by $\boldsymbol{X}_t$, I pick the $\boldsymbol{D}$ that maximizes expected household utility. This determines the employment and DI application decisions of the household, and it indirectly determines the household's consumption decision $C_t^*(\boldsymbol{X}_t, S_t)$. These decisions along with the household's current state variables are used to make a draw of the household's state variables in the next year, $\boldsymbol{X}_{t+1}$ and S_{t+1}. The process is then repeated for period $t + 1$.

With the full life cycle simulated out for a household, I lastly simulate the sampling design of the PSID in order to account for its effects on empirical moments. To do this, I draw set of periods in which the household is observed, from the empirical distribution of periods in the life cycle at which households are observed in the PSID conditional on the household type g. I simulate out the sample of PSID households 10 times in this way, providing a total of 27,260 simulated households. I finally construct the targeted moments from the simulated data, comparing them to their empirical values.

F2. Simulated method of moments

The previous sub-section describes how the model simulates household behavior for a given set of model parameters Θ. Let $\hat{\beta}(\Theta)$ denote the simulated moments generated by the model with parameter Θ and $\hat{\beta}^0$ their empirically observed values, where the moments used are outlined in Section IV.D. I fit Θ by simulated method of moments, setting it so that $\hat{\beta}(\Theta)$ closely matches to $\hat{\beta}^0$ according to the criterion:

$$\min_{\Theta}(\hat{\beta}^0 - \hat{\beta}(\Theta))'\Sigma^{-1}(\hat{\beta}^0 - \hat{\beta}(\Theta)) \tag{F1}$$

where Σ^{-1} is a weighting matrix. Following Autor et al. (2019) and Blundell et al. (2016a), the weighting matrix I use is the inverse of the diagonal matrix of the empirical standard deviations for each moment. This choice is motivated by evidence from Altonji and Segal (1996) that the asymptotically efficient weighting matrix has poor small-sample properties.

To numerically solve this problem, I use a modified version of the particle swarm algorithm in the "psoptim" function published in version 1.0 of the *psoptim* R package (Ciupke, 2016). My modification to the algorithm introduces multi-threading, allowing for the particles to be evaluated in parallel. Solving and simulating of the model for a given set of parameter values, coded in C++, is also multi-threaded. This setup allows for the particles in the particle swarm to be distributed to different multicore machines, which then use the available cores on each machine to solve and simulate the model in parallel. I use 20 particles to search the parameter space. Once $\hat{\Theta}$ has been found, I compute standard errors on the parameters using the formula provided by Gourieroux et al. (1993). This procedure treats as fixed the parameters of the model that are not estimated by SMM.

Simulating Potential Welfare Benefits

As an exogenous shifter of the propensity to work, I construct a measure consistent with the measure used by Low and Pistaferri (2015). This is a measure of potential welfare benefits in the event that a reference person works an average of 20 hours per week at the federal minimum wage, with no other earned income in the household. This measure is constructed using safety net programs which experience reforms over time, and some of which vary across states as well. The programs included in this measure are: the EITC (including state supplements), food stamp payments, AFDC/TANF benefits administered by states, and unemployment insurance benefits administered by states.[44]

$$Z_{it} = Z_{it}^{EITC} + Z_{it}^{AFDC} Z_{it}^{TANF} + Z_{it}^{FS} + Z_{it}^{UI}$$

The EITC payments are a kinked function of gross earned income w_t, which depends on t through the federal minimum wage:

$$Z_{it}^{EITC} = \tau_{it}^1 w_t \mathbb{1}\{0 < w_t \leq k_t^1\} + \tau_{it}^1 k_t^1 \mathbb{1}\{k_t^1 < w_t \leq k_t^2\} + (\tau_{it}^1 k_t^1 - \tau_{it}^2 (w_t - k_t^2))\mathbb{1}\{k_t^2 < w_t < k_t^3\}$$

The parameters k_t^1, k_t^2, and k_t^3 are kinks in the EITC formula determined by federal rules and which vary from year to year. The parameters τ_{it}^1 and τ_{it}^2 are the phase-in and phase-out rates for EITC payments respectively. They vary across households i through both the number of dependent children in the household and state of residence, as many states supplement the federal EITC payment. Both federal phase-in and phase-out rates, as well as state supplements, vary over time. I follow Low and Pistaferri (2015) in disallowing EITC payments from households belonging to demographic groups (based on family size, education, state of residence, age, and calendar year) with average assets that exceed EITC limits.[45]

Prior to 1996, AFDC payments are determined as a share $r_{s(i)}^{NS}$ of the state-determined need standard (NS_{it}, varying over states, time, and family size), reduced by countable income (y_{it}^{AFDC}) according to a state-specific rate $r_{s(i)}^{y}$. Payments are also subject to an upper limit M_{it}^{AFDC} which depends on state and family size, and varies over time.

$$Z_{it}^{AFDC} = \min\{\max\{0, r_t^{NS} NS_{it} - r_t^y y_{it}^{AFDC}\}, M_{it}^{AFDC}\}$$

Countable income y_{it}^{AFDC} represents earned household income net of deductions. Those deductions include a \$90 flat deduction, in addition to a \$30 deduction for the first year of benefit receipt and a deduction equal to 1/3 of earned income for the first 4 months on a job. For the purpose of simulated benefits, a worker is always assumed to be on the first year of their job. Consistent with real program rules, benefits are 0 for households with no children, gross income above 185 percent of the need standard, or countable income above 100 percent of the need standard. AFDC benefits are only non-zero prior to welfare reform in 1996.

TANF replaces AFDC after 1996 welfare reform. TANF payments are equal to some maximum M_{it}^{TANF} which depends on state of residence, family size, and year. This amount is reduced by

[44] The federal minimum wage enters the computation of these benefits, as each of them depends on earned income.

[45] Historical federal EITC parameters are taken from Congressional Research Service reports. Historical state EITC parameters are taken from NBER TAXSIM documentation.

earned income net of a state-specific deduction formula (y_{it}^{TANF}). Households earning above an income limit (varying by state and family size, and over time) are ineligible for TANF payments.

$$Z_{it}^{TANF} = \max\{0, M_{it}^{TANF} - y_{it}^{TANF}\}$$

Food stamp benefits are equal to a maximum amount varying by year and family size (M_{it}^{FS}), which is reduced by \$0.30 for each \$1.00 of family income net of deductions (y_{it}^{FS}). Family income includes earned income, as well as other transfer payments. Deductions include a standard deduction, and deductions for: earned income (20%), dependent care, medical costs, child support payments, and excess shelter expenses. The excess shelter deduction is equal to any rent payments over 50% of family income net of other deductions and subject to a limit which varies over time. The standard deduction and dependent care deduction vary over time. Following Low and Pistaferri (2015), I set dependent care and medical deductions to zero because information on these expenses is unavailable prior to 1999. Child support and rent payments are set at average values within demographic cells (based on family size, education, state of residence, age, and calendar year).

$$Z_{it}^{FS} = \max\{10, M_{it}^{FS} - 0.3 y_{it}^{FS}\}$$

For unemployment insurance, I follow Low and Pistaferri (2015) in assuming full coverage and receipt of payments for 26 weeks. UI pays a minimum amount $m_{s(i)t}^{UI}$, up to a maximum amount $M_{s(i)t}^{UI}$ which varies across state of residence $s(i)$ and time. Payments are based on a replacement rate $r_{s(i)t}^{UI}$ for reference earnings $y_{s(i)t}^{UI}$. Additionally, some states provide an allowance for dependents, denoted b_{it}, which may vary over time. The definition of reference earnings $y_{s(i)t}$ vary across states and over time.[46] The replacement rate $r_{s(i)t}^{UI}$ likewise varies over states, and within states, over time.

$$Z_{it}^{UI} = \max\{\min\{r_{s(i)t}^{UI} y_{s(i)t}^{UI}, M_{s(i)t}^{UI}\}, m_{s(i)t}^{UI}\} + b_{it}$$

Selection Correction in the Wage Equation

The expression for the selection correction is motivated by distributional assumption on unobserved errors in the wage equation, based on the standard Roy model. Work decisions and the log wage process are respectively defined as follows:

$$\begin{aligned}
L_{it}^s =& \mathbb{1}\{\alpha_0^{s\prime} \boldsymbol{Z}_{it}^s + \alpha_1^{s\prime} \boldsymbol{A}_{it}^s + \alpha_2^{s\prime} \boldsymbol{H}_{it}^s + k_{g(i)}^s + \nu_{it}^s > 0\} \\
\ln W_{it}^s =& \beta_1^{s\prime} \boldsymbol{A}_{it}^s + \beta_2^{s\prime} \boldsymbol{H}_{it}^s + f_{g(i)}^s + \omega_{it}^s + \epsilon_{it}^s \\
\omega_{it}^s =& \omega_{it-1}^s + \xi_{it}^s, \quad \xi_{it}^s = \sum_\ell \Gamma_\ell \nu_{it-\ell}^s + \eta_{it}^s \\
\nu_{it}^s \sim& N(0,1) \\
\begin{pmatrix} \epsilon_{it}^1 \\ \epsilon_{it}^2 \end{pmatrix} \sim& N\left(\mathbf{0}, \begin{pmatrix} (\sigma_\epsilon^1)^2 & \theta_\epsilon \\ \theta_\epsilon & (\sigma_\epsilon^2)^2 \end{pmatrix}\right), \quad \begin{pmatrix} \xi_{it}^1 \\ \xi_{it}^2 \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} (\sigma_\xi^1)^2 & \theta_\xi \sigma_\xi^1 \sigma_\xi^2 \\ \theta_\xi \sigma_\xi^1 \sigma_\xi^2 & (\sigma_\xi^2)^2 \end{pmatrix}\right)
\end{aligned}$$

The unobservable terms ($\epsilon_{it}^1, \epsilon_{it}^2$) are interpreted as measurement error. They are normally distributed and strictly exogenous with mean zero, variance $(\sigma_\epsilon^s)^2$, and inter-partner correlation θ_ϵ. The terms ω_{it}^s are also unobservable, and follow a random walk.

[46] More than half of states now base payments on a worker's highest quarter of earnings, converted to average weekly earnings. The period over which earnings are taken is most often the first four of the last five completed quarters prior to the filing of a claim. State parameters are taken from reports of the US Department of Labor.

The innovations in the random walk are denoted ξ_{it}^s, and are drawn from a multivariate normal distribution, with variances $(\sigma_\xi^1)^2$ and $(\sigma_\xi^2)^2$, and inter-partner correlation θ_ξ. The stated structure allows for flexible correlations in some important ways. It allows for individuals to differently select into work as a function of their type, through the term k_g^s. The innovations of partner s may be correlated with past, contemporaneous, and future shocks to preferences of work for partner s (ν_{it}^s) according to the correlation parameters Γ_ℓ. This allows for the classic selection problem, whereby shocks to potential earnings may affect one's decision to work. However, the innovations of partner s may not be correlated with the shocks to work preferences of the other partner.

The structure also imposes restrictions on the joint selection problem of spouses. It rules out the possibility that shocks to preferences for work, ν_{it}^s, are correlated *across* partners. It also implies that innovations to one's wage do not affect selection into work for one's partner, hence why ξ_{it}^s can be expressed as a function of only $\boldsymbol{\nu}_i^s$ and not both $\boldsymbol{\nu}_i^1$ and $\boldsymbol{\nu}_i^2$. The term η_{it}^s represents variation in wages which is uncorrelated with one's own work preferences by construction. The specification of ξ_{it}^s and ν_{it}^s together imply that η_{it}^1 and η_{it}^2 are jointly normally distributed with a correlation that depends on θ_ξ.

H1. Estimation

The model for work is estimated for both partners using a probit regression. The main coefficients are presented in Appendix Table H.1. I estimate parameters of the wage process by taking first differences, which eliminates the group effects $f_{g(i)}^s$ and accumulated values of the random walk error term. To accommodate the movement from annual to biannual surveying in 1997, I take differences over p periods $p \in \{1, 2\}$:

$$\Delta^p \ln W_{it}^s = \beta_1^{s\prime} \Delta^p \boldsymbol{A}_{it}^s + \beta_2^{s\prime} \Delta^p \boldsymbol{H}_{it}^s + \sum_{j=0}^{p-1} \xi_{it-j}^s + \Delta^p \epsilon_{it}^s$$

An individual chooses to work only when $\nu_{it}^s > -\alpha_0^{s\prime} \boldsymbol{Z}_{it}^s - \alpha_1^{s\prime} \boldsymbol{A}_{it}^s - \alpha_2^{s\prime} \boldsymbol{H}_{it}^s - k_{g(i)}^s \equiv -h_{it}^s$. Selection in the wage equation enters through the ξ_{it-j}^s terms.

$$\begin{aligned}
E[\Delta^p \ln W_{it}^s | L_{it} = 1, L_{it-p} = 1] =& \beta_1^{s\prime} \Delta^p \boldsymbol{A}_{it}^s + \beta_2^{s\prime} \Delta^p \boldsymbol{H}_{it}^s \\
& + \sum_{j=0}^{p-1} E[\xi_{it-j}^s | \nu_{it}^s > -h_{it}^s, \nu_{it-p}^s > -h_{it-p}^s] \\
=& \beta_1^{s\prime} \Delta^p \boldsymbol{A}_{it}^s + \beta_2^{s\prime} \Delta^p \boldsymbol{H}_{it}^s \\
& + \sum_{j=0}^{p-1} E[\sigma_\xi^s \rho_j^s \nu_{it}^s + \sigma_\xi^s \rho_{j-p}^s \nu_{it-p}^s | \nu_{it}^s > -h_{it}^s, \nu_{it-p}^s > -h_{it-p}^s] \\
=& \beta_1^{s\prime} \Delta^p \boldsymbol{A}_{it}^s + \beta_2^{s\prime} \Delta^p \boldsymbol{H}_{it}^s \\
& + \sum_{j=0}^{p-1} \sigma_\xi^s \rho_j^s \lambda_{it}^s + \sigma_\xi^s \rho_{j-p}^s \lambda_{it-p}^s
\end{aligned} \tag{H1}$$

where $\lambda_{it}^s = \frac{\phi(h_{it}^s)}{\Phi(h_{it}^s)}$ is the inverse Mills' ratio and ρ_ℓ^s is the correlation between the shock to preferences for work and the ℓth lag (or lead if $k < 0$) of the permanent wage shock.[47]

[47] With an instrument used to estimate the first stage decision to work, identification of the work and wage system does not rely on the distributional assumption of normality. The selection expression can be viewed as a non-parametric approximation

The wage growth parameters are estimated consistently by including estimates of λ^s_{it} and λ^s_{it-p} (derived from the model for the first stage decision to work) as covariates on the right hand side of the log wage growth regression:

$$\text{(H2)} \qquad \Delta^p \ln W^s_{it} = \beta^{s\prime}_1 \Delta^p \boldsymbol{A}^s_{it} + \beta^{s\prime}_2 \Delta^p \boldsymbol{H}^s_{it} + \sum_{j=0}^{p-1} \alpha^j_1 \hat{\lambda}^s_{it} + \alpha^j_2 \hat{\lambda}^s_{it-p} + \text{residual}_{it}$$

For the (male) reference persons who have high labor force participation, there is little room for selection to introduce bias into the wage regressions. For the (female) partners though, labor force participation is lower and the potential for selection is greater.

H2. Estimation of Productive Fixed Effects

Given consistent estimators of the wage growth parameters as described previosuly, consider the expected residual log wage for partner s of household i:

$$\begin{aligned} E[\widetilde{\ln W^s_{it}}|i, L_{it} = 1] = & \ E[\ln W^s_{it} - \hat{\beta}^{s\prime}_1 \boldsymbol{A}^s_{it} - \hat{\beta}_2{}' \boldsymbol{L}^s_{it}|i, L_{it} = 1] = f^s_{g(i)} + E[\omega^s_{it} + \epsilon^s_{it}|i, L_{it} = 1] \\ = & \ f^s_{g(i)} + \sum_{\ell=0}^{t} E[\xi^s_{it-\ell}|i, L_{it} = 1] \\ = & \ f^s_{g(i)} + \underbrace{\lambda^s_{it} \sum_{\ell=0}^{t} \sigma^s_\xi \rho^s_\ell}_{\text{bias}} \end{aligned}$$

The second equality uses that productive type effects are constant conditional on i. The second line uses the exogeneity of ϵ_{it} and the recursive definition of ω_{it}. The final equality provides an expression for the bias term arising from selection in $\widetilde{\ln W^s_{it}}$.

I account for selection bias in the household average of $\widetilde{\ln W^s_{it}}$ by subtracting $\hat{\alpha}^0_1 \hat{\lambda}^s_{it}$ from it, which is an estimate for $\sigma^s_\xi \rho^s_0 \lambda^s_{it}$. This is only part of the full bias term, but it is the largest of the components if shocks to work preferences are more correlated with contemporaneous shocks to wages than shocks to wages in past years.

H3. Estimation of Wage Variance

As described above, the two partners draw their innovations for the wage equation from a joint normal distribution, allowing for a correlation (θ_ξ) in contemporaneous productivity shocks in the household. This distribution has the following variance-covariance matrix.

$$\Sigma \equiv \begin{bmatrix} (\sigma^1_\xi)^2 & \theta_\xi \sigma^1_\xi \sigma^2_\xi \\ \theta_\xi \sigma^1_\xi \sigma^2_\xi & (\sigma^2_\xi)^2 \end{bmatrix}$$

Estimation of the three parameters defining Σ, and of the covariance matrix of the measurement error process, is based on the p-period difference of residual log wages ($p \in \{1, 2\}$). That observed variable is defined in Equation 10, and its mean is an estimator for $\Delta^p(\omega^s_{it} + \epsilon^s_{it})$.

Given the assumptions on ϵ^s_{it}, ξ^s_{it}, and ν^s_{it} stated at the beginning of this section, the variance-covariance parameters are estimated alongside several nuisance parameters by generalized method

of a true selection term in the case that the distributional assumptions are violated.

of moments. The system used in GMM includes the second order moment, first order moment and autocovariance $\Delta^p \widetilde{\ln W_{it}^s}$. These respectively take the form:

$$
\begin{aligned}
E[(\Delta^p(\omega_{it}^s + \epsilon_{it}^s))^2 | \mathbf{L}_{i,(t-p,t)}^s = \mathbf{1}] = \sigma_\xi^{s2} & \left(p - h_{it}^s \lambda_{it}^s \sum_{j=0}^{p-1} (\rho_j^s)^2 - h_{it-p}^s \lambda_{it-p}^s \sum_{j=1}^{p} (\rho_{-j}^s)^2 \right) \\
& + 2\sigma_\xi^{s2} \left(\sum_{j=0}^{p-1} \rho_j^s \rho_{j-p}^s \lambda_{it}^s \lambda_{it-p}^s \right) \\
& + \sigma_\xi^{s2} \left(\sum_{j=0}^{p-1} \sum_{\substack{\ell=0 \\ \ell \neq j}}^{p-1} (\rho_j^s \lambda_{it}^s + \rho_{j-p}^s \lambda_{it-p}^s)(\rho_\ell^s \lambda_{it}^s + \rho_{\ell-p}^s \lambda_{it-p}^s) \right) \\
& + 2\sigma_\epsilon^{s2} \\
E[\Delta^p(\omega_{it}^s + \epsilon_{it}^s) | \mathbf{L}_{i,(t-p,t)}^s = \mathbf{1}] = & \sum_{j=0}^{p-1} \sigma_\xi^s \rho_j^s \lambda_{it}^s + \sigma_\xi^s \rho_{j-p}^s \lambda_{it-p}^s \\
E[\Delta^p(\omega_{it}^s + \epsilon_{it}^s) \Delta^r(\omega_{it-p}^s + \epsilon_{it-p}^s) \mathbf{L}_{i,(t-p,t)}^s = \mathbf{1}] = & -\sigma_\epsilon^{s2}
\end{aligned}
$$

The vector $\mathbf{L}_{i,(t-p,t)}^s$ contains indicators for work of partner s of household i in periods $t-p$ and t.[48] These three equations along with instruments derived from variation in simulated welfare benefits are sufficient to pin down the variance of the wage innovations and the variance of the noise terms, as in Low and Pistaferri (2015).

In order to pin down the intra-household correlation in wage innovations and in measurement errors (the parameters θ_ξ and θ_ϵ), the above system of equations is augmented with a final set of equations: the intra-household correlation in $\Delta^p \widetilde{\ln W_{it}^s}$:

$$
\begin{aligned}
E[\Delta^p(\omega_{it}^1 + \epsilon_{it}^1) \Delta^p(\omega_{it}^2 + \epsilon_{it}^2) | \mathbf{L}_{(t-p,t)}^{(1,2)} = \mathbf{1}] &= E\left[\left(\sum_{j=0}^{p-1} \xi_{it-j}^1 \right) \left(\sum_{j=0}^{p-1} \xi_{it-j}^2 \right) | \mathbf{L}_{(t-p,t)}^{(1,2)} \right] \\
& \quad + E[\epsilon_{it}^1 \epsilon_{it}^2 + \epsilon_{it-p}^1 \epsilon_{it-p}^2] \\
&= \left(\sum_{j=0}^{p-1} \rho_j^1 \lambda_{it}^1 + \rho_{j-p}^1 \lambda_{it-p}^1 \right) \left(\sum_{j=0}^{p-1} \rho_j^2 \lambda_{it}^2 + \rho_{j-p}^2 \lambda_{it-p}^2 \right) \\
& \quad + p \sigma_\xi^1 \sigma_\xi^2 \theta_\xi + 2\sigma_\epsilon^1 \sigma_\epsilon^2 \theta_\epsilon
\end{aligned}
$$

where $\mathbf{L}_{(t,t-p)}^{(1,2)}$ is a vector containing indicators for work of both partners in household i in both periods $t-p$ and t.

[48] Note that for some random variable X, $E[X | \mathbf{L}_{(t-p,t)}^s = \mathbf{1}] \equiv E[X | \nu_{it}^s > -h_{it}^s, \nu_{it-p}^s > -h_{it-p}^s]$

Table H.1—: Work probit model parameters

	Head	Partner
Severe Dis.	-1.387**	
	(0.560)	
Mod Dis.	-0.954*	-2.054***
	(0.498)	(0.410)
Age	0.130***	0.119***
	(0.017)	(0.016)
$Age^2/100$	-0.198***	-0.168***
	(0.019)	(0.020)
Married	0.275***	
	(0.059)	
f_1	-1.046**	0.587
	(0.449)	(0.405)
f_2	0.174	0.687*
	(0.451)	(0.391)
f_3	-0.006	0.638
	(0.454)	(0.389)

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are clustered at the household level.