# Appendix (for online publication)

## A Survey and administrative data

**Key measures in the PSID.** The Panel Study of Income Dynamics (PSID) is a panel survey of between 5,000 and 8,000 households. The PSID was annual from 1968 to 1997 and has been administered every two years since then. The PSID has two key advantages for our analysis. First, it includes measures of consumption expenditures, which enables us to quantify the insurance value of U.S. disability programs in a highly flexible way without modeling the budget constraint in detail. Second, its long panel structure allows us to measure the characteristics and accumulated experiences of households before USDP entry or disability onset.

In much of our analysis, we classify households with a working-age head (25–65 years old) by their USDP status and health status in 2016 (measured in the 2017 wave), the latest available year, and then track them back to their entry onto USDP and the onset of a health condition. We follow the disability literature and use a measure of severity in the PSID based on self-reported work-limiting health problems (e.g., Low and Pistaferri, 2015). We classify households as "more-severe" if they report that a health condition limits "a lot" the amount of work that the head or spouse can do or that the head or spouse "can do nothing" as a result of that condition. We classify households as "less-severe" if they report that they have no health condition or that their health conditions limit "somewhat," "just a little," or "not at all" the amount of work the head and spouse can do. Appendix Table I1 reproduces the exact questions and responses. We also consider a broader definition that includes "somewhat" in addition to "a lot," and a definition that incorporates mental health in addition to "a lot."[42]

A household is "more-severe" (M) if either the head or the spouse has a more-severe health condition, and "less-severe" (L) otherwise. Similarly, we classify households as disability recipients or non-recipients based on their self-reported USDP status; a household is a recipient (DI) if either the head or the spouse reports receiving disability benefits, and a non-recipient (NDI) otherwise (so the share of households receiving disability benefits is larger than the share of individuals receiving disabil-

[42] We define having a mental health condition as reporting that a doctor or other health professional ever told the respondent that they had depression, psychiatric issues, or "loss of memory or mental ability."

ity benefits). From the latest wave of the PSID (2017), the share of working-age households in each USDP-by-severity receipt group is the following: 4.3% disability recipients with more-severe conditions (M-DI), 6.1% disability recipients with less-severe conditions (L-DI), 3.4% non-recipients with more-severe conditions (M-NDI), and 86% non-recipients with less-severe conditions (L-NDI).[43]

These figures indicate that, at least based on the measures in the PSID (measurement error in which we discuss shortly), there are non-negligible mismatches with respect to health relative to a benchmark of a household receiving disability benefits if and only if the head or spouse has a more-severe health condition: Among households reporting a more-severe condition, over one third (40%) report not receiving disability benefits. Among households reporting that they receive disability benefits, more than half (53%) report not having a more-severe condition. Given these magnitudes, the characteristics of these mismatch groups are critical to the overall value of USDP.

We use the PSID to measure the following household characteristics, with the exact questions and response options reproduced in Appendix Tables I1–I5. For *consumption*, we use the PSID's detailed consumption expenditures data, which includes spending on food, housing and utilities, transportation, education, child care, and health care. We use total consumption spending excluding health care expenditures to avoid artificially inflating the consumption (and therefore the utility) of households with more-severe health conditions. For *adverse non-health events*, we measure head and spouse job loss, distressed moves, and divorce. We take steps to focus these measures on external events or shocks, rather than choices. For head and spouse job loss, we limit the measure to job separations that the household reports as involuntary, namely layoffs or firings, firm closings or moves, and strikes. For moves, we limit the measure to what the PSID calls "involuntary reasons" (eviction, health reasons, divorce, etc.), "purposive consumptive reasons" (less space, less rent); and ambiguous or other reasons (e.g., need to save money). The PSID does not include any information about the reasons for divorce, so we include all divorces. For *resources*, we

[43]Household shares are weighted by the family weight variable, as are other PSID statistics that we report. Fewer than 1% of households have more than one more-severe member or more than one disability recipient. PSID sample sizes are relatively small (316 M-DI, 443 L-DI, and 250 M-NDI) but large enough that our main results are statistically significant (see Appendix Figure F1). The sample sizes change to 429 M-DI, 330 L-DI, and 523 M-NDI for the "a lot + somewhat" severity definition; and 463 M-DI, 296 L-DI, and 1088 M-NDI for the "physical + mental health" definition.

use marriage, banking relationship, and homeownership as our primary measures of resources in the PSID. We also use health insurance coverage (public, private, or any); help from relatives and non-relatives; and other public transfer programs, including food stamps, TANF, UI, and workers' compensation.

**Key measures in SSA administrative data.** The PSID has important limitations: a relatively small sample size, self-reported USDP status, and no information about applications for SSDI or SSI. To address these shortcomings, we supplement it with SSA administrative data. In the SSA administrative data, our measure of severity is at which stage of the disability determination process the recipient is allowed. We classify disability recipients as "more-severe" if they are allowed at the initial state DDS level and as "less-severe" if they are allowed upon appeal. We get nearly identical results when we instead use as the measure of severity the SSA adjudicator's judgment of the likelihood of medical improvement. We use the administrative records to determine USDP status.

Sample sizes are much larger in SSA administrative data. For mass layoffs, which are measured for a random 10% of applicants between 1990 and 2016, the sample size is 1.9 million. The financial outcomes are measured for a larger share of applicants but over fewer and more recent years: 1.4 million for mass layoffs, 13 million for bankruptcy, and 2.2 million for foreclosure and eviction. For each sample, between two-thirds and three-quarters of disability recipients are classified as more-severe.

We use SSA administrative data for measures of two types of adverse non-health events. For *adverse financial events*, we use the Master Beneficiary Record linked to nationwide financial records from Deshpande, Gross and Su (2019) on bankruptcies, evictions, and foreclosures to measure these adverse financial events for disability recipients and applicants. For *mass layoffs*, we use the Continuous Work History Sample, a 10% sample of all SSNs in the United States and their earnings histories (including employer EINs), to identify firms where mass layoffs occur, and link disability applicants and recipients to these mass layoff events. We define mass layoffs as events in which the number of workers at a firm that had at least 150 workers declines by at least 30 percent.[44]

We also use SSA administrative data linked to the Survey of Income and Program Participation (SIPP) to verify our main cross-sectional results and to observe SSDI

[44]We use a higher threshold than Von Wachter, Song and Manchester's (2011) threshold of 50 workers because the CWHS has only a 10% sample of individuals.

and SSI applications. In the SIPP-SSA linked data, we use self-reported work-limiting health problems as the measure of severity. We classify households as "more-severe" if they report that the head or spouse has a physical, mental, or other health condition that prevents them from working at all, and "less-severe" otherwise. Appendix Table I1 reproduces the exact questions and responses. We use the 1996, 2001, 2004, 2008, and 2014 panels from the SIPP. Each panel has a sample of around 20,000 to 37,000 households, and we combine panels to further increase precision. The SIPP-SSA linked data also allow us to observe applications in the SSA administrative data, which is important for understanding mechanisms behind mismatches with respect to health. The main disadvantages of the SIPP are that each panel lasts only a few years and it does not have a broad consumption measure.

# B Robustness to measurement error

There is no single ideal measure of severity, since there is no single view about what types of conditions or states of the world are "more-severe." Some might think of severity in terms of health alone, while others might consider work limitations; some might consider only physical disabilities, while others might include mental health conditions. Moreover, only a limited set of measures is available in the relevant datasets, and these measures, like other variables, likely include measurement error. In the PSID, our main measure is based on the question about work limitations. While not the only way to measure severity, it is closely related to what disability programs aim to insure and the standard question used in the disability insurance literature (e.g., Low and Pistaferri, 2015). In supplementary analyses, we also account for mental health conditions and less strict definitions of more-severe. In the SSA administrative data, the measure of severity is whether the recipient is allowed at the initial state DDS level, rather than upon appeal. We get nearly identical results when we use an alternative definition based on the adjudicator's judgment of how often the recipient's medical condition should be reviewed to determine whether it has improved. For both samples, our goal is not to determine the exact number of more- and less-severe recipients, but rather to compare, using the best available data, the characteristics of more- and less-severe recipients. In practice, the severity distributions of the more- and less-severe groups likely overlap.

Meyer, Mok and Sullivan (2015) show that there is non-negligible measurement error in disability receipt in the PSID. Although we cannot quantify this error directly in the PSID, we can use the SIPP-SSA linked data using pooled SIPP waves (1996, 2001, 2004, 2008) to get a sense of the likely bias by investigating analogous errors in the self-reported measure in the SIPP. Our findings suggest that measurement error usually, though not always, works against our main conclusions, meaning that measurement error makes the evidence appear less supportive of our key findings than it really is:

- Robustness of Fact 2: For the PSID results, one concern is that households that experience serious non-health shocks might be more likely to incorrectly report having a more-severe health condition. In this case, we would likely understate the non-health shocks of L-DI both overall and relative to M-DI, which works against the Fact 2 result. Another concern is that disability recipients might be more likely to incorrectly report having a more-severe health condition, perhaps to rationalize their disability receipt. In this case, some L-DI would be mistakenly categorized as M-DI. If this error is random, then it would again work against this result. Measurement error in USDP status is also a concern for the PSID results. We investigate this issue and find that measurement error likely makes L-DI look better off than they actually are (which works against this result), but likely makes M-DI look worse off (which works toward this result).[45] For the SSA results, there could be error in the severity measure that we use based on adjudicator judgment, but this error is harder to characterize without access to third-party medical records.

[45] Although we cannot quantify this error directly in the PSID, we can measure it in the SIPP-SSA linked data using pooled SIPP waves (1996, 2001, 2004, 2008) to get a sense of the likely bias arising from our use of the PSID measure. The true L-DI group differs in two ways from the self-reported L-DI: It includes L households that incorrectly report not receiving disability benefits (67% of true L-DI in the SIPP), and it excludes L households that incorrectly report receiving disability benefits (33% of the size of true L-DI). These misclassifications appear to work against the L-DI vs. L-NDI and L-DI vs. M-DI results, since true L-DI appear more disadvantaged than self-reported L-DI in the SIPP. For example, true L-DI are less likely to be married (58% vs. 62%) and less likely to own a home (59% vs. 64%). The true M-DI group differs in two ways from the self-reported M-DI: It includes M households that incorrectly report not receiving disability benefits (39% of true M-DI), and it excludes M households that incorrectly report receiving disability benefits (14% of the size of true M-DI). In contrast to the L-DI misclassifications, these misclassifications appear to work toward the L-DI vs. M-DI result, since true M-DI appear more disadvantaged than self-reported M-DI in the SIPP. For example, true M-DI are less likely to be married (50% vs. 54%) and less likely to own a home (51% vs. 58%). The L-NDI group is so large that USDP status misclassifications are tiny relative its size.

- Robustness of Fact 3: With respect to the PSID severity measure, non-recipients likely have fewer incentives to misreport their severity than recipients. However, given that L-NDI is a large group, if some fraction of them misreport as M-NDI, then this would make M-NDI appear more advantaged than they truly are and work toward the finding in Fact 3. We also consider measurement error in USDP status in the PSID. We find that measurement error likely makes M-NDI look worse off than they are and M-DI look better off than they are, both of which work against the finding in Fact 3.[46]

- Robustness of normative results: Measurement error in self-reported *disability receipt* in the PSID is both substantial and likely to work against our key normative findings on net (see footnotes 45 and 46 for details). Its main effect is likely to make reported L-DI appear to be less valuable disability recipients than actual L-DI, which works against our finding that L-DI benefits are valuable. Its secondary effect is likely to make M-DI and M-NDI appear to be more similar than they really are, which works against our findings that M-DI benefits are valuable and that M-NDI benefits would have been costly though toward our finding that L-DI benefits have roughly comparable value to M-DI benefits. Measurement error in self-reported *severity* tends to work toward our results, especially our finding that M-NDI benefits would have been costly. However, bias created by measurement error in severity seems likely to be smaller in magnitude (see footnote 46 for details).

# C Estimates incorporating health insurance

As a baseline we focus on the cash benefit component of USDP. Excluding the health insurance component as a baseline follows the literature on the welfare effects of USDP (e.g., Bound et al., 2004; Chandra and Samwick, 2009; Low and Pistaferri,

[46] Again using the SIPP-SSA linked data, we find that the true M-NDI group in the SIPP differs in two ways from the self-reported M-NDI: It includes M households that incorrectly report receiving disability benefits (18% of true M-NDI), and it excludes M households that incorrectly report not receiving disability benefits (52% of the size of true M-NDI). These misclassifications appear to work against the M-NDI vs. M-DI result, since true M-NDI appear more advantaged than self-reported M-NDI in the SIPP. For example, true M-NDI are more likely to be married (66% vs. 59%) and less likely to own a home (60% vs. 53%). M-DI misclassifications, reported in footnote 45, also appear to work against the the M-NDI vs. M-DI result, since true M-DI appear more disadvantaged than self-reported M-DI.

2015; Cabral and Cullen, 2019; Meyer and Mok, 2019), since for many recipients the health insurance component mainly displaces other forms of subsidized health care cost sharing they otherwise would have received (see, e.g., Liebman, 2015). In this appendix section, we present alternative scenarios that include the health insurance component of USDP. We test robustness to making different assumptions about the health insurance component, with results reported in Appendix Table F3.

The most important factor is likely the extent to which the health insurance component of USDP displaces other forms of health care cost-sharing, including private health insurance (see the thoughtful discussion in Bound et al., 2004), other sources of government health insurance, and "informal health insurance" from charity care, bankruptcy, and bad debt. Displacement of other sources of government insurance has likely become even more important in recent years. As Liebman (2015) notes, "[G]iven the expansions of Medicaid eligibility and subsidies for insurance purchase enacted as part of the Patient Protection and Affordable Care Act of 2010, many disability recipients would today be receiving free or heavily subsidized health insurance even if they were not receiving disability benefits" (131).[47] As a result, we view our baseline analysis, which assumes a zero net value and zero net cost of the health insurance component, as a fairly close approximation to the full impact of USDP, including its health insurance component.

Still, we test the robustness of our conclusions to a wide range of alternative assumptions about health insurance, especially since the health insurance component may be more valuable to more-severe recipients. In these robustness tests, our main aim is to test the robustness of our finding about the relatively high value of L-DI benefits *compared to* M-DI benefits. We make assumptions that are conservative relative to that aim. In particular, assumptions that overstate the value of the health insurance component will overstate the value of M-DI benefits relative to L-DI benefits,

[47]Note that the key input to our analysis is the value of the health insurance component not to the average recipient but to "inframarginal" recipients who would earn below the USDP earnings limit even if they were not receiving disability benefits, since to first-order the value of the health insurance component, like the value of the cash component, is zero in states of the world in which the individual would otherwise, if not for receiving disability benefits, have earned more than the earnings limit. These states in which counterfactual no-benefit earnings are very low are likely states in which counterfactual no-health-insurance-component-of-disability-benefits protection against health care costs (from Medicaid, means-tested subsidies, and the rest of the safety net) is substantial—and so the net value and cost of the health insurance component of disability benefits is small.

since more-severe individuals tend to benefit more from health insurance.[48,49]

**Medical Expenditure Panel Survey.** For some of these robustness tests, we require richer, higher-quality data on health care costs than are available in the PSID. To this end, we use the Household Component of the Medical Expenditure Panel Survey, a nationally representative survey of the U.S. civilian non-institutionalized population with rich data on health care consumption, health care costs, and health care payments. We use the 2018 wave, the most recent wave available as of this writing.[50] We focus on individuals aged 25–64 with non-missing values of the relevant variables. This leaves us with a sample of 14,775 individuals. We use questions about "Reason not working" to classify individuals as having more- or less-severe health conditions. We classify individuals who respond "Unable to work because ill/disabled" in all three survey rounds as having a more-severe health condition.[51] We classify individuals whose responses to the "Reason not working" question in all three survey rounds are valid responses *other than* "Unable to work because ill/disabled" as less-severe. By these measures, approximately 4% of the population is more-severe (and 688 individuals in the sample) and 94% is less-severe (13,673 individuals in the sample), with the remaining 2% having responded "Unable to work because ill/disabled" in one or two of the three rounds.[52]

[48]Overstating the value of the health insurance component also tends to overstate the value of USDP relative to cost-equivalent tax cuts and other policies.

[49]Note that health insurance, because of its moral hazard cost and insurance value, is potentially worth a different amount to the individual than the cost to the government of providing it. So unlike the equations in the main text, which assume that the USDP benefit $b$ takes the same value in the equations on the value of USDP to recipients and those on the cost of USDP to the government (e.g., equations (7) and (8)), with health insurance, the USDP benefit in a particular state can have unequal values in the value and cost equations. Moreover, with health insurance the ex post value and cost of USDP benefits can differ across more- versus less-severe states, due to differences in the value and cost of the health insurance component.

[50]We leave the health care spending variables in 2018 dollars rather than deflating to 2016 dollars, the units of the rest of the monetary variables in our normative analysis, in order to err on the side of overstating the value of the health insurance component and so to be conservative with respect to the value of L-DI relative to M-DI benefits.

[51]Other possible responses to this question are "Could not find work," "Retired," "Going to school," "Taking care of home or family," "Don't want to work," and "Other."

[52]This classification of more- versus less-severe in MEPS yields a lower share of more-severe individuals than the PSID classification, which suggests a higher severity threshold in MEPS than in our baseline analysis. This would tend to overstate the value of the health insurance component to more-severe individuals, which is conservative with respect to our result on the relative value of L-DI to M-DI benefits.

**Value and net cost of going from no health insurance to full health insurance.** We take three main routes from available data and evidence to rough estimates of the net value to recipients and net cost to the government of the health insurance component of USDP in different types of states of the world. The first is based on out-of-pocket health spending by households, the second on total health care payments, and the third on Finkelstein, Hendren and Luttmer's (2019) estimates of the value and net cost of Medicaid based on the Oregon Health Insurance Experiment.

*Measures based on out-of-pocket health spending.* Perhaps the most natural approach to quantifying the value and cost of health insurance is based on the out-of-pocket health spending of households that do not have health insurance. To first-order and in the absence of private costs of relying on the safety net, the ex post value of going from status quo health insurance to full health insurance coverage of all health care costs in a particular state of the world is status quo out-of-pocket health spending in that state.[53] This idea is the basis for one of our sets of measures of the value of health insurance, as follows. We use the MEPS data to estimate mean out-of-pocket health spending by individuals without health insurance in each of the two severity categories. We find that average annual out-of-pocket health spending by individuals without health insurance is $521 among individuals with less-severe health conditions and $1,471 among individuals with more-severe health conditions. We assume that these are the net costs of providing health insurance to less- and more-severe individuals, respectively. As for the value of health insurance, we make two different assumptions. For our "main" analysis based on out-of-pocket health spending, we assume that the ex ante value of health insurance is 0.85 times its net cost, since 0.85 is the midpoint of Finkelstein, Hendren and Luttmer's (2019) range of estimates of the ratio of ex-ante value to net cost of 0.5–1.2. This yields a value of $443 for less-severe individuals and $1,250 for more-severe individuals. In alternative "ratio-minimizing" scenarios, we assume that the ex ante value of health insurance is 1.2 times its net cost, the maximum of Finkelstein, Hendren and Luttmer's (2019) range of estimates. This yields a value of $625 for less-severe individuals and $1,765 for more-severe individuals.

*Measures based on total health care payments.* In the second approach, we start

[53]This follows from the usual envelope theorem logic that to first order the value of a change in constraints is the associated "mechanical effect"—the reduction in net expenditure that would occur if not for behavioral responses—since re-optimization gains are second order.

with total, gross health care costs and scale them down to reflect that even individuals without formal health insurance receive substantial help in paying their health care costs from the safety net, including charity care, bad debt, and bankruptcy. We use the MEPS data to estimate mean total payments for health care, by health insurers and households, for individuals with health insurance in each severity category. We estimate that mean annual total payments for health care are $5,662 for less-severe individuals and $20,459 for more-severe individuals. In order to go from these gross, total health care costs to net health care costs—net of support the individual would receive from the safety net if they did not have formal health insurance—we multiply by the share of total costs that are paid by uninsured individuals (as opposed to by other parties). In their review, Finkelstein, Mahoney and Notowidigdo (2018) cite estimates of the share of total costs paid by uninsured individuals of one-fifth to one-third. For our main analysis based on total health care payments, we use the midpoint of this range, which is about 27%. This yields a net cost of $1,510 for less-severe individuals and $5,456 for more-severe individuals. In the ratio-minimizing scenarios, we use the maximum of this range (one-third). This yields a net cost of $1,887 for less-severe individuals and $6,820 for more-severe individuals. Once again, to estimate the values of the health insurance component, we scale the net costs by Finkelstein, Hendren and Luttmer's (2019) estimates of the ratio of value to net costs. For our main analysis based on total health care payments, we use the midpoint of Finkelstein, Hendren and Luttmer's (2019) range of 0.85. This yields a value of $1,283 for less-severe individuals and $4,637 for more-severe individuals. In the ratio-minimizing scenarios, we assume that the ex ante value of health insurance is 1.2 times its net cost, the maximum of Finkelstein, Hendren and Luttmer's (2019) range of estimates (while at the same time assuming that the share of total costs paid by uninsured individuals is the maximum of its range, one third). This yields a value of $2,265 for less-severe individuals and $8,184 for more-severe individuals.

*Measure based on Finkelstein, Hendren and Luttmer (2019).* We use Finkelstein, Hendren and Luttmer's (2019) maximum estimate of the value of health insurance based on the Oregon Health Insurance Experiment, which is $1,675 (from their "complete information" approach). We also use their estimate of the net cost to the government, which is $1,448. Finkelstein, Hendren and Luttmer (2019) do not estimate heterogeneity in value or cost, so in this specification we use the same value and cost for both more- and less-severe individuals. This specification is useful in part by pro-

viding benchmarks for the values and costs based on the more roundabout approaches (based on out-of-pocket health spending and total health care payments) that we use in order to account for heterogeneity across more- and less-severe individuals.

**Displacement of other forms of health care cost sharing.** As discussed above, for many disability recipients the health insurance component of USDP likely mainly displaces other sources of subsidized health insurance that they otherwise would have received. In our main analyses with health insurance, we assume that the average, across all disability recipients, of the net value to recipients and cost to the government of the health insurance component is one-fourth of the value and cost of going from no health insurance to full health insurance. One-fourth is the share of new SSDI recipients that lacked health insurance during the waiting period between when they were awarded disability benefits and when their Medicare benefits started in the late 1990s (Riley, 2006; Short and Weaver, 2008; Livermore et al., 2009). The corresponding share is likely to be significantly lower today given subsequent expansions of Medicaid and health insurance subsidies. For the ratio-minimizing scenarios, we report results in which the average value of the health insurance component equals the full value of going from no health insurance to full health insurance, implicitly assuming zero displacement of other forms of subsidized health insurance.[54]

**Results.** We present results for the different health insurance scenarios in Appendix Table F2. We find that plausible assumptions about the health insurance component tend to modestly increase the surplus from both M-DI and L-DI benefits. However, since the health insurance component is worth more in M-DI than in L-DI states, the ratio of the per-recipient surplus of L-DI to M-DI benefits decreases from its baseline of 0.78. Under our baseline health insurance scenarios, the ratio decreases to 0.77 (based on out-of-pocket health spending) and 0.74 (based on total health spending). Under the ratio-minimizing scenarios that assume zero displacement of alternative sources of subsidized health insurance, the ratio decreases to 0.71 (based on out-of-pocket health spending) and 0.54 (based on total health spending). These results suggest that accounting for the health insurance component of USDP does not materially change the conclusion that L-DI benefits have significant value not only relative to cost-equivalent tax cuts but even relative to M-DI benefits.

[54]Note that prior to disability receipt, L-DI are less likely than M-DI to have health insurance (67% vs. 75%, from Appendix Figure G8b), so assuming a common displacement share will likely again tend to overstate the value of health insurance to M-DI relative to L-DI.

# D Policy analysis

The normative analysis in Section 5 estimates the ex ante value and cost of receiving disability benefits in different states of the world. This analysis does not speak directly to the welfare implications of reforms to USDP. In this appendix section, we consider three commonly proposed policy reforms: limiting receipt by less-severe individuals, decreasing benefit levels, and decreasing the allowance rate. The underlying theory and empirical implementation are very similar to those of the main analysis in Section 5. The key difference is that here we consider the costs (or cost savings) associated with application responses. In the main analysis, the cost to the government of providing disability benefits to the individual in a particular state comprises the costs of both the disability benefit itself and the induced behavioral responses in the "directly affected state" (i.e., that in which the benefit is received). Here, we also include the costs (cost savings) of the additional applications and awards induced (discouraged) by the reform in states in which the reform affects the individual's decision about whether to apply for disability benefits. Application responses tend to increase the costs of expansions of disability insurance, since such expansions tend to increase applications, and increase the cost savings of contractions, since such contractions tend to decrease applications. To quantify such costs, we use evidence from the literature on the responsiveness of disability applications to changes in disability policy and in economic conditions. Section D.4 provides details about these cost calculations.

Appendix Table F5 reports the results. The following sections provide details of the analysis for each of the three policies in turn.

## D.1 Reducing receipt among less-severe individuals

This hypothetical policy reform reduces the number of less-severe recipients (L-DI). This could potentially be achieved by increasing investments in health assessments during the adjudication process or in reassessments of existing recipients. Many of the reforms that have been proposed to reduce SSDI spending (see, e.g., McCrery and Pomeroy, 2016; Greszler, Moffit and Schaefer, 2019) aim to reduce receipt among less-severe individuals in different ways, such as stricter medical eligibility criteria (reduce mental and back pain allowances and allowances based on vocational grid); temporary disability insurance (time-limit benefits of individuals with less-severe conditions); full funding of continuing disability reviews (remove recipients who have medically

improved). Here, we abstract from the particular means used to reduce receipt among less-severe individuals to try to shed light on the main efficiency-related costs and benefits of such changes. In order to be conservative relative to our conclusion about L-DI benefits being valuable, we ignore any costs associated with the investments necessary to reduce receipt among less-severe individuals, including any unintended effects on individuals with more-severe health conditions (e.g., from higher application costs or higher likelihoods of having their applications rejected).

We assume that the affected states are representative of less-severe recipient states as a whole in terms of marginal utility and counterfactual earnings. We estimate the foregone ex ante value of receiving disability benefits in the less-severe states that no longer receive disability benefits as a result of the reform exactly as in Section 5: See equation (1), plugging in $\Omega_b = \Omega_{L\text{-}DI}$. The cost savings to the government comprise not only the foregone costs of the benefits and associated behavioral responses in the directly-affected states (as in Section 5) but also the foregone costs of applications and awards in other states in which the individual is discouraged from applying for disability benefits as a result of the reform. See Section D.4 for details.

## D.2 Decreasing the allowance rate

This reform decreases the allowance rate of adjudicators. For example, a recent Heritage Foundation proposal calls for eliminating vocational grid allowances and reviewing outlier judges (Greszler, Moffit and Schaefer, 2019). We consider a decrease in the allowance rate of all examiners, meaning that applicants on the margin of allowance are less likely to be allowed. Maestas, Mullen and Strand (2013) estimate that the SSA examiner to which an applicant is assigned affects the ultimate award of 23 percent of applicants, those between the 20th and 43rd percentiles of the distribution in terms of disability severity. In this counterfactual, the 23 percent of "applicants on the margin of program entry" all become marginally less likely to be awarded disability benefits, while the award decisions on the 77 percent of infra-marginal applicants are unchanged. We consider a decrease in the allowance rate that reduces the household recipiency rate by 10%, roughly from 11% to 10% of states.

We assume that the states in which the individual is an applicant on the margin of program entry, of which direct measures do not exist, are representative of less-severe recipient states in terms of marginal utility and counterfactual earnings. This

is based on the idea that applicants on the margin of program entry tend to be those with less-severe health conditions relative to inframarginal approved applicants. We estimate the ex ante value of receiving disability benefits in the states in which the individual is an applicant on the margin of program entry exactly as in Section 5: See equation (1), plugging in $\Omega_b = \Omega_{ampe}$, the set of states in which the individual is an applicant on the margin of program entry with respect to this reform. We assume that the probability of being on the margin of program entry with respect to this reform is ten percent of the disability recipiency rate $p(\Omega_{ampe}) = 0.1 \times p(\Omega_{DI})$. The cost savings to the government comprise not only the foregone costs of the benefits and associated behavioral responses in the directly-affected states (as in Section 5) but also the foregone costs of applications and awards in other states in which the individual is discouraged from applying for USDP as a result of the reform. See Section D.4 for details.

## D.3 Decreasing benefit levels

This reform marginally decreases disability benefit levels. For example, the President's FY2021 budget proposes cutting SSDI benefits in various ways, including reducing retroactive benefits and offsetting payments from other programs, and a recent Heritage Foundation proposal calls for a "flat antipoverty benefit" that would cut payments substantially (Greszler, Moffit and Schaefer, 2019; United States Office of Management and Budget, 2020). We consider a reform that decreases benefit levels by 10%, from \$13,000 to \$11,700. The analysis of such a reform is closely related to the main analysis of the value of receiving disability benefits in different states described in Section 5, here focused on all states in which the individual receives disability benefits. The key differences are that here we consider application responses and that the ex post value and cost of a change in the benefit level differ from those of receiving benefits versus not.

We estimate the ex ante value of receiving disability benefits in the states in which the individual is an "inframarginal recipient" of disability benefits—i.e., receives disability benefits under status quo policies—similarly to as in Section 5. The only difference is that the ex post value of the benefit decrease equals the full benefit decrease amount in all inframarginal recipient states, not just those states in which counterfactual no-benefit earnings were below the USDP limit. Formally, the ex post

value in state of the world $\omega$ of a $\Delta b$ decrease in the disability benefit is

$$WTP_\omega = -DI_\omega \times \Delta b, \tag{11}$$

where $DI_\omega$ equals one if the individual receives disability benefits in state $\omega$ under status quo policies (i.e., before the benefit decrease). Recall from equation (7) that the ex post marginal value in state $\omega$ of a USDP-like expansion of the constraint is

$$WTP_\omega = \mathbb{1}(z_\omega^{DI=0} \leq \bar{z}) \times b,$$

where $z_\omega^{DI=0}$ is counterfactual earnings without disability benefits in state $\omega$, $\bar{z}$ is the earnings limit for disability recipients, and $b$ is the disability benefit. The difference arises because any costs to inframarginal recipients of the earnings limit or other restrictions, though diminishing the full value to them of receiving disability benefits, do not change the marginal value of receiving lower benefits. We plug this ex post value into equation (1), plugging in $\Omega_b = \Omega_{DI}$, the set of states in which the individual is an inframarginal recipient.

The cost savings to the government comprise not only the foregone costs of the benefits and associated behavioral responses in the directly-affected states (in which the individual is an inframarginal recipient) but also the foregone costs of applications and awards in other states in which the individual is discouraged from applying for disability benefits as a result of the reform. See Section D.4 for details.

## D.4 The fiscal externality of changes in U.S. disability programs

For each policy counterfactual, we calculate the associated costs to the government based on the available evidence from the literature on behavioral responses to disability benefits and to economic conditions. We follow Bound et al.'s (2004) thorough procedure for mapping key behavioral elasticities into the marginal cost to the government of SSDI expansions, and we use recent evidence on such elasticities from quasi-experimental studies. Although this is the best evidence we know of, we emphasize the substantial uncertainty about the fiscal externality costs of reforms to USDP. Producing additional evidence on these key parameters is a high priority for future research.

Bound et al. (2004) combine evidence from a wide range of sources into a microsimulation model to estimate the net marginal cost to the government of a one percent increase in disability benefits. This model quantifies the effects of the increase in SSDI benefits in terms of: (i) reducing SSI and Food Stamp benefits received by current beneficiaries, (ii) increasing federal and state taxes paid by current beneficiaries, (iii) increasing disability benefits due to successful new applicants, (iv) decreasing taxes paid by recipients and new applicants and their families, and (v) increasing transfers from SSI, Food Stamps, and AFDC to new applicants. In magnitude, (iii) disability benefits paid to new applicants is by far the largest, followed by (iv) reduced taxes paid by recipients and new applicants and their families. See Bound et al.'s (2004) Table 1 (page 2500).

We depart from Bound et al. (2004) in two main ways. First, our counterfactuals consider changes in disability (both SSDI and SSI) awards and benefits to non-elderly beneficiaries, whereas Bound et al. (2004) consider an increase in SSDI benefits to all beneficiaries. Since the behavioral responses to changes in disability benefits are driven by younger people, our focus on changes in disability awards and benefits to non-elderly beneficiaries tends to increase the fiscal externality as a fraction of the "mechanical cost" (the cost were there no behavioral responses) of the increase in benefits relative to Bound et al. (2004).

Second, we use more recent evidence on the key behavioral elasticity. The key elasticity determining the size of the fiscal externality is the elasticity of applications with respect to benefits. Unfortunately, the evidence on this key elasticity is sparse. Bound et al.'s (2004) preferred estimates are based on time series evidence from the 1960s and 1970s, which suggest elasticities around 0.5. More recently, Black, Daniel and Sanders (2002) and Charles, Li and Stephens Jr (2018) have used local economic booms and busts to estimate an elasticity of disability receipt with respect to earnings. We follow Meyer and Mok (2019) in translating their elasticity with respect to earnings (0.3 in absolute value) into an elasticity with respect to disability benefits, based on the assumption that what determines applications is the replacement rate. This suggests an elasticity of around 0.2. We scale the subset of cost estimates in Bound et al. (2004) that should scale in this key elasticity (the ones driven by new applications and awards) by the ratio of our preferred estimate of 0.2 to Bound et al.'s (2004) preferred estimate of 0.5, i.e., by $0.2/0.5 = 0.4$. This tends to reduce the fiscal externality relative to Bound et al.'s (2004) preferred estimate.

Based on this evidence, we estimate a fiscal externality per \$1 of greater net transfers to inframarginal recipients—increased disability benefits less decreased means-tested transfers received and increased tax payments made—of \$0.34. This means that increasing net transfers to inframarginal disability recipients by \$1 costs the government \$1.34. This is somewhat smaller than Bound et al.'s (2004) baseline estimate of \$1.50 because the effect of using a smaller elasticity dominates the effect of not increasing the benefits of elderly recipients. For each of our disability reform counterfactuals and each of our ex ante risk types, we scale the mechanical cost of the reform by one plus the fiscal externality, $(1 + FE) = 1.34$, to estimate the full cost to the government associated with the reform. In principle, the fiscal externalities of different reforms or of the same reform for different risk types could differ, but there is little evidence on this important issue.

# E Decomposing the markup on disability benefits

Section 5.4 decomposes the markup associated with increasing different sets of actual and hypothetical disability benefits into components that reflect the underlying across- and within-health transfers. Here, we derive the key equation in Section 5.4, equation (10), which is

$$M_\Omega(\Omega_b) = \underbrace{Cov_h\left[\frac{p(\Omega_b|h)}{p(\Omega_b)}, E\left(\widehat{\lambda}|h\right)\right]}_{\text{Insurance against health risk}} + \underbrace{E_h\left\{\frac{p(\Omega_b|h)}{p(\Omega_b)}[1-p(\Omega_b|h)]\left[E\left(\widehat{\lambda}|\Omega_b,h\right) - E\left(\widehat{\lambda}|\sim\Omega_b,h\right)\right]\right\}}_{\text{Insurance against non-health risk}}.$$

For a given set of actual or hypothetical disability benefits in states $\Omega_b \subseteq \Omega_\theta$, decompose the benefit received in a particular state $\omega$ into the sum of (i) the mean benefit received in states in the same health category as $\omega$, $h_\omega \in \{\text{L, M}\}$, and (ii) a within-health category transfer from states in which the individual does not receive a benefit ($h = h_\omega$ and $\sim \Omega_b$) to those in which they do ($h = h_\omega$ and $\Omega_b$):

$$b_\omega = \underbrace{E(b|h = h_\omega)}_{\text{Health-contingent benefit}} + \underbrace{[b_\omega - E(b|h = h_\omega)]}_{\text{Within-health transfer}}.$$

To ease notation, we suppress the conditioning on risk type $\theta$. The equations should be understood to apply to a particular risk type.

**Health-contingent benefits, $E(b|h = h_\omega)$.** State $\omega$'s health-contingent benefit is the mean benefit received in states in the same health category as $\omega$, $h_\omega \in \{\text{L, M}\}$, $E(b|h = h_\omega) = p(\Omega_b|h)b$. The "mechanical effect"—the expected marginal value—of an increase in the overall benefit level $b$ operating through the health-contingent benefit is the benefit recipiency rate in states in the same health category, $p(\Omega_b|h)$.

The ex-ante marginal value of an increase in a single health-contingent benefit component (i.e., the health-contingent benefit of a particular health category), measured in terms of non-labor income in all states $\Omega$ (i.e., willingness to pay out of non-labor income in all states for the increase in the health-contingent benefit), is

$$MV_h^{across} = p(h)p(\Omega_b|h)E\left(\widehat{\lambda}|h\right),$$

where $\widehat{\lambda}$ is the marginal utility of income normalized to be mean one, $\widehat{\lambda} \equiv \lambda/E(\lambda)$.

Summing across health categories, the ex-ante marginal value of an increase in all health-contingent benefits components is

$$\begin{aligned}
\sum_h MV_h^{across} &= \sum_h p(h)p(\Omega_b|h)E\left(\widehat{\lambda}|h\right) \\
&= E_h\left[p(\Omega_b|h)E\left(\widehat{\lambda}|h\right)\right] \\
&= E_h\left[p(\Omega_b|h)\right]E_h\left[E\left(\widehat{\lambda}|h\right)\right] + Cov_h\left[p(\Omega_b|h), E\left(\widehat{\lambda}|h\right)\right] \\
&= p(\Omega_b)\left[1 + Cov_h\left(\frac{p(\Omega_b|h)}{p(\Omega_b)}, E\left(\widehat{\lambda}|h\right)\right)\right],
\end{aligned}$$

where we have used the fact that $E\left(\widehat{\lambda}\right) = 1$. The left-most term in the last line, $p(\Omega_b)$, is the mechanical effect, so $Cov_h\left(\frac{p(\Omega_b|h)}{p(\Omega_b)}, E\left(\widehat{\lambda}|h\right)\right)$ is the markup. This markup comes from the across-health targeting due to differential recipiency in some health categories relative to others. It is increasing in the extent to which recipiency is—and so health-contingent benefits are—concentrated in health categories with higher mean marginal utility.

**Within-health category transfer, $b_\omega - E(b|h = h_\omega)$.** The within-health category transfers shift resources from those states in a given health category in which the individual does not receive a benefit ($h = h_\omega$ and $\sim \Omega_b$) to those in which they do ($h = h_\omega$ and $\Omega_b$). State $\omega$'s within-health category transfer is $(b - E(b|h = h_\omega)) \geq 0$ if the individual receives a benefit in $\omega$ and $-E(b|h = h_\omega) \leq 0$ if the individual does

not receive a benefit in $\omega$. Note that the ex ante expected value of the within-health category transfer is zero. The mechanical effect of an increase in the overall benefit level $b$ operating through the within-health transfers is $(1-p(\Omega_b|h))$ if the individual is a disability recipient and $-p(\Omega_b|h)$ if not.[55]

The ex-ante marginal value of an increase in the within-health category transfer component is

$$\begin{aligned} MV_h^{within} &= p(h)p(\Omega_b|h)E\left(\widehat{\lambda}|h,\Omega_b\right)[1-p(\Omega_b|h)] + p(h)p(\sim\Omega_b|h)E\left(\widehat{\lambda}|h,\sim\Omega_b\right)[-p(\Omega_b|h)] \\ &= p(h)p(\Omega_b|h)[1-p(\Omega_b|h)]\left[E\left(\widehat{\lambda}|h,\Omega_b\right) - E\left(\widehat{\lambda}|h,\sim\Omega_b\right)\right]. \end{aligned}$$

Summing across health categories, the ex-ante marginal value of an increase in all within-health transfers components is

$$\begin{aligned} \sum_h MV_h^{within} &= \sum_h p(h)p(\Omega_b|h)[1-p(\Omega_b|h)]\left[E\left(\widehat{\lambda}|h,\Omega_b\right) - E\left(\widehat{\lambda}|h,\sim\Omega_b\right)\right] \\ &= E_h\left[p(\Omega_b|h)[1-p(\Omega_b|h)]\left[E\left(\widehat{\lambda}|h,\Omega_b\right) - E\left(\widehat{\lambda}|h,\sim\Omega_b\right)\right]\right]. \\ &= p(\Omega_b)E_h\left[\frac{p(\Omega_b|h)}{p(\Omega_b)}[1-p(\Omega_b|h)]\left[E\left(\widehat{\lambda}|h,\Omega_b\right) - E\left(\widehat{\lambda}|h,\sim\Omega_b\right)\right]\right]. \end{aligned}$$

This entire value is insurance value, since the ex ante expected within-health transfer is zero.

**Combined value of health-contingent benefits and within-health transfers.** Summing the marginal values of the health-contingent benefits and the within-health category transfers, the full ex-ante marginal value of increasing the overall benefit level $b$ of benefits in states $\Omega_b$, $\sum_h \left\{MV_h^{across} + MV_h^{within}\right\}$ is

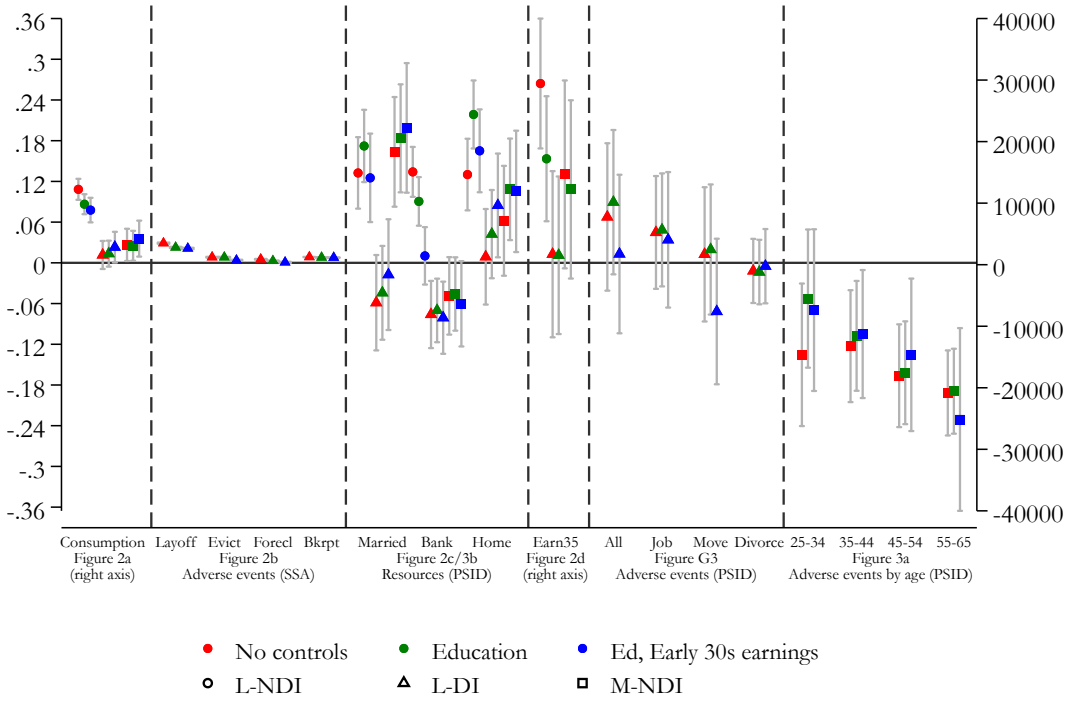
$$= p(\Omega_b)\left\{1 + \underbrace{Cov_h\left[\frac{p(\Omega_b|h)}{p(\Omega_b)}, E\left(\widehat{\lambda}|h\right)\right] + E_h\left\{\frac{p(\Omega_b|h)}{p(\Omega_b)}[1-p(\Omega_b|h)]\left[E\left(\widehat{\lambda}|h,\Omega_b\right) - E\left(\widehat{\lambda}|h,\sim\Omega_b\right)\right]\right\}}_{M_\Omega(\Omega_b)}\right\}.$$

The first term, $p(\Omega_b)$, is the mechanical effect of an increase in benefits in $\Omega_b$ states. The terms within the large curly braces are therefore one plus the markup on transfers from all states to recipients, $M_\Omega(\Omega_b)$, as in equation (10) in the main text.

[55] The sum of the mechanical effects of an increase in the overall benefit level $b$ operating through the health-contingent benefit and the within-health category transfers is 1 for recipient states and 0 for non-recipient states, which is the mechanical effect of an increase in benefit levels, as it must be.

# F Appendix Figures and Tables (Online)

Figure F1: Facts 2 and 3 with Ability-Group Controls and Confidence Intervals




Notes: Figure presents all outcomes from Figures 2, 3, and G3 with controls for ability group. The figure plots the coefficients on indicators for each USDP-by-severity group (L-NDI, L-DI, and M-NDI; M-DI is the excluded group) from a regression of the outcome on those indicators and ability controls. The three groups are shown using different shapes. We use three specifications for ability controls: no controls (which reproduces the coefficients from Figures 2 and 3), education controls (less than high school, high school graduate, some college, and college plus), and education plus average annual earnings (deciles) between ages 31–35. The three specifications are shown using different colors. The "Consumption" and "Earn35" outcomes use the right axis, and the other outcomes use the left axis. The variable "All" for Figure G3 indicates experiencing at least one of the adverse events (job loss, involuntary move, or divorce). Controlling for age 31–35 earnings over-controls for differences in ex ante risk type to the extent that earnings at that time reflect not only ex ante earning ability but also ex post realizations of shocks.

Table F1: Value of disability benefits in different states: Heterogeneity by education level

| | Mean markup ($E(M_\omega)$) | Counterfactual no-benefit earnings ($) | Surplus per household ($) ($p(\Omega_\tau)s$) | Surplus per recipient HH ($) ($p(\Omega_\tau)s/p(\Omega_b)$) | USDP value/value of cost-equiv tax cuts ($(s+EAG)/EAG$) | MVPF |
|---|---|---|---|---|---|---|
| Full pop | | | | | | |
| DI | 0.91 | 5,191 | 924 | 8,718 | 1.64 | 1.76 |
| M-DI | 0.95 | 3,557 | 488 | 9,858 | 1.73 | 1.85 |
| L-DI | 0.88 | 6,623 | 436 | 7,719 | 1.57 | 1.68 |
| M-NDI | 0.64 | 22,591 | -74 | -2,203 | 0.86 | 0.93 |
| HS dropouts | | | | | | |
| DI | 0.46 | 3,775 | 1,255 | 5,567 | 1.41 | 1.39 |
| M-DI | 0.50 | 3,751 | 647 | 6,067 | 1.45 | 1.43 |
| L-DI | 0.43 | 3,796 | 608 | 5,118 | 1.38 | 1.36 |
| M-NDI | 0.44 | 11,011 | 108 | 1,541 | 1.11 | 1.10 |
| HS graduates | | | | | | |
| DI | 0.66 | 5,135 | 931 | 5,977 | 1.44 | 1.52 |
| M-DI | 0.62 | 3,416 | 432 | 6,187 | 1.46 | 1.53 |
| L-DI | 0.70 | 6,530 | 499 | 5,806 | 1.43 | 1.50 |
| M-NDI | 0.44 | 23,088 | -174 | -3,855 | 0.76 | 0.80 |
| Some college | | | | | | |
| DI | 1.56 | 5,094 | 1,663 | 15,787 | 2.16 | 2.36 |
| M-DI | 1.60 | 3,525 | 942 | 17,289 | 2.28 | 2.48 |
| L-DI | 1.53 | 6,776 | 721 | 14,176 | 2.04 | 2.22 |
| M-NDI | 0.74 | 25,059 | -90 | -2,906 | 0.82 | 0.90 |
| College or above | | | | | | |
| DI | 1.13 | 8,415 | 271 | 7,764 | 1.57 | 1.95 |
| M-DI | 1.27 | 3,682 | 142 | 10,153 | 1.75 | 2.17 |
| L-DI | 1.03 | 11,596 | 129 | 6,158 | 1.45 | 1.80 |
| M-NDI | 1.15 | 33,704 | -49 | -3,025 | 0.83 | 1.03 |

Notes: Table presents statistics (given by the column names) associated with receiving disability benefits in each of several sets of states of the world (row names). These states are those in which the household receives USDP (DI), including when more-severe (M-DI) and less-severe (L-DI), and those in which the household does not receive USDP when more-severe (M-NDI). Markup is defined in equation (3) and calculated using PSID consumption excluding health care expenditures. Counterfactual no-benefit earnings are actual observed earnings for NDI states and are inferred from French and Song's (2014) estimates of the effects of disability benefits on earnings for DI states, as described in the text. Surplus $s$, defined in equation (6), is in units of government revenue per $\Omega_\tau$ state, so $p(\Omega_\tau)s$ is in units of government revenue per household and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient household. Value relative to that of cost-equivalent tax cuts is derived in footnote 27. USDP benefits for recipient households in the "some college" education category create the biggest surplus, perhaps because these households have substantial earnings to lose but lack the alternative insurance options of households with a college education. Monetary amounts are in 2016 dollars per year. Sample sizes for M-DI, L-DI, and M-NDI: 313, 438, 248 for "Full pop;" 74, 125, 51 for "HS dropouts;" 120, 166, 91 for "HS graduates;" 93, 101, 69 for "Some college;" and 26, 46, 37 for "College or above."

Table F2: Value of disability benefits in different states: Robustness to assumptions about health insurance

| | HI displacement share | HI value ($) | Surplus per household ($) $(p(\Omega_\tau)s)$ | Surplus per recipient HH ($) $(p(\Omega_\tau)s/p(\Omega_b))$ | USDP value / value of cost-equiv tax cuts $((s+EAG)/EAG)$ | $MVPF$ | Ratio of L-DI to M-DI surplus per recipent |
|---|---|---|---|---|---|---|---|
| Baseline | | | | | | | |
| DI | 1 | N/A | 924 | 8,718 | 1.64 | 1.76 | 0.78 |
| M-DI | 1 | N/A | 488 | 9,858 | 1.73 | 1.85 | |
| L-DI | 1 | N/A | 436 | 7,719 | 1.57 | 1.68 | |
| M-NDI | 1 | N/A | -74 | -2,203 | 0.86 | 0.93 | |
| Health insurance based on out-of-pocket health care spending (main) | | | | | | | |
| DI | 0.75 | 820 | 937 | 8,837 | 1.64 | 1.76 | 0.77 |
| M-DI | 0.75 | 1,250 | 498 | 10,054 | 1.72 | 1.85 | |
| L-DI | 0.75 | 443 | 439 | 7,771 | 1.57 | 1.68 | |
| M-NDI | 0.75 | 1,250 | -75 | -2,236 | 0.86 | 0.93 | |
| Health insurance based on total health care spending (main) | | | | | | | |
| DI | 0.75 | 2,849 | 969 | 9,136 | 1.63 | 1.75 | 0.74 |
| M-DI | 0.75 | 4,637 | 524 | 10,583 | 1.71 | 1.83 | |
| L-DI | 0.75 | 1,283 | 445 | 7,868 | 1.56 | 1.68 | |
| M-NDI | 0.75 | 4,637 | -78 | -2,325 | 0.87 | 0.94 | |
| Health insurance based on out-of-pocket health care spending (ratio-minimizing) | | | | | | | |
| DI | 0 | 1,158 | 1,038 | 9,786 | 1.67 | 1.80 | 0.71 |
| M-DI | 0 | 1,765 | 573 | 11,568 | 1.77 | 1.90 | |
| L-DI | 0 | 625 | 465 | 8,224 | 1.58 | 1.70 | |
| M-NDI | 0 | 1,765 | -60 | -1,783 | 0.90 | 0.97 | |
| Health insurance based on total health care spending (ratio-minimizing) | | | | | | | |
| DI | 0 | 5,029 | 1,420 | 13,394 | 1.75 | 1.88 | 0.54 |
| M-DI | 0 | 8,184 | 881 | 17,784 | 1.87 | 2.01 | |
| L-DI | 0 | 2,265 | 540 | 9,548 | 1.62 | 1.74 | |
| M-NDI | 0 | 8,184 | -9 | -257 | 0.99 | 1.07 | |
| Health insurance based on Finkelstein, Hendren and Luttmer's (2019) maximum estimates | | | | | | | |
| DI | 0 | 1,675 | 1,076 | 10,145 | 1.67 | 1.80 | 0.79 |
| M-DI | 0 | 1,675 | 566 | 11,428 | 1.76 | 1.89 | |
| L-DI | 0 | 1,675 | 510 | 9,020 | 1.60 | 1.72 | |
| M-NDI | 0 | 1,675 | -62 | -1,857 | 0.89 | 0.97 | |

Notes: Table presents statistics associated with receiving disability benefits in each of several sets of states of the world. "Baseline" is the baseline specification in which the health insurance (HI) component contributes zero to the value and cost of disability benefits. See Appendix Section C for details on the specifications with HI. "HI displacement share" is the share of the gross value and cost of the HI component that displaces other sources of HI and so does not contribute to the net value and cost of disability benefits. In the "main" specifications, the estimate of 0.75 comes from estimates that one-fourth of new SSDI recipients in the late 1990s lacked HI during the waiting period before the HI benefit started (Livermore et al., 2009). In the "ratio-minimizing" (i.e., L-DI to M-DI surplus ratio) specifications, we report results based on zero assumed displacement as extreme upper bounds on the net value of the HI component. The "main" specifications based on out-of-pocket and total health spending assume that HI is worth 0.85 times its net cost, the midpoint of Finkelstein, Hendren and Luttmer's (2019) ("FHL's") range of estimates. The "ratio-minimizing" specifications assume that HI is worth 1.2 times its net cost, the maximum of FHL's range of estimates. The specifications based on out-of-pocket health spending (oop) assume that the net cost of HI for severity group $sev \in \{L, M\}$ is $E(oop|HI = 0, sev)$, which we estimate in the Medical Expenditure Panel Survey (MEPS). The specifications based on total health spending (tot) assume that the net cost of HI for severity group $sev \in \{L, M\}$ is $\alpha \times E(tot|HI = 1, sev)$. We estimate $E(tot|HI = 1, sev)$ in the MEPS. The "main" specification based on total health spending assumes that $\alpha$ is the midpoint of the range of the share of total costs that uninsured individuals pay out-of-pocket reported by Finkelstein, Mahoney and Notowidigdo (2018) ("FMN") in their review, which is about 27%. The "ratio-minimizing" specification based on total health spending assumes that $\alpha$ is the maximum of the range FMN report: one-third. The specifications based on FHL's estimates of the value and cost of Medicaid use their maximum estimated value. See Table 1 for the definitions of the sets of states and estimated parameters. Monetary amounts are in 2016 dollars per year.

Table F3: Value of disability benefits in different states: Robustness to assumptions about earnings responses

| | Mean markup ($E(M_\omega)$) | Counterfactual no-benefit earnings ($) | Surplus per household ($) ($p(\Omega_\tau)s$) | Surplus per recipient HH ($) ($p(\Omega_\tau)s/p(\Omega_b)$) | USDP value/value of cost-equiv tax cuts ($(s+EAG)/EAG$) | MVPF |
|---|---|---|---|---|---|---|
| Baseline | | | | | | |
| DI | 0.91 | 5,191 | 924 | 8,718 | 1.64 | 1.76 |
| M-DI | 0.95 | 3,557 | 488 | 9,858 | 1.73 | 1.85 |
| L-DI | 0.88 | 6,623 | 436 | 7,719 | 1.57 | 1.68 |
| M-NDI | 0.64 | 22,591 | -74 | -2,203 | 0.86 | 0.93 |
| Earnings responses based on severity rather than diagnosis | | | | | | |
| DI | 0.91 | 6,366 | 877 | 8,273 | 1.60 | 1.72 |
| M-DI | 0.95 | 4,434 | 473 | 9,547 | 1.69 | 1.82 |
| L-DI | 0.88 | 8,060 | 404 | 7,157 | 1.51 | 1.63 |
| M-NDI | 0.64 | 22,591 | -74 | -2,215 | 0.86 | 0.93 |
| Earnings responses multiplied by four | | | | | | |
| DI | 0.91 | 14,028 | 432 | 4,073 | 1.27 | 1.36 |
| M-DI | 0.95 | 12,027 | 330 | 6,659 | 1.44 | 1.54 |
| L-DI | 0.88 | 15,782 | 102 | 1,806 | 1.12 | 1.20 |
| M-NDI | 0.64 | 22,591 | -82 | -2,441 | 0.85 | 0.92 |

Notes: Table presents statistics (given by the column names) associated with receiving disability benefits in each of several sets of states of the world (row names). "Baseline" is the baseline specification. "Earnings responses based on severity rather than diagnosis" uses French and Song's (2014) estimates of earnings responses to disability benefits based on severity rather than diagnosis (as in the baseline specification). "Earnings responses multiplied by four" assumes that earnings responses to disability benefits are four times those estimated by French and Song (2014). The sets of states are those in which the household receives USDP (DI), including when more-severe (M-DI) and less-severe (L-DI), and those in which the household does not receive USDP when more-severe (M-NDI). Markup is defined in equation (3) and calculated using PSID consumption excluding health care expenditures. Counterfactual no-benefit earnings are actual observed earnings for NDI states and are inferred from French and Song's (2014) estimates of the effects of disability benefits on earnings for DI states, as described in the text. Surplus $s$, defined in equation (6), is in units of government revenue per $\Omega_\tau$ state, so $p(\Omega_\tau)s$ is in units of government revenue per household and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient household. Value relative to that of cost-equivalent tax cuts is derived in footnote 27. Monetary amounts are in 2016 dollars per year. Sample sizes: 313 M-DI, 438 L-DI, 248 M-NDI.

Table F4: Value of disability benefits in different states: Robustness to assumptions about marginal utility

| | Mean markup ($E(M_\omega)$) | Counterfactual no-benefit earnings ($) | Surplus per household ($) ($p(\Omega_\tau)s$) | Surplus per recipient HH ($) ($p(\Omega_\tau)s/p(\Omega_b)$) | USDP value/value of cost-equiv tax cuts ($(s+EAG)/EAG$) | MVPF |
|---|---|---|---|---|---|---|
| Baseline | | | | | | |
| DI | 0.91 | 5,191 | 924 | 8,718 | 1.64 | 1.76 |
| M-DI | 0.95 | 3,557 | 488 | 9,858 | 1.73 | 1.85 |
| L-DI | 0.88 | 6,623 | 436 | 7,719 | 1.57 | 1.68 |
| M-NDI | 0.64 | 22,591 | -74 | -2,203 | 0.86 | 0.93 |
| Lower risk aversion ($\gamma = 1$, i.e., log utility) | | | | | | |
| DI | 0.41 | 5,191 | 310 | 2,927 | 1.22 | 1.30 |
| M-DI | 0.44 | 3,557 | 186 | 3,763 | 1.28 | 1.36 |
| L-DI | 0.39 | 6,623 | 124 | 2,195 | 1.16 | 1.24 |
| M-NDI | 0.28 | 22,591 | -208 | -6,180 | 0.62 | 0.66 |
| Higher risk aversion ($\gamma = 3$) | | | | | | |
| DI | 1.41 | 5,191 | 1,530 | 14,425 | 2.06 | 2.22 |
| M-DI | 1.43 | 3,557 | 772 | 15,587 | 2.15 | 2.32 |
| L-DI | 1.39 | 6,623 | 758 | 13,407 | 1.99 | 2.15 |
| M-NDI | 1.02 | 22,591 | 76 | 2,253 | 1.14 | 1.24 |
| State-dependent utility: Lower when more-severe | | | | | | |
| DI | 0.72 | 5,191 | 676 | 6,372 | 1.47 | 1.58 |
| M-DI | 0.49 | 3,557 | 215 | 4,349 | 1.32 | 1.42 |
| L-DI | 0.92 | 6,623 | 460 | 8,144 | 1.60 | 1.72 |
| M-NDI | 0.25 | 22,591 | -184 | -5,471 | 0.66 | 0.71 |
| State-dependent utility: Higher when more-severe | | | | | | |
| DI | 1.10 | 5,191 | 1,164 | 10,979 | 1.81 | 1.94 |
| M-DI | 1.39 | 3,557 | 751 | 15,162 | 2.12 | 2.27 |
| L-DI | 0.85 | 6,623 | 413 | 7,314 | 1.54 | 1.65 |
| M-NDI | 1.01 | 22,591 | 32 | 940 | 1.06 | 1.14 |
| Markup x 0.25 | | | | | | |
| DI | 0.23 | 5,191 | 79 | 748 | 1.06 | 1.13 |
| M-DI | 0.24 | 3,557 | 66 | 1,342 | 1.10 | 1.17 |
| L-DI | 0.22 | 6,623 | 13 | 228 | 1.02 | 1.09 |
| M-NDI | 0.16 | 22,591 | -252 | -7,488 | 0.54 | 0.57 |

Notes: Table presents statistics (given by the column names) associated with receiving disability benefits in each of several sets of states of the world (row names). "Baseline" is the baseline specification. "Lower risk aversion" uses $\gamma = 1$ (i.e., log utility). "Higher risk aversion" uses $\gamma = 3$. "State-dependent utility: Lower when more-severe" assumes that marginal utility is 25% lower in states with more-severe health conditions at a given level of consumption, the upper bound effect of bad health on marginal utility estimated by Finkelstein, Luttmer and Notowidigdo (2013). This affects the markup and surplus from transfers not only to more-severe states but to less-severe states as well, by changing marginal utility in some of the states in which the individual pays for benefits (more-severe non-recipient states). "State-dependent utility: Higher when more-severe" assumes that marginal utility is 25% higher in states with more-severe health conditions at a given level of consumption. "Markup $\times$ 0.25" sets all markups to 25% of their estimated baseline values. The sets of states are those in which the household receives USDP (DI), including when more-severe (M-DI) and less-severe (L-DI), and those in which the household does not receive USDP when more-severe (M-NDI). Markup is defined in equation (3) and calculated using PSID consumption excluding health care expenditures. Counterfactual no-benefit earnings are actual observed earnings for NDI states and are inferred from French and Song's (2014) estimates of the effects of disability benefits on earnings for DI states, as described in the text. Surplus $s$, defined in equation (6), is in units of government revenue per $\Omega_\tau$ state, so $p(\Omega_\tau)s$ is in units of government revenue per household and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient household. Value relative to that of cost-equivalent tax cuts is derived in footnote 27. Monetary amounts are in 2016 dollars per year. Sample sizes: 313 M-DI, 438 L-DI, 248 M-NDI.

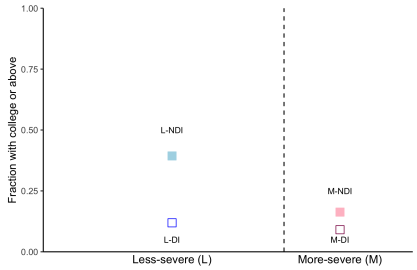
Table F5: Policy analysis

| | $E(M_\omega)$ | Surplus per household (\$) $(p(\Omega_\tau)s)$ | Surplus per recipient HH (\$) $(p(\Omega_\tau)s/p(\Omega_b))$ | $MVPF$ |
|---|---|---|---|---|
| Eliminate L-DI benefits | 0.88 | -292 | -5,164 | 1.42 |
| Decrease benefit levels | 0.91 | -61 | -571 | 1.43 |
| Decrease allowance rate | 0.88 | -29 | -5,164 | 1.42 |

Notes: Table presents mean markup, social surplus per household, social surplus per recipient, and MVPF associated with different policies. Each of the policies contracts USDP, so a negative surplus means that contracting USDP in that way reduces social surplus. The MVPF, the marginal value of public funds associated with each policy, can be viewed either as the ex ante cost to the individual per dollar of net savings to the government of *contracting* USDP along that dimension or, equivalently, as the ex ante value to the individual per dollar of net cost to the government of *expanding* USDP along that dimension, $MVPF(\Omega_b) = \frac{EAWTP(\Omega_b)}{EAG(\Omega_b)}$. "Eliminate L-DI benefits" is a hypothetical, infeasible policy that eliminates L-DI benefits (benefits to less-severe) at no administrative cost. "Decrease benefit levels" decreases benefit levels by 10%, from \$13,000 to \$11,700. (The scale matters for the surplus results.) "Decrease allowance rate" decreases the allowance rate such that USDP recipiency decreases by 10%, from about 11% to about 10%. We assume that this decrease in allowances affects "USDP applicants on the margin of program entry," those applicants whose award decisions depend on the examiner to which they are (quasi-randomly) assigned (as in Maestas, Mullen and Strand, 2013). We assume that mean marginal utility among those so denied equals that among less-severe recipients. We assume that each of these contractions of USDP produce cost savings to the government from reducing applications to USDP as well as through their direct effect on inframarginal recipients, with an overall fiscal externality savings of 34 cents per dollar of reduced transfers to inframarginal recipients. Surplus $s$ is in units of government revenue per $\Omega_\tau$ state per year, so $p(\Omega_\tau)s$ is in units of government revenue per household per year and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient per year. Monetary amounts are in 2016 dollars per year. See Section D for details.

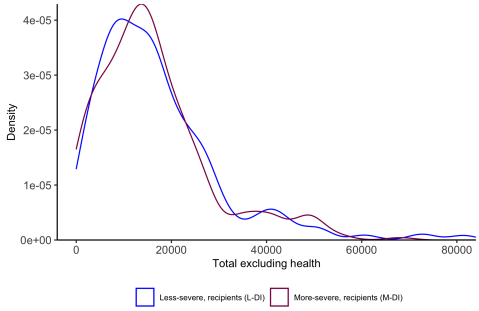
# G Supplemental Materials (not for publication): Additional Figures

Figure G1: PSID: College or above



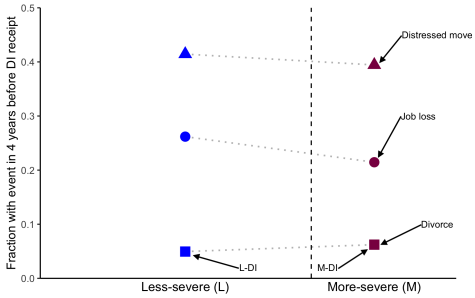

Notes: Figure presents rates of college degree or above in the 2017 PSID for each of the four severity-by-USDP receipt groups: less-severe non-DI-recipients (6,312 L-NDI), more-severe non-recipients (250 M-NDI), less-severe recipients (443 L-DI), and more-severe recipients (316 M-DI). "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe."

Figure G2: PSID: Consumption distribution at application



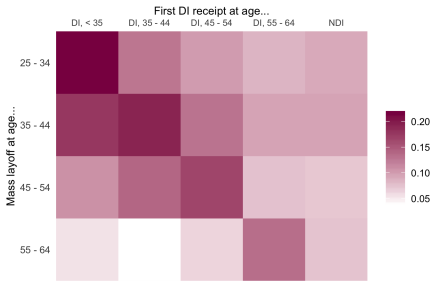

Notes: Figure presents the distribution of consumption (excluding health care expenditures) one year prior to the year of initial USDP receipt for each of the two severity-by-DI-receipt groups in the 2017 PSID: more-severe DI-recipients (213 M-DI), and less-severe recipients (286 L-DI). Household consumption excludes health care and is divided by the square root of household size. "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe."

Figure G3: PSID: Adverse events prior to USDP receipt



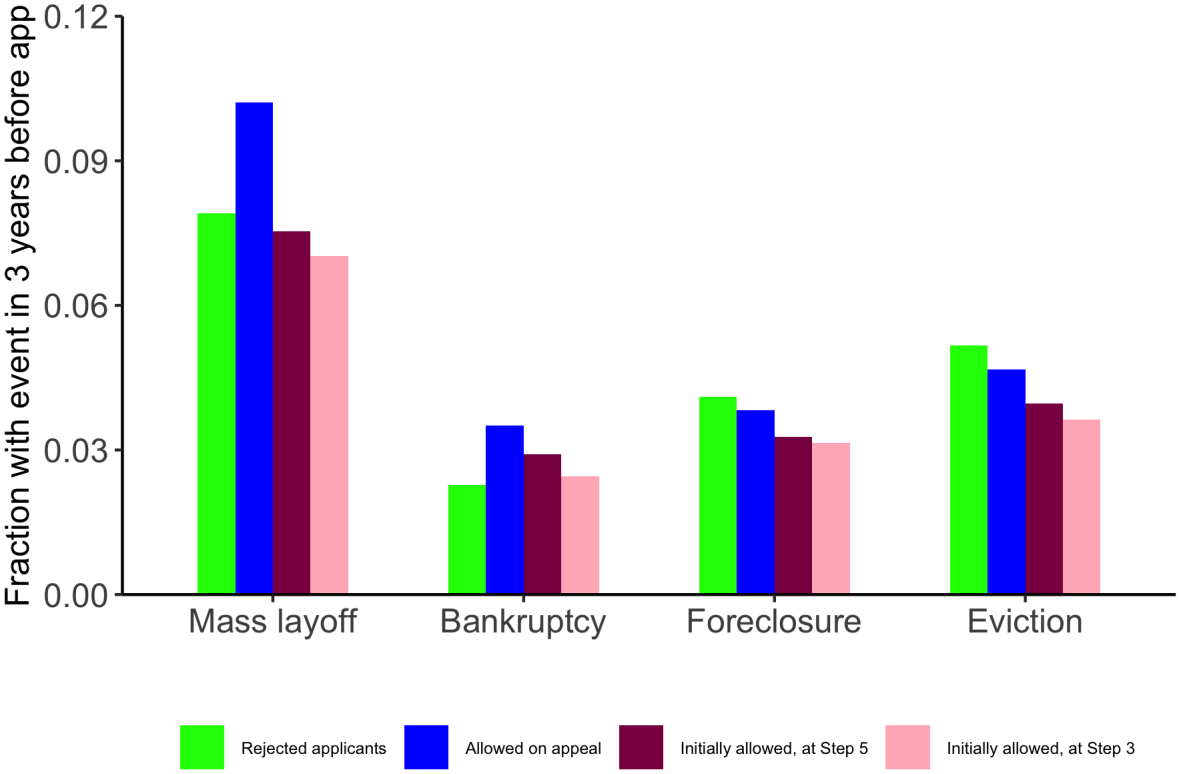

Notes: Figure presents rates of various adverse life events in the four years before receiving USDP for households receiving USDP in the 2017 PSID, by more-severe and less-severe. Job loss includes "involuntary" reasons for separation: strike/lockout, laid off/fired, or company going out of business or leaving town (exact PSID question in Appendix Table I2). The sample for job loss is households currently receiving USDP who can be observed in the four years before initial USDP receipt (175 M-DI, 240 L-DI). Distressed move includes external events like eviction, contraction of housing (less space/less rent), and other reasons potentially indicating distress, such as saving money (exact PSID question in Appendix Table I3). The sample for distressed move is households currently receiving USDP who can be observed in the four years before initial USDP receipt (168 M-DI, 221 L-DI). Divorce is defined as being married in a previous survey year but not in this survey year (exact PSID question in Appendix Table I2). The divorce sample is households currently receiving USDP who can be observed in the four years before initial USDP receipt (162 M-DI, 210 L-DI). "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe."

Figure G4: CWHS: Mass layoffs by USDP status



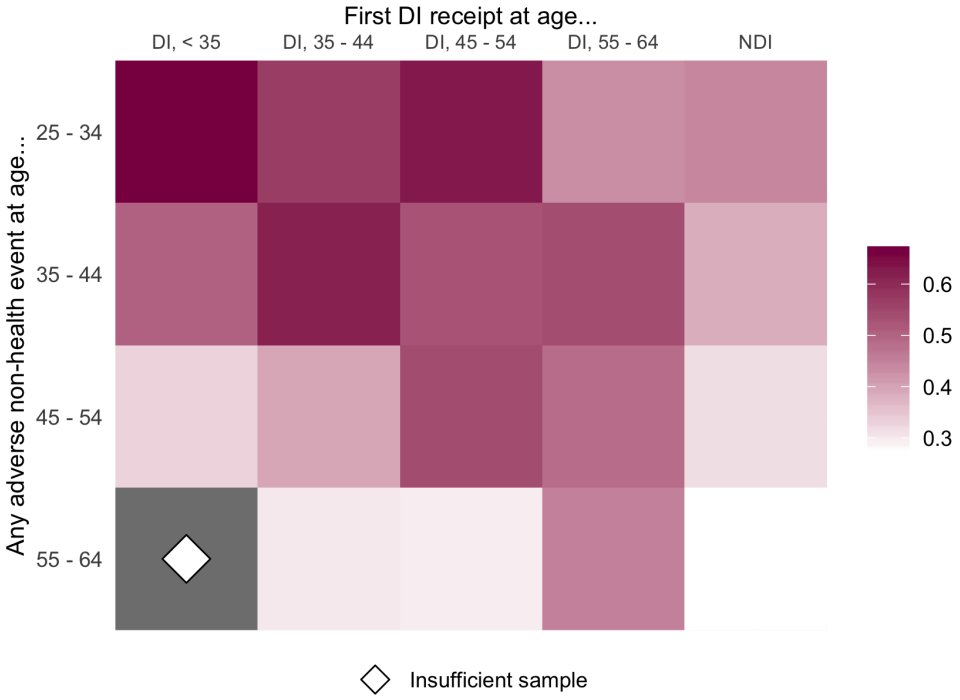

Notes: Figure represents rates of mass layoff at various ages for individuals in the CWHS, by USDP status, and age at USDP entry. Individuals are first categorized by their USDP status: whether they ever received USDP (DI) or never received USDP (NDI). The DI group is further categorized by age at USDP entry: <35, 35–44, 45–54, 55–64. We organize these groups along the horizontal axis and then plot the rate of experiencing any adverse event at the following ages along the vertical axis: 25–34, 35–44, 45–54, and 55–64. Darker shades indicate higher rates, as shown in the legend. The sample of each cell is individuals in DI/age category indicated in the horizontal axis that are observed in the CWHS during the age range indicated in the vertical axis.

Figure G5: Admin data: Financial events in 3 years before app



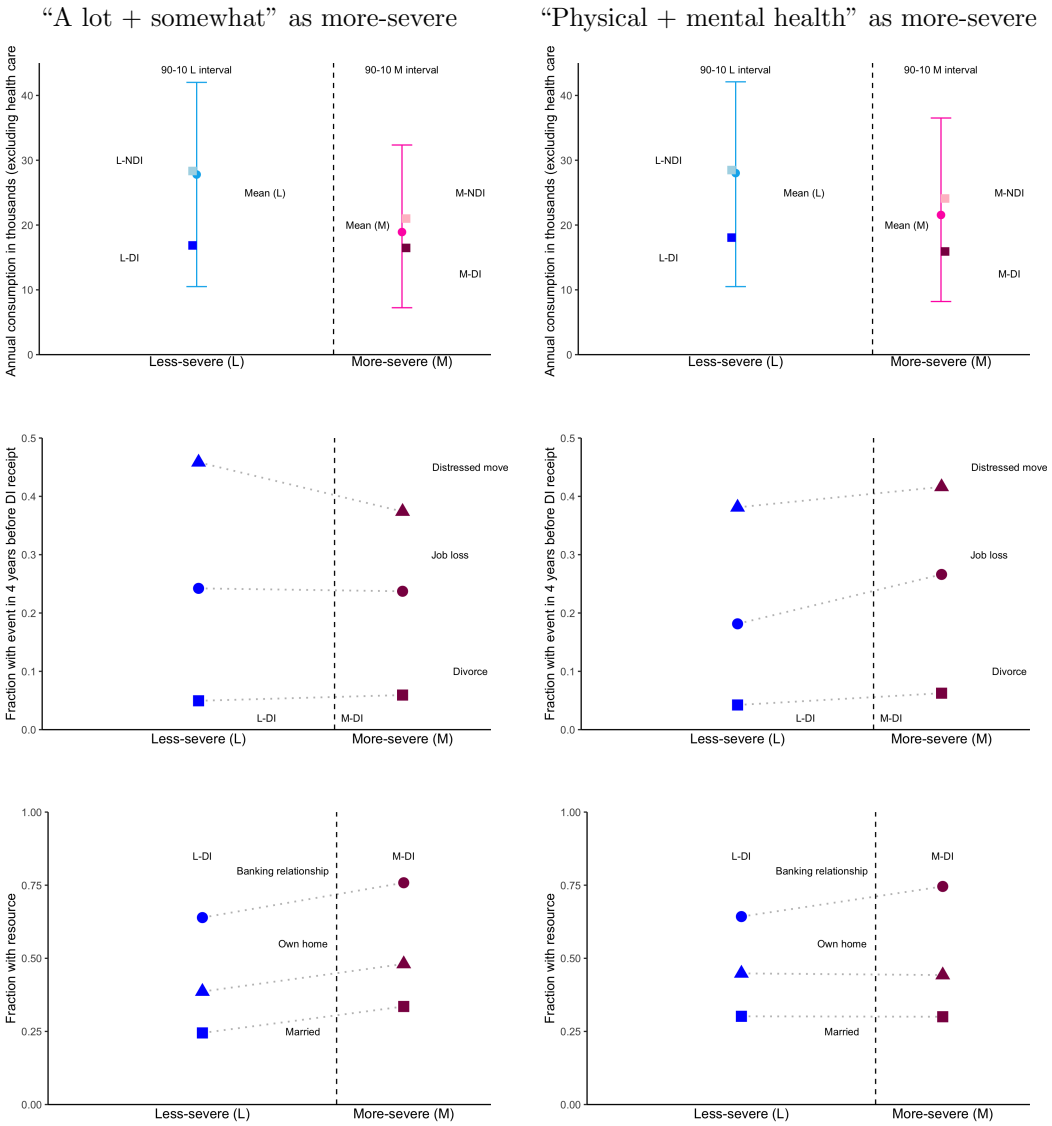

Notes: Figure presents rates of adverse financial events in the three years before receiving USDP (SSDI and SSI) for individuals who receive USDP benefits in SSA administrative data. "Initially allowed, at Step 3" indicates that recipient was allowed at initial state DDS level at the earliest possible step (meeting medical listings). "Initially allowed, at Step 5" indicates that the recipient was allowed at initial state DDS level, but not until Step 5 (vocational grid considering age, education, experience). "Allowed upon appeal" indicates that the recipient was rejected at the initial state DDS level and allowed upon appeal. The mass layoff sample is USDP applicants in the 831 records between 1990 and 2016 that appear in the Continuous Work History (CWHS), a 10% sample of individuals in the U.S. (640,000 Step 3; 770,000 Step 5; 510,000 appeal; 500,000 rejected). We identify EINs in the CWHS that experience a drop in employees of at least 30 percent from a base of at least 150 employees and consider an applicant to have experienced a mass layoff in the three years prior to application if they experienced a separation from a mass-layoff EIN in the same year as the mass layoff event. "Bankruptcy," "foreclosure," and "eviction" indicate experiencing these events in the three years prior to USDP application. We link USDP applicants in the 831 files to nationwide financial records using the methods described in Deshpande, Gross and Su (2019). The bankruptcy sample is USDP applicants in the 831 files between 1995 and 2009 (4.7M Step 3; 3.7M Step 5; 4.9M appeal; 12M rejected). The foreclosure sample is approved USDP applicants in the 831 files between 2005 and 2014 who are homeowners (630,000 Step 3; 810,000 Step 5; 700,000 appeal; 1.3M rejected). The eviction sample is approved USDP applicants in the 831 files between 2005 and 2014 who are not homeowners (710,000 Step 3; 750,000 Step 5; 590,000 appeal; 2.4M rejected).

Figure G6: PSID: Any adverse life event by USDP status



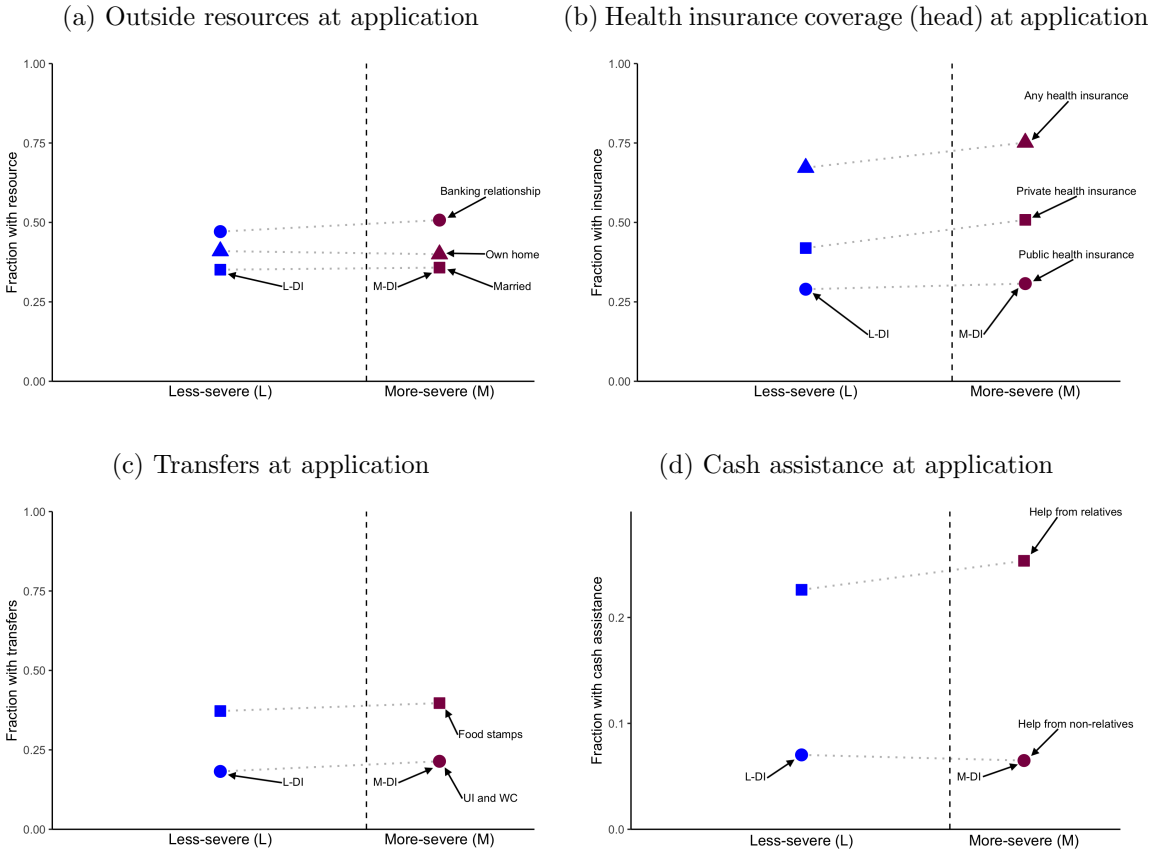

Notes: Figure presents rates of experiencing an "adverse life event"—head or spouse job loss, involuntary move, or divorce—at various ages for households in the PSID, by USDP status, and age at USDP entry. Households are first categorized by their USDP status: whether they ever received USDP (DI) or never received USDP (NDI). The DI group is further categorized by age at USDP entry: <35, 35–44, 45–54, 55–64. We organize these groups along the horizontal axis and then plot the rate of experiencing any adverse event at the following ages along the vertical axis: 25–34, 35–44, 45–54, and 55–64. Darker shades indicate higher rates, as shown in the legend. Gray squares with a diamond indicate that the sample size is less than 50. The sample of each cell is households in DI/age category indicated in the horizontal axis that are observed in the PSID during the age range indicated in the vertical axis.

Figure G7: PSID: Positive results using alternative definitions of more-severe




Notes: These figures replicate Figures 2a, G3, and 2c using two alternative definitions of more-severe. In the left column, we use a broader definition of more-severe that includes the "somewhat" response (in addition to the baseline "a lot" response) to the question of how much a health condition limits work. In the right column, we use a definition that incorporates mental health on top of the baseline severity measure—in particular, having depression, psychiatric issues, or "loss of memory or mental ability." The figures in the first row represent the 90-10 percentile interval (and average) of consumption for more-severe and less-severe households in the 2017 PSID, as in Figure 2a. The figures in the second row represent rates of various adverse life events in the four years before receiving USDP for households receiving USDP in the 2017 PSID, by more-severe and less-severe, as in Figure G3. The figures in the third row present rates of marriage, banking relationship, and homeownership for households receiving USDP in the 2017 PSID, by more-severe and less-severe, as in Figure 2c. See notes in original figures for more details on each measure.

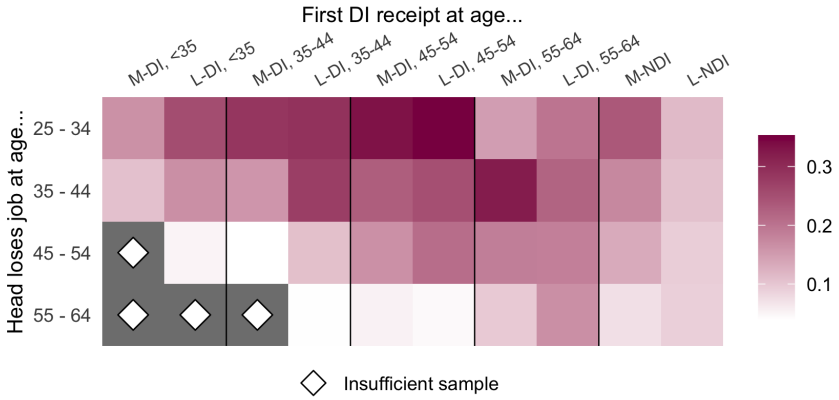
Figure G8: Resources available at application



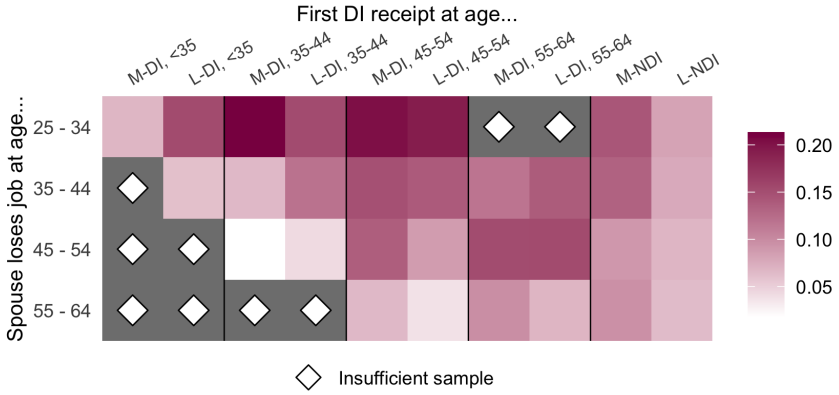

Notes: Figures present rates of outside resources one year prior to the year of initial USDP receipt for each of two severity-by-USDP-receipt groups in the 2017 PSID: more-severe USDP-recipients (243 M-DI), and less-severe recipients (320 L-DI). Figure G8a presents rates of marriage, banking relationship, and homeownership, Figure G8b presents rates of any health insurance, private health insurance, and public health insurance, Figure G8c presents rates of SNAP (food stamps) and unemployment insurance (UI) and workers' compensation (WC), and Figure G8d presents rates of cash assistance from relatives and non-relatives. "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe." Appendix Tables I4 and I5 present the exact PSID questions for each measure.

Figure G9: PSID: Job loss by USDP and health status

(a) Head job loss by USDP and health status



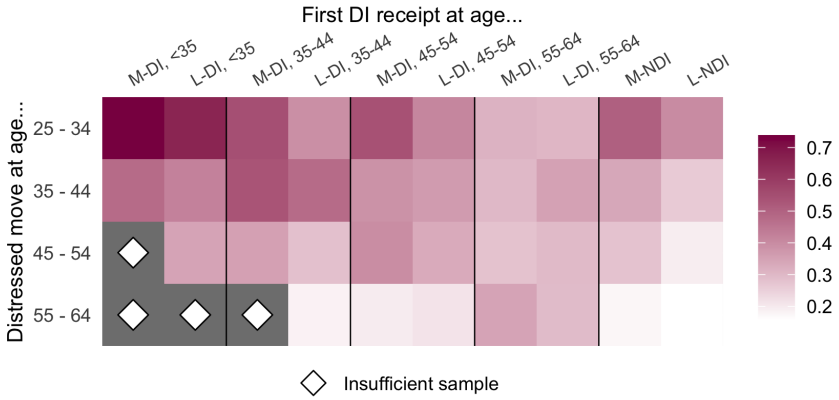

(b) Spouse job loss by USDP and health status



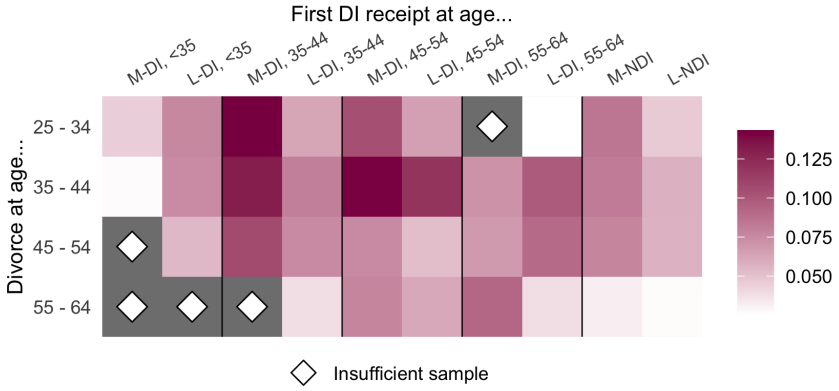

Notes: Figures present rates of head and spouse job loss at various ages for households in the PSID, by USDP status, health status, and age at USDP entry. Households are first categorized by their USDP status: whether they ever received USDP (DI) or never received USDP (NDI). DI are further classified by their health status at USDP entry (more-severe M-DI, or less-severe L-DI), and NDI by whether they ever had a more-severe health condition (ever more-severe M-NDI, never more-severe L-NDI). The M-DI and L-DI groups are further categorized by their age at USDP entry: <35, 35–44, 45–54, 55–64. We organize these groups along the horizontal axis and then plot the rate of experiencing any adverse event at the following ages along the vertical axis: 25–34, 35–44, 45–54, and 55–64. Darker shades indicate higher rates, as shown in the legend. Gray squares with a diamond indicate that the sample size is less than 50. The sample of each cell is households in DI/age category indicated in the horizontal axis that are observed in the PSID during the age range indicated in the vertical axis. Job loss includes "involuntary" reasons for separation: strike/lockout, laid off/fired, or company going out of business or leaving town (exact PSID question in Appendix Table I2). "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe."

Figure G10: PSID: Distressed move and divorce by USDP and health status

(a) Distressed move by USDP and health status




(b) Divorce by USDP and health status




Notes: Figures present rates of experiencing distressed move and divorce at various ages for households in the PSID, by USDP status, health status, and age at USDP entry. Households are first categorized by their USDP status: whether they ever received USDP (DI) or never received USDP (NDI). DI are further classified by their health status at USDP entry (more-severe M-DI, or less-severe L-DI), and NDI by whether they ever had a more-severe health condition (ever more-severe M-NDI, never more-severe L-NDI). The M-DI and L-DI groups are further categorized by their age at USDP entry: <35, 35–44, 45–54, 55–64. We organize these groups along the horizontal axis and then plot the rate of experiencing any adverse event at the following ages along the vertical axis: 25–34, 35–44, 45–54, and 55–64. Darker shades indicate higher rates, as shown in the legend. Gray squares with a diamond indicate that the sample size is less than 50. The sample of each cell is households in DI/age category indicated in the horizontal axis that are observed in the PSID during the age range indicated in the vertical axis. Distressed move includes external events like eviction, contraction of housing (less space/less rent), and other reasons potentially indicating distress, such as saving money (exact PSID question in Appendix Table I3). Divorce is defined as being married in a previous survey year but not in this survey year (exact PSID question in Appendix Table I2). "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe."

# H Supplemental Materials (not for publication): Additional Tables

Table H1: Value of disability benefits in different states: Robustness to alternative severity

| | Mean markup ($E(M_\omega)$) | Counterfactual no-benefit earnings ($) | Surplus per household ($) ($p(\Omega_\tau)s$) | Surplus per recipient HH ($) ($p(\Omega_\tau)s/p(\Omega_b)$) | USDP value/value of cost-equiv tax cuts ($(s+EAG)/EAG$) | MVPF |
|---|---|---|---|---|---|---|
| Severity: A lot (baseline) | | | | | | |
| DI | 0.91 | 5,191 | 924 | 8,718 | 1.64 | 1.76 |
| M-DI | 0.95 | 3,557 | 488 | 9,858 | 1.73 | 1.85 |
| L-DI | 0.88 | 6,623 | 436 | 7,719 | 1.57 | 1.68 |
| M-NDI | 0.64 | 22,591 | -74 | -2,203 | 0.86 | 0.93 |
| Severity: A lot + Somewhat | | | | | | |
| DI | 0.91 | 5,191 | 924 | 8,718 | 1.64 | 1.76 |
| M-DI | 0.92 | 3,942 | 614 | 9,362 | 1.69 | 1.82 |
| L-DI | 0.91 | 7,213 | 311 | 7,677 | 1.56 | 1.68 |
| M-NDI | 0.45 | 30,558 | -495 | -6,408 | 0.63 | 0.69 |
| Severity: A lot + Mental health | | | | | | |
| DI | 0.91 | 5,191 | 924 | 8,718 | 1.64 | 1.76 |
| M-DI | 0.99 | 4,385 | 714 | 9,961 | 1.73 | 1.87 |
| L-DI | 0.76 | 6,875 | 210 | 6,121 | 1.45 | 1.55 |
| M-NDI | 0.29 | 48,866 | -2,059 | -13,012 | 0.37 | 0.41 |

Notes: Table presents statistics (given by the column names) associated with receiving disability benefits in each of several sets of states of the world (row names). "A lot" is the baseline severity measure: a household is classified as "more-severe" if it reports that a health condition limits "a lot" the amount of work that the head or spouse can do or that the head or spouse "can do nothing" as a result of that condition. "A lot + Somewhat" is a broader definition that includes households that report that a health condition limits "somewhat" the amount of work that the head or spouse can do. "A lot + Mental health" is the union of the baseline severity definition and households that report that a health professional ever told the head or spouse that they have depression, psychiatric issues, or "loss of memory or mental ability." The sets of states are those in which the household receives USDP (DI), including when more-severe (M-DI) and less-severe (L-DI), and those in which the household does not receive USDP when more-severe (M-NDI). Markup is defined in equation (3) and calculated using PSID consumption excluding health care expenditures. Counterfactual no-benefit earnings are actual observed earnings for NDI states and are inferred from French and Song's (2014) estimates of the effects of disability benefits on earnings for DI states, as described in the text. Surplus $s$, defined in equation (6), is in units of government revenue per $\Omega_\tau$ state, so $p(\Omega_\tau)s$ is in units of government revenue per household and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient household. Value relative to that of cost-equivalent tax cuts is derived in footnote 27. Monetary amounts are in 2016 dollars per year. Sample sizes for M-DI, L-DI, M-NDI: 313, 438, 248 for "A lot (baseline);" 424, 327, 519 for "A lot + Somewhat;" 458, 293, 1077 for "A lot + Mental health."

Table H2: Value of disability benefits in different states: Robustness to different strategies for isolating risk from heterogeneity

| | Mean markup ($E(M_\omega)$) | Counterfactual no-benefit earnings ($) | Surplus per household ($) ($p(\Omega_\tau)s$) | Surplus per recipient HH ($) ($p(\Omega_\tau)s/p(\Omega_b)$) | USDP value/value of cost-equiv tax cuts ($(s+EAG)/EAG$) | MVPF |
|---|---|---|---|---|---|---|
| Risk types: education (baseline) | | | | | | |
| DI | 0.91 | 5,191 | 924 | 8,718 | 1.64 | 1.76 |
| M-DI | 0.95 | 3,557 | 488 | 9,858 | 1.73 | 1.85 |
| L-DI | 0.88 | 6,623 | 436 | 7,719 | 1.57 | 1.68 |
| M-NDI | 0.64 | 22,591 | -74 | -2,203 | 0.86 | 0.93 |
| Risk types: earnings | | | | | | |
| DI | 0.65 | 5,587 | 684 | 6,091 | 1.45 | 1.51 |
| M-DI | 0.85 | 3,047 | 440 | 9,258 | 1.68 | 1.77 |
| L-DI | 0.51 | 7,445 | 245 | 3,774 | 1.28 | 1.33 |
| M-NDI | 0.58 | 27,812 | -173 | -4,873 | 0.71 | 0.75 |
| Estimate markups with controls | | | | | | |
| DI | 0.33 | 5,191 | 117 | 1,106 | 1.08 | 1.21 |
| M-DI | 0.34 | 3,557 | 91 | 1,846 | 1.14 | 1.27 |
| L-DI | 0.32 | 6,623 | 26 | 460 | 1.03 | 1.15 |
| M-NDI | 0.33 | 22,591 | -210 | -6,279 | 0.61 | 0.68 |

Notes: Table presents statistics (given by the column names) associated with receiving disability benefits in each of several sets of states of the world (row names). The risk type definition determines the ex ante risk facing different individuals and so the insurance value of transfers into different states. "Risk type: education," the baseline specification, defines four risk types: one for each of the following education categories: high school dropout, high school graduate, some college, and college or more. "Risk type: earnings" defines ten risk types: one for each of the deciles of the distribution of age-35 earnings among individuals not receiving disability benefits at that time. "Estimate markups with controls" uses just a single risk type but estimates the markups on transfers to different states with regressions that include controls for education, age, whether the head of household has a spouse, and dummies for household income quartiles. This specification aims to be conservative in terms of understating the value of transfers to USDP-receiving states, as the controls for (current) household income likely "over-control" for heterogeneity across ex ante risk types by incorporating much of the ex post realization of risk. The sets of states are those in which the household receives USDP (DI), including when more-severe (M-DI) and less-severe (L-DI), and those in which the household does not receive USDP when more-severe (M-NDI). Markup is defined in equation (3) and calculated using PSID consumption excluding health care expenditures. Counterfactual no-benefit earnings are actual observed earnings for NDI states and are inferred from French and Song's (2014) estimates of the effects of disability benefits on earnings for DI states, as described in the text. Surplus $s$, defined in equation (6), is in units of government revenue per $\Omega_\tau$ state, so $p(\Omega_\tau)s$ is in units of government revenue per household and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient household. Value relative to that of cost-equivalent tax cuts is derived in footnote 27. Monetary amounts are in 2016 dollars per year. Counterfactual earnings differ somewhat in the "Risk types: earnings" specification because it excludes households with missing values of age-35 earnings. Sample sizes for M-DI, L-DI, and M-NDI: 313, 438, 248 for "Risk types: education (baseline)" and "Estimate markups with controls;" 181, 276, 148 for "Risk types: earnings."

# I Supplemental Materials (not for publication): PSID definitions

Table I1: Severity questions

| | |
|---|---|
| Severity (PSID) | |
| (1) | Do you have any physical or nervous condition that limits the type of work or the amount of work you can do?<br>- Yes<br>- No<br>- Can do nothing |
| (2) | For work you can do, how much does it limit the amount of work you can do – a lot, somewhat, or just a little?<br>- A lot<br>- Somewhat<br>- Just a little<br>- Not at all |
| More-severe | if "Yes" in (1) and "A lot" in (2), or "Can do nothing" in (1) |
| Less-severe | otherwise |
| Severity (SIPP) | |
| (1) | Does ... have a physical, mental, or other health condition that limits the kind or amount of work ... can do at a job or business?<br>- Yes<br>- No |
| (2) | Does ... health or condition prevent ... from working at a job or business?<br>- Yes<br>- No |
| More-severe | if "Yes" in both (1) and (2) |
| Less-severe | otherwise |

Table I2: Job loss and divorce/marriage questions

| | |
|---|---|
| Job loss | Why did you stop working for (Name of employer)? – Did the company go out of business, were you laid off, did you quit, or what? – Most recent main job |
| ✓ | 1. Company folded/changed hands/moved out of town; employer died/went out of business |
| ✓ | 2. Strike; lockout |
| ✓ | 3. Laid off; fired |
| | 4. Quit; resigned; retired; pregnant; needed more money; just wanted a change |
| | 5. Other; transfer; any mention of armed services |
| | 6. Job was completed; seasonal work; was a temporary job |
| Divorce/marriage | Are you (Head) married, widowed, separated, or have you never been married? |
| | 1. Married |
| | 2. Never married |
| | 3. Widowed |
| | 4. Divorced, annulled |
| | 5. Separated |
| Divorced | if married in the previous observation but not married in this observation |

Notes: In our measure of job loss, we consider those reasons indicated by the check marks.

Table I3: Distressed move question

| | | |
|---|---|---|
| Moving | Why did you move? | |
| | 1. | Purposive productive reasons: to take another job; transfer: stopped going to school |
| | 2. | To get nearer to work |
| | 3. | Purposive consumptive reasons – expansion of housing: more space; more rent; better place |
| ✓ | 4. | Purposive consumptive reasons – contraction of housing: less space; less rent |
| | 5. | Purposive consumptive – other house-related: get own home/place; got married; physical conditions of the previous housing unit |
| | 6. | Purposive consumptive – neighborhood-related: better neighborhood; go to school; to be closer to friends and/or relatives |
| ✓ | 7. | Response to outside events (involuntary reasons): HU coming down; being evicted; armed services, etc.; health reasons; divorce; retiring because of health |
| ✓ | 8. | Ambiguous, mixed, or other reasons, including reasons such as to save money, all my old neighbors moved away, retiring |
| | 9. | Homeless |

Notes: In our measure of distressed moves, we consider those reasons indicated by the check marks.

Table I4: Resource, cash assistance, and other transfer questions

| | |
|---|---|
| Home own | Do you (or anyone else in your family living there) own the (home/apartment), pay rent, or what?<br>1. Owns or is buying home, either fully or jointly; mobile home owners who rent lots are included here<br>2. Pays rent<br>3. Neither owns nor rents |
| Banking relationship | If you added up all such accounts (for all of your family living there) about how much would they amount to right now? |
| Help from relatives | Did you receive any help from relatives? |
| Help from non-relatives | Did anyone (else) not living with you help you and your family out by giving you money? |
| Food stamp | Did you (or anyone else in your family) receive food stamp benefits (that is, either food stamps or a food stamp benefit card,) at any time last year? |
| UI | Did you receive any income from unemployment compensation? |
| Workers compensation | Did you receive any income from workers compensation? |

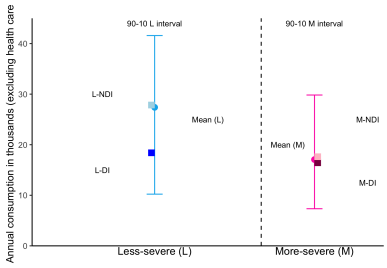
Table I5: Health insurance question

| Insurance | | What type of health insurance plan or health insurance coverage did you have in 2015 and 2016? If you had more than one type of plan or coverage please tell me about each of them. |
|---|---|---|
| Private | Public | |
| ✓ | | 1. Employer provided health insurance |
| ✓ | | 2. Private health insurance purchased directly |
| | ✓ | 3. Medicare |
| | ✓ | 4. Medi-Gap/Supplemental |
| | ✓ | 5. Medicaid/Medical Assistance/[STATE PROGRAM] |
| | ✓ | 6. Military health care/TRICARE (Active) |
| | ✓ | 7. CHAMPUS/TRICARE/CHAMP-VA (Dependents, Veterans) |
| | ✓ | 8. Indian Health Insurance |
| | ✓ | 9. Other state-sponsored plan (not Medicaid) |
| ✓ | | 10. Other government program |
| ✓ | | 11. Other |

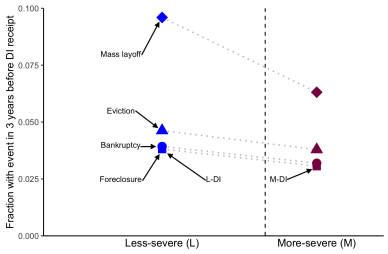
# J Supplemental Materials (not for publication): SSDI-Only Figures/Tables

Figure J1: Fact 2: L-DI similar to or worse off than M-DI on non-health measures

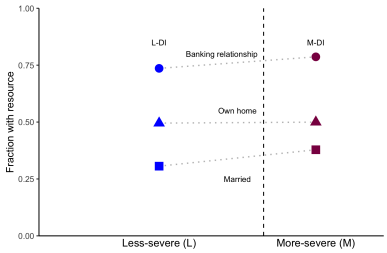
(a) PSID: Consumption



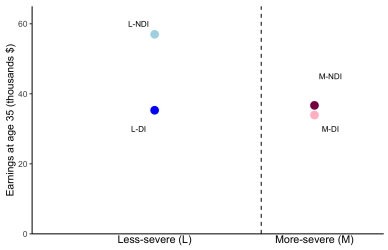

(b) Admin data: Prior to SSDI receipt




(c) PSID: Outside resources



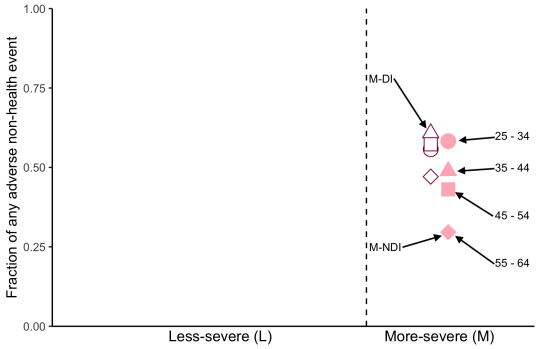

(d) PSID: Earnings at age 35



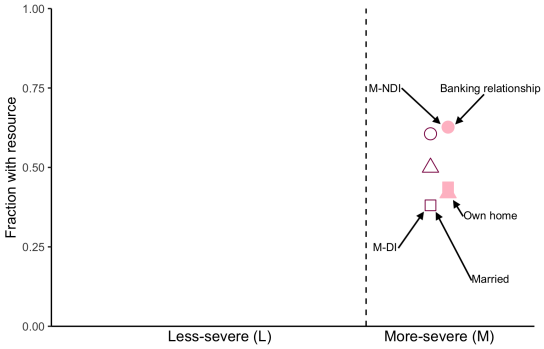

Notes: Figures J1a, J1c, and J1d present data from the PSID. Figure J1a presents statistics on consumption in the 2017 PSID for less-severe non-SSDI-recipients (6,442 L-NDI), more-severe non-recipients (316 M-NDI), less-severe recipients (312 L-DI), and more-severe recipients (250 M-DI). Household consumption excludes health care and is divided by the square root of household size. Figure J1c presents rates of marriage, banking relationship, and homeownership for M-DI and L-DI (exact PSID questions in Appendix Table I4). Figure 2d presents average earnings at head age 35 (in 2016 dollars) for households that did not receive SSDI benefits before age 36 and can be tracked back to age 35: 1,967 L-NDI, 108 M-NDI, 93 M-DI, and 75 L-DI. In all three PSID figures, "more-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do (vs. "somewhat," "just a little," or "not at all" [or no condition]). Figure J1b presents rates of adverse financial events in the three years before receiving SSDI for individuals who receive SSDI benefits in SSA administrative data. The mass layoff sample (1.1M L-DI, 800,000 M-DI) is SSDI recipients in the 831 records between 1990 and 2016 that appear in the Continuous Work History (CWHS). The bankruptcy sample is SSDI recipients in the 831 files between 1995 and 2009 (3.7M L-DI, 5.3M M-DI). The foreclosure sample is approved SSDI applicants in the 831 files between 2005 and 2014 who are homeowners (640,000 L-DI, 1.3M M-DI). The eviction sample is approved SSDI applicants in the 831 files between 2005 and 2014 who are not homeowners (430,000 L-DI, 880,000 M-DI). "More-severe" in Figure 2b indicates recipients allowed at the initial state DDS level, and "less-severe" indicates recipients allowed upon appeal.

Figure J2: Fact 3: M-NDI better off than M-DI on non-health measures

(a) PSID: Any adverse life event by SSDI status and age

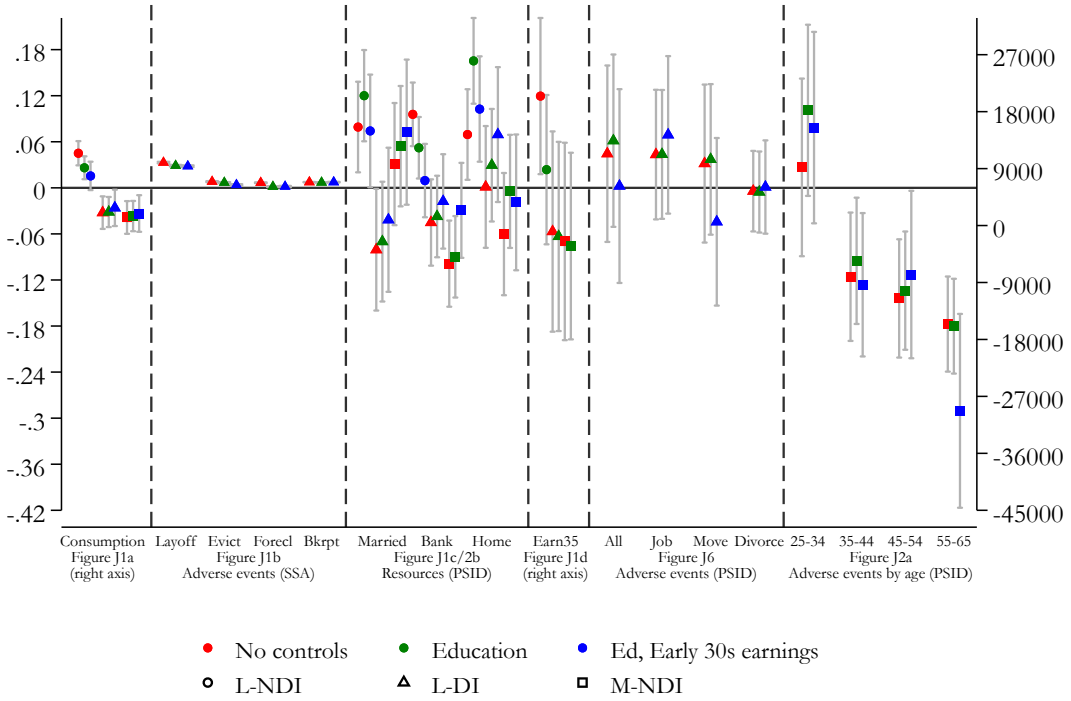



(b) PSID: Resources at more-severe onset



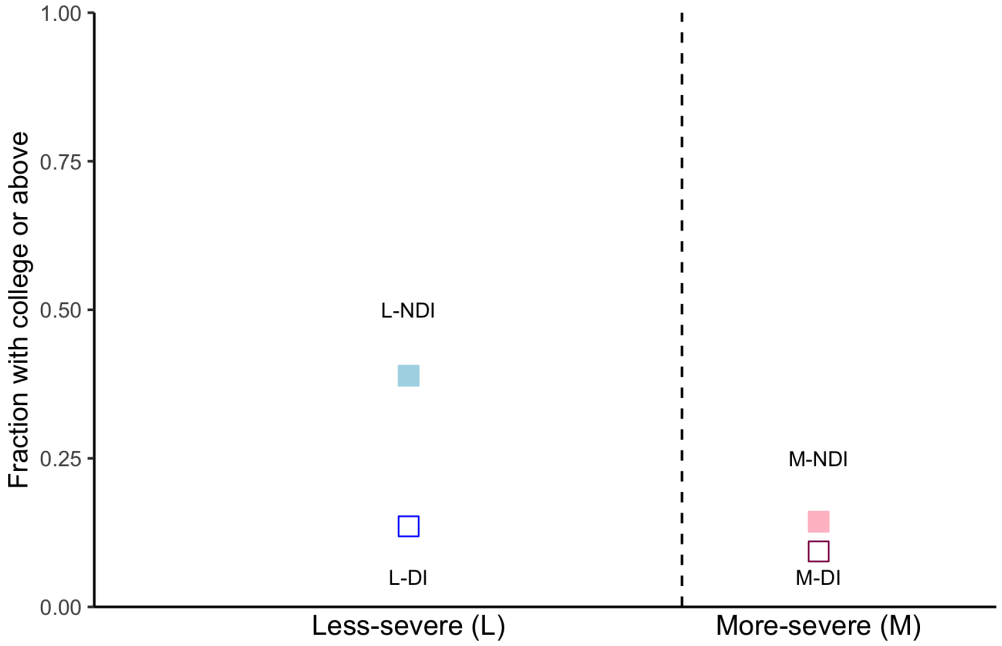

Notes: Figure J2a presents rates of experiencing an "adverse life event"—head or spouse job loss, involuntary move, or divorce—at various ages for households in the PSID, by SSDI status, health status, and age at SSDI entry. Households are first categorized by their SSDI status: whether they ever received SSDI (DI) or never received SSDI (NDI). DI are further classified by their health status at SSDI entry and their age at entry, and NDI are further classified by whether they ever had a more-severe health condition. The figure plots the share experiencing an adverse life event in a specific age range (e.g., 25–34 and 35–44) among M-NDI and among M-DI who also entered SSDI at that age. The sample sizes for age groups 25–34, 35–44, 45–54, and 55–64 are 1039, 1407, 1254, and 1039 for M-NDI; and 75, 172, 223, and 339 for M-DI. Job loss includes "involuntary" reasons for separation: strike/lockout, laid off/fired, or company going out of business or leaving town (see Appendix Table I2). Distressed move includes external events like eviction, contraction of housing (less space/less rent), and other reasons potentially indicating distress, such as saving money (see Appendix Table I3). Divorce is defined as being married in a previous survey year but not in this survey year (see Appendix Table I2). Figure J2b presents rates of marriage, banking relationship, and homeownership one year prior to the year of more-severe onset for M-NDI (204) and M-DI (152) (exact PSID questions in Appendix Table I4). "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do (as opposed to "somewhat," "just a little," or "not at all" [or no health condition]). Markers for the two groups are offset to facilitate easier reading of the graph, not as a representation of relative severity.

Figure J3: Robustness to Ability-Group Controls



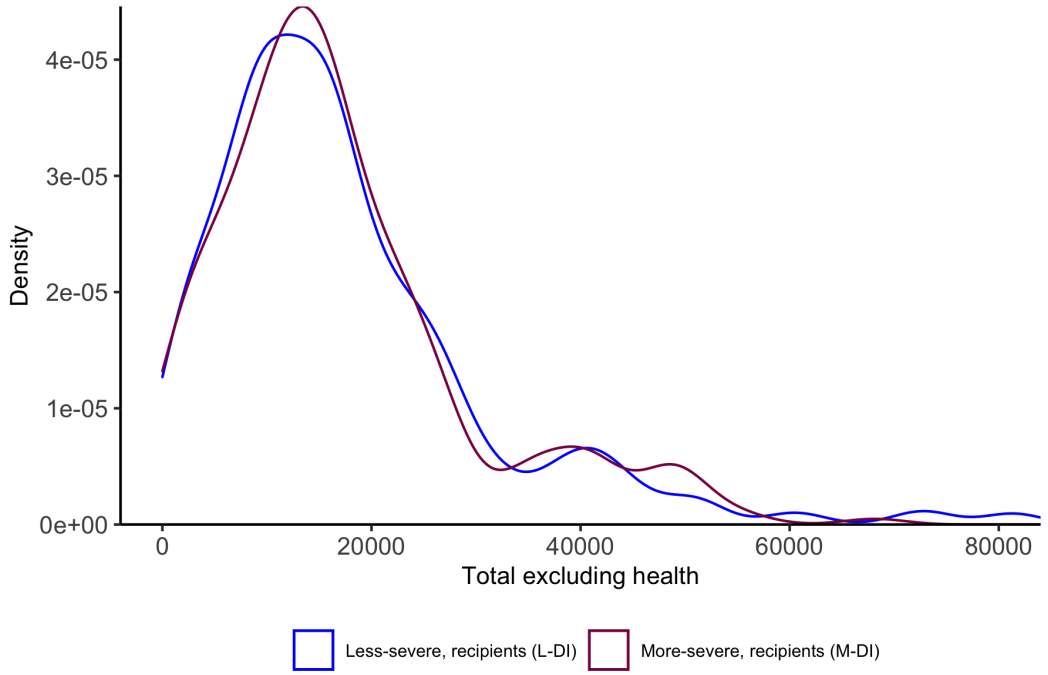

Notes: Figure presents all outcomes from Figures J1, J2, and J6 with controls for ability group. The figure plots the coefficients on indicators for each SSDI-by-severity group (L-NDI, L-DI, and M-NDI; M-DI is the excluded group) from a regression of the outcome on those indicators and ability controls. The three groups are shown using different shapes. We use three specifications for ability controls: no controls (which reproduces the coefficients from Figures J1 and J2), education controls (less than high school, high school graduate, some college, and college plus), and education plus average annual earnings (deciles) between ages 31-35. The three specifications are shown using different colors. The "Consumption" and "Earn35" outcomes use the right axis, and the other outcomes use the left axis. The variable "All" for Figure J6 indicates experiencing at least one of the adverse events (job loss, involuntary move, or divorce).

Figure J4: PSID: College or above



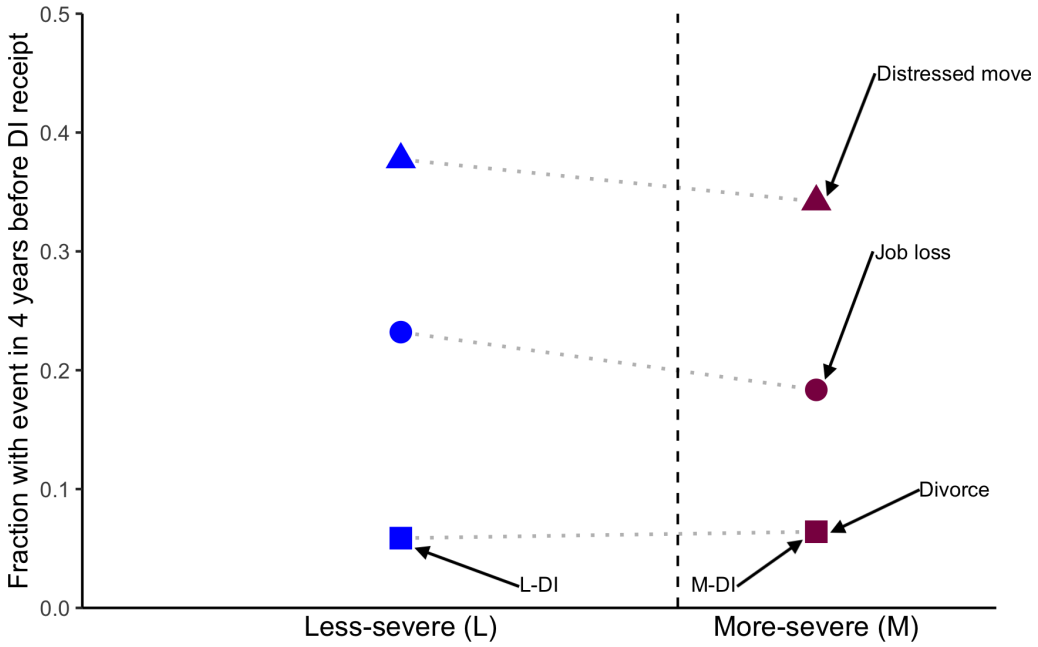

Notes: Figure presents rates of college degree or above in the 2017 PSID for each of the four severity-by-SSDI receipt groups: less-severe non-SSDI-recipients (6,442 L-NDI), more-severe non-recipients (316 M-NDI), less-severe recipients (312 L-DI), and more-severe recipients (250 M-DI). "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe."

Figure J5: PSID: Consumption distribution at application



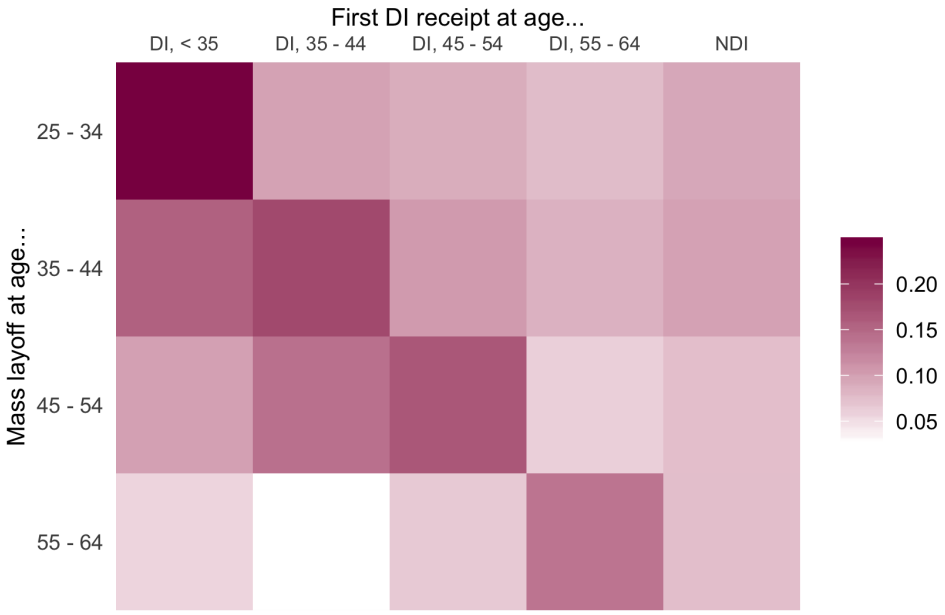

Notes: Figure presents the distribution of consumption (excluding health care expenditures) one year prior to the year of initial SSDI receipt for each of the two severity-by-SSDI-receipt groups in the 2017 PSID: more-severe SSDI-recipients (199 M-DI), and less-severe recipients (242 L-DI). Household consumption excludes health care and is divided by the square root of household size. "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe."

Figure J6: PSID: Adverse events prior to SSDI receipt by health status



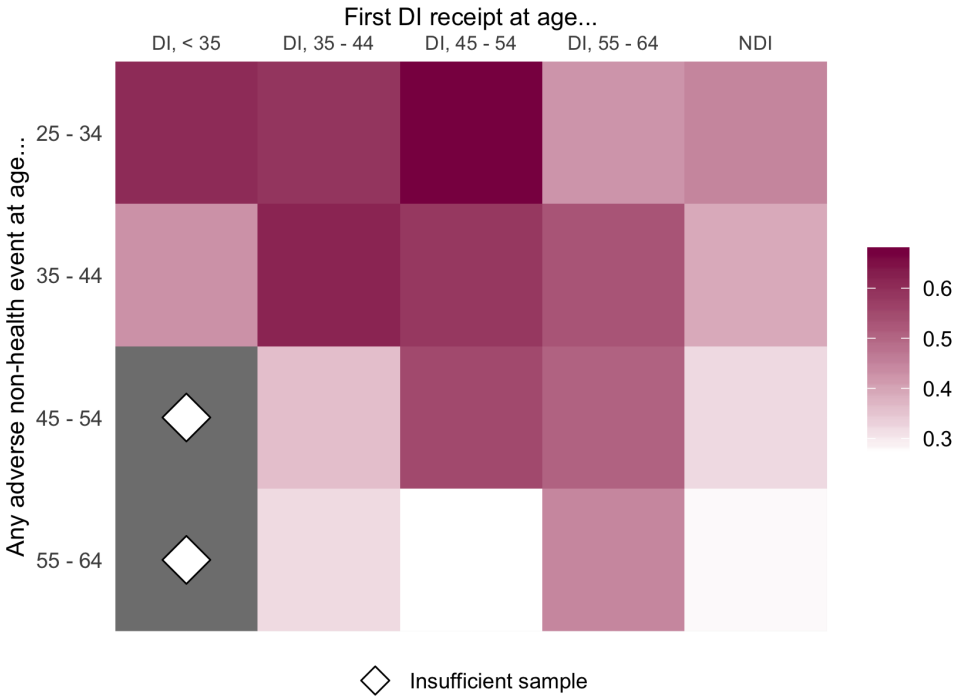

Notes: Figure presents rates of various adverse life events in the four years before receiving SSDI for households receiving SSDI in the 2017 PSID, by more-severe and less-severe. Job loss includes "involuntary" reasons for separation: strike/lockout, laid off/fired, or company going out of business or leaving town (see Appendix Table I2). The sample for job loss is households currently receiving SSDI who can be observed in the four years before initial SSDI receipt (161 more-severe, 201 less-severe). Distressed move includes external events like eviction, contraction of housing (less space/less rent), and other reasons potentially indicating distress, such as saving money (see Appendix Table I3). The sample for distressed move is households currently receiving SSDI who can be observed in the four years before initial SSDI receipt (186 more-severe, 156 less-severe). Divorce is defined as being married in a previous survey year but not in this survey year. The divorce sample is households currently receiving SSDI who can be observed in the four years before initial SSDI receipt (175 more-severe, 154 less-severe).

Figure J7: CWHS: Mass layoffs by SSDI status



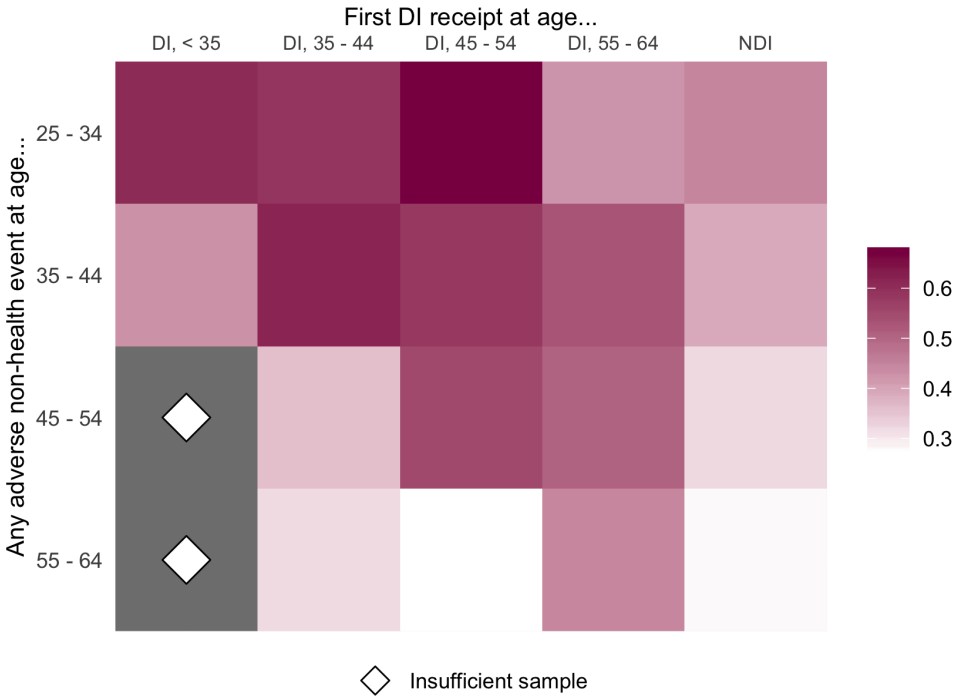

Notes: Figure represents rates of mass layoff at various ages for individuals in the CWHS, by SSDI status, and age at USP entry. Individuals are first categorized by their SSDI status: whether they ever received SSDI (DI) or never received SSDI (NDI). The SSDI group is further categorized by age at SSDI entry: <35, 35–44, 45–54, 55–64. We organize these groups along the horizontal axis and then plot the rate of experiencing any adverse event at the following ages along the vertical axis: 25–34, 35–44, 45–54, and 55–64. Darker shades indicate higher rates, as shown in the legend. The sample of each cell is individuals in SSDI/age category indicated in the horizontal axis that are observed in the CWHS during the age range indicated in the vertical axis.

Figure J8: PSID: Any adverse life event by SSDI status



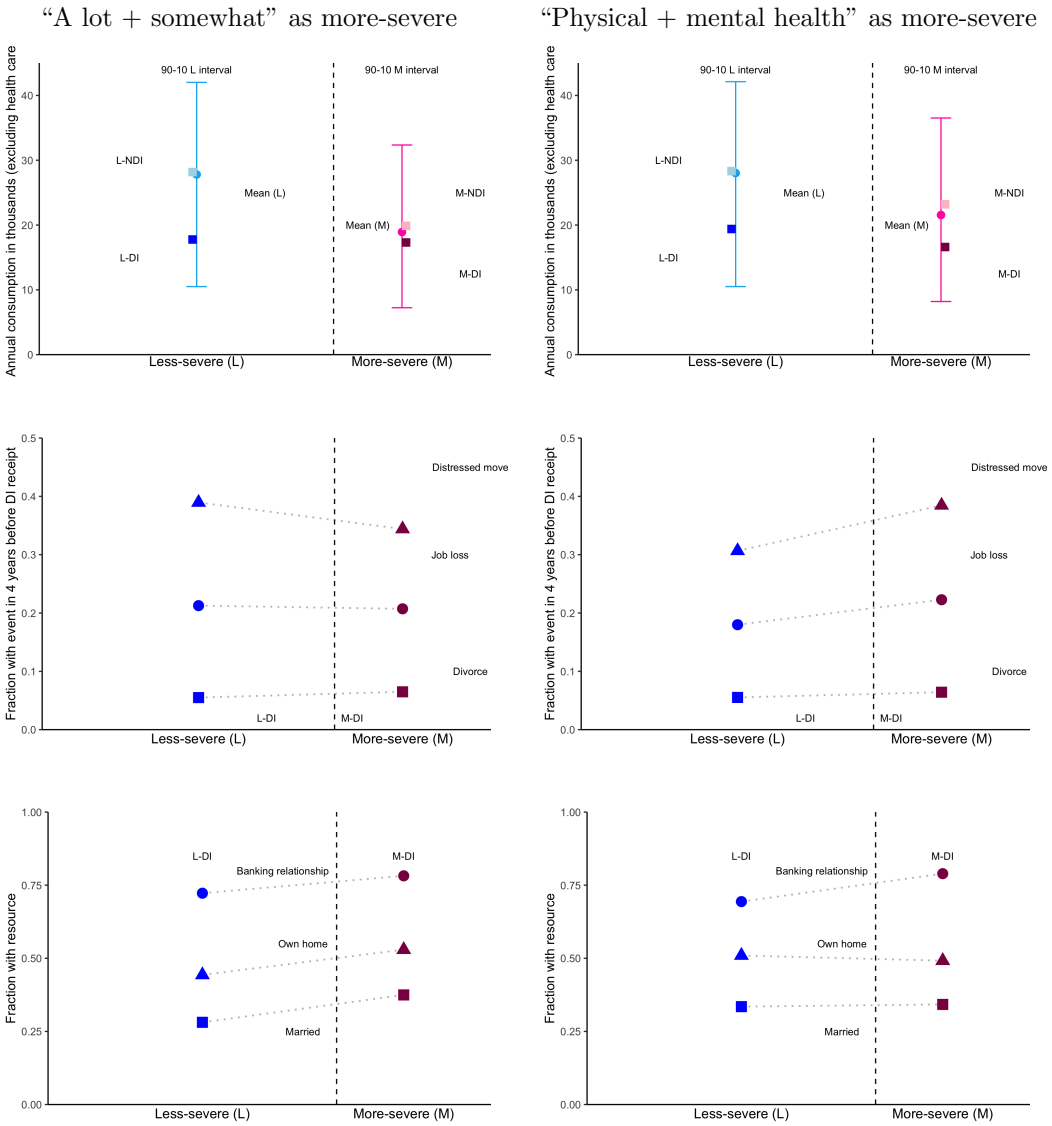

Notes: Figure presents rates of experiencing an "adverse life event"—head or spouse job loss, involuntary move, or divorce—at various ages for households in the PSID, by SSDI status, and age at SSDI entry. Households are first categorized by their SSDI status: whether they ever received SSDI (DI) or never received SSDI (NDI). The DI group is further categorized by age at SSDI entry: <35, 35–44, 45–54, 55–64. We organize these groups along the horizontal axis and then plot the rate of experiencing any adverse event at the following ages along the vertical axis: 25–34, 35–44, 45–54, and 55–64. Darker shades indicate higher rates, as shown in the legend. Gray squares with a diamond indicate that the sample size is less than 50. The sample of each cell is households in SSDI/age category indicated in the horizontal axis that are observed in the PSID during the age range indicated in the vertical axis.

Figure J9: PSID: Any adverse life event by SSDI status



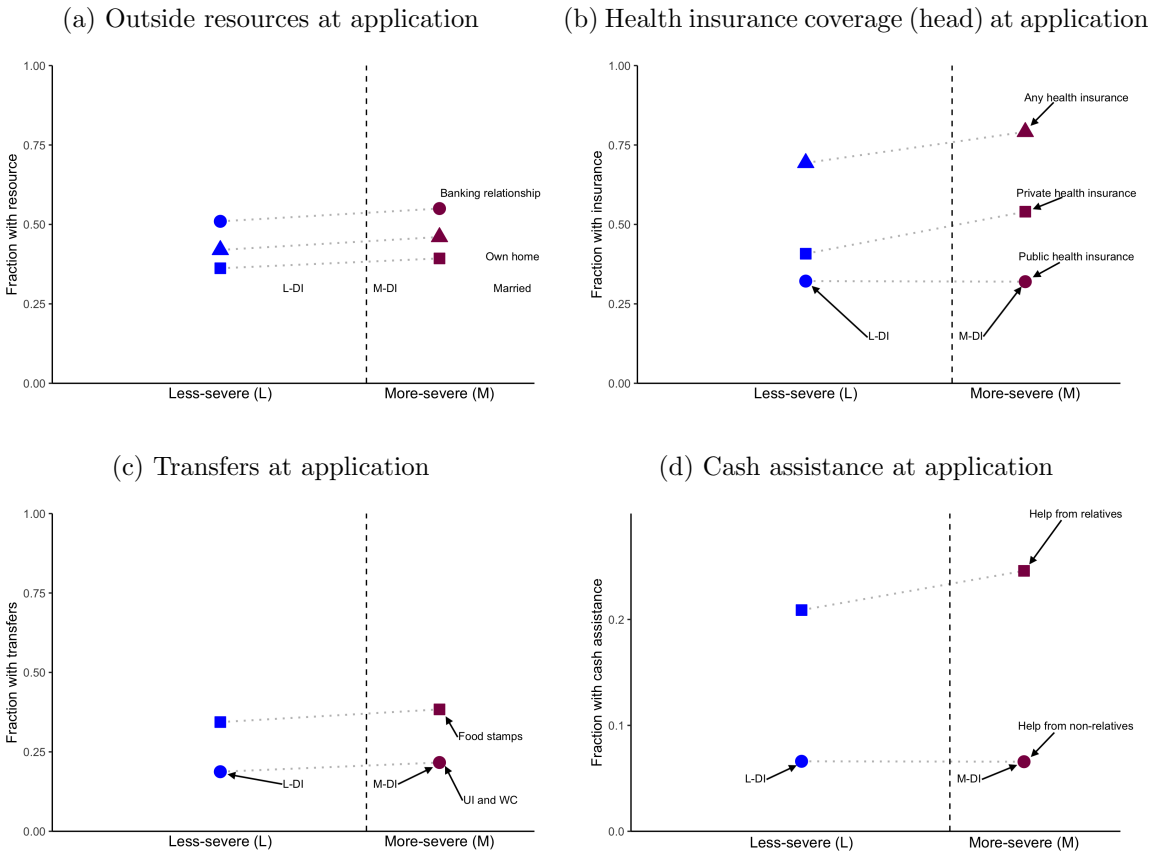

Notes: Figure presents rates of experiencing an "adverse life event"—head or spouse job loss, involuntary move, or divorce—at various ages for households in the PSID, by SSDI status, and age at SSDI entry. Households are first categorized by their SSDI status: whether they ever received SSDI (DI) or never received SSDI (NDI). The DI group is further categorized by age at SSDI entry: <35, 35–44, 45–54, 55–64. We organize these groups along the horizontal axis and then plot the rate of experiencing any adverse event at the following ages along the vertical axis: 25–34, 35–44, 45–54, and 55–64. Darker shades indicate higher rates, as shown in the legend. Gray squares with a diamond indicate that the sample size is less than 50. The sample of each cell is households in SSDI/age category indicated in the horizontal axis that are observed in the PSID during the age range indicated in the vertical axis.

Figure J10: PSID: Positive results using alternative definitions of more-severe




Notes: These figures replicate Figures J1a, J6, and J1c using two alternative definitions of more-severe. In the left column, we use a broader definition of more-severe that includes the "somewhat" response (in addition to the baseline "a lot" response) to the question of how much a health condition limits work. In the right column, we use a definition that incorporates mental health on top of the baseline severity measure—in particular, having depression, psychiatric issues, or "loss of memory or mental ability." The figures in the first row represent the 90-10 percentile interval (and average) of consumption for more-severe and less-severe households in the 2017 PSID, as in Figure J1a. The figures in the second row represent rates of various adverse life events in the four years before receiving SSDI for households receiving SSDI in the 2017 PSID, by more-severe and less-severe, as in Figure J6. The figures in the third row present rates of marriage, banking relationship, and homeownership for households receiving SSDI in the 2017 PSID, by more-severe and less-severe, as in Figure J1c. See notes in original figures for more details on each measure.

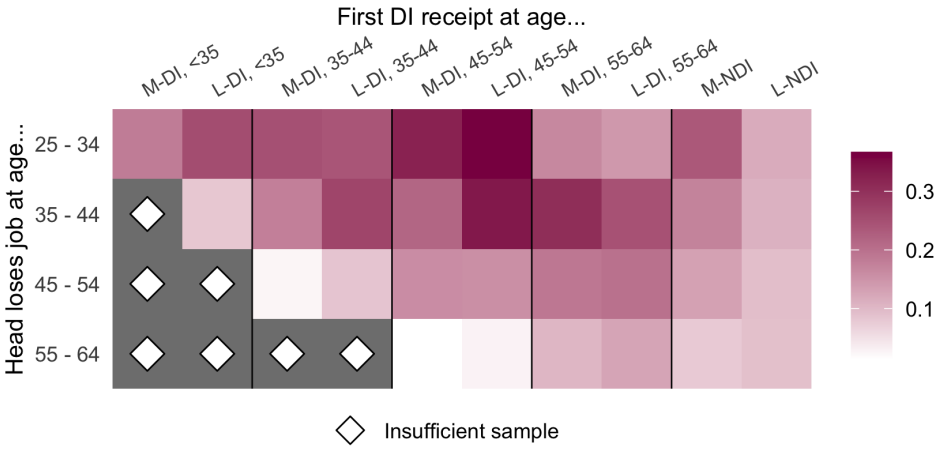
Figure J11: PSID: Resources available at application



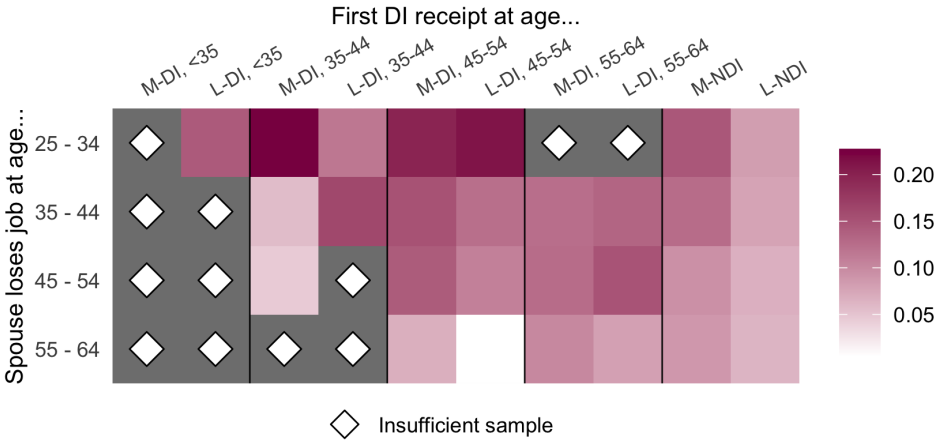

Notes: Figures present rate of outside resources one year prior to the year of initial SSDI receipt for each of the two-severity-by-SSDI-receipt groups in the 2017 PSID: more-severe SSDI-recipients (216 M-DI), and less-severe recipients (274 L-DI). Figure J11a presents rates of marriage, banking relationship, and homeownership, Figure J11b presents rates of any health insurance, private health insurance, and public health insurance, Figure J11c presents rates of SNAP (food stamps) and unemployment insurance (UI) and workers' compensation (WC), and Figure J11d presents rates of cash assistance from relatives and non-relatives. "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe." Appendix Tables I4 and I5 present the exact PSID questions for these measures.

Figure J12: PSID: Job loss by SSDI and health status

(a) Head job loss by SSDI and health status



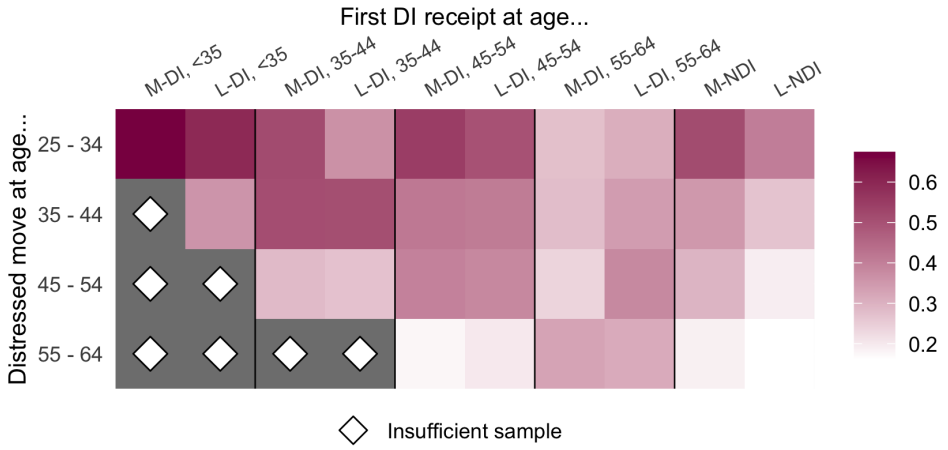

(b) Spouse job loss by SSDI and health status



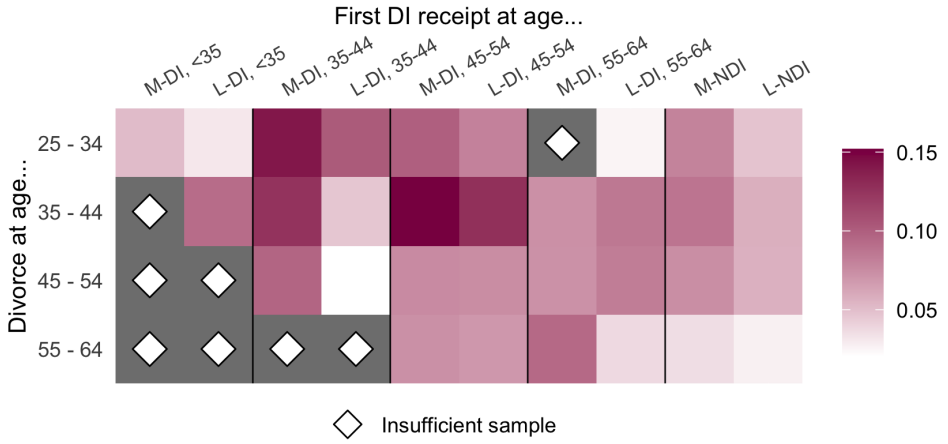

Notes: Figures present rates of experiencing head and spouse job loss at various ages for households in the PSID, by SSDI status, health status, and age at SSDI entry. Households are first categorized by their SSDI status: whether they ever received SSDI (DI) or never received SSDI (NDI). SSDI are further classified by their health status at SSDI entry (more-severe M-DI, or less-severe L-DI), and NDI by whether they ever had a more-severe health condition (ever more-severe M-NDI, never more-severe L-NDI). The M-DI and L-DI groups are further categorized by their age at SSDI entry: <35, 35–44, 45–54, 55–64. We organize these groups along the horizontal axis and then plot the rate of experiencing any adverse event at the following ages along the vertical axis: 25–34, 35–44, 45–54, and 55–64. Darker shades indicate higher rates, as shown in the legend. Gray squares with a diamond indicate that the sample size is less than 50. The sample of each cell is households in SSDI/age category indicated in the horizontal axis that are observed in the PSID during the age range indicated in the vertical axis. Job loss includes "involuntary" reasons for separation: strike/lockout, laid off/fired, or company going out of business or leaving town (exact PSID question in Appendix Table I2). "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe."

Figure J13: PSID: Distressed move and divorce by SSDI and health status

(a) Distressed move by SSDI and health status




(b) Divorce by SSDI and health status




Notes: Figures present rates of experiencing distressed move and divorce at various ages for households in the PSID, by SSDI status, health status, and age at SSDI entry. Households are first categorized by their SSDI status: whether they ever received SSDI (DI) or never received SSDI (NDI). SSDI are further classified by their health status at SSDI entry (more-severe M-DI, or less-severe L-DI), and NDI by whether they ever had a more-severe health condition (ever more-severe M-NDI, never more-severe L-NDI). The M-DI and L-DI groups are further categorized by their age at SSDI entry: <35, 35–44, 45–54, 55–64. We organize these groups along the horizontal axis and then plot the rate of experiencing any adverse event at the following ages along the vertical axis: 25–34, 35–44, 45–54, and 55–64. Darker shades indicate higher rates, as shown in the legend. Gray squares with a diamond indicate that the sample size is less than 50. The sample of each cell is households in SSDI/age category indicated in the horizontal axis that are observed in the PSID during the age range indicated in the vertical axis. Distressed move includes external events like eviction, contraction of housing (less space/less rent), and other reasons potentially indicating distress, such as saving money (exact PSID question in Appendix Table I3). Divorce is defined as being married in a previous survey year but not in this survey year (exact PSID question in Appendix Table I2). "More-severe" is defined as self-reporting that a health condition limits "a lot" the amount of work one can do, as opposed to "somewhat," "just a little," or "not at all" (or no health condition) for "less-severe."