Washington Center for Equitable Growth
Evidence for a stronger economy

equitablegrowth.org

Working paper series

Beyond Health: Non-Health Risk and the Value of Disability Insurance

Manasi Deshpande
Lee M. Lockwood

November 2021

https://equitablegrowth.org/working-papers/beyond-health-non-health-risk-and-the-value-of-disability-insurance/

facebook.com/equitablegrowth
@equitablegrowth
equitablegrowth.org/feed
info@equitablegrowth.org

1156 15th St NW Suite 700
Washington DC 20005
202-545-6002

Beyond Health: Non-Health Risk and the Value of Disability Insurance*

Manasi Deshpande Lee M. Lockwood

August 30, 2021

Abstract

The public debate over disability insurance has centered on concerns about individuals without severe health conditions receiving benefits. We go beyond health risk alone to quantify the overall insurance value of U.S. disability programs, including value from insuring non-health risk. We find that disability recipients, especially those with less-severe health conditions, are much more likely to have experienced a wide variety of non-health shocks than non-recipients. Selection into disability receipt on the basis of non-health shocks is so strong among individuals with less-severe health conditions that by many measures less-severe recipients are worse off than more-severe recipients. As a result, under baseline assumptions, benefits to less-severe recipients have an annual surplus value (insurance benefit less efficiency cost) over cost-equivalent tax cuts of $7,700 per recipient, about three-fourths that of benefits to more-severe recipients ($9,900). Insurance against non-health risk accounts for about one-half of the value of U.S. disability programs.

*Deshpande: University of Chicago and NBER, mdeshpande@uchicago.edu; Lockwood: University of Virginia and NBER, leelockwood@virginia.edu. Deshpande conducted this research as a Visiting Economist at the Social Security Administration (SSA). We thank the Office of Research, Evaluation, and Statistics at SSA for providing access to administrative data, and Katherine Bent, Ben Danforth, Jeffrey Hemmeter, Steve Robinson, Mark Sarney, Alexander Strand, Robert Weathers, and several SSA components for feedback. We are grateful to David Autor, Katherine Baicker, Julie Berry Cullen, Joseph Doyle, Arindrajit Dube, Mark Duggan, Liran Einav, Amy Finkelstein, Leora Friedberg, Michael Greenstone, Steven Haider, Jill Horwitz, William Johnson, Daniel Kaliski, Henrik Kleven, Jonathan Kolstad, Magne Mogstad, Timothy Moore, Austin Nichols, Matthew Notowidigdo, Edgar Olsen, Luigi Pistaferri, James Poterba, Adam Sacarny, Joseph Shapiro, Isaac Sorkin, Stefanie Stantcheva, Melanie Wasserman, and seminar participants at Federal Reserve Board, Michigan State, MIT, Stanford, Illinois-Chicago, and Wisconsin-Madison for helpful comments. We thank Yalun Su and Jiada Ye for excellent research assistance. Deshpande thanks the Ronzetti Initiative for the Study of Labor Markets at the Becker-Friedman Institute for financial support. This material is based upon work supported by the National Science Foundation under Grant No. 1941538.

1 Introduction

Disability insurance is a major component of the social safety net in developed countries. In the United States, the Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI) programs (which we collectively refer to as U.S. disability programs, or USDP) together provide access to health insurance and $200 billion annually in cash benefits to nearly 13 million Americans. The main purpose of these programs is to provide assistance to people who cannot work because of severe health conditions. Yet the expansion of U.S. disability programs in recent decades has been attributed at least in part to non-health factors like stagnating wages (Autor and Duggan, 2003, 2006; Liebman, 2015), and there is widespread concern that providing benefits to individuals without severe health conditions is diluting the value of these programs.[1]

In this paper, we go beyond health risk alone to quantify the overall insurance value U.S. disability programs, including value from insuring non-health risk. We quantify the extent to which these programs insure different risks by comparing disability recipients and non-recipients along a wide variety of health and non-health dimensions, including consumption, adverse events like job loss, and resources available to cope with adverse events. Building on these comparisons, we make additional assumptions to quantify the (ex ante) value of receiving disability benefits in different states of the world: the net of the insurance benefit from receiving transfers in high-marginal utility states less the efficiency cost from distorting behavior. Reflecting the public debate over USDP, we focus on "mismatches with respect to health": individuals who receive benefits despite not having a health condition that exceeds a certain severity threshold ("less-severe recipients") and individuals who do not receive benefits despite having such a condition ("more-severe non-recipients").[2]

Going "beyond health" is crucial for making informed judgments about disability programs and their mismatches with respect to health. Health, though likely a strong indicator of the value of receiving disability benefits, is not a *perfect* indicator

[1] More recently, growth in U.S. disability programs has slowed. Examples of concern include, "Unfit for Work: The startling rise of disability in America," Chana Joffe-Walt, *NPR*, March 23, 2013, and "How Americans Game the $200 Billion-a-Year 'Disability-Industrial Complex,'" Avik Roy, *Forbes*, April 8, 2013.

[2] Such mismatches are not necessarily errors as defined by USDP rules, which explicitly account for certain non-health factors in eligibility determination (see Section 3 for details). We use "less-/more-severe individual" as shorthand for "individual with a less-/more-severe health condition."

because individuals face major non-health risks as well, including job loss, productivity shocks, and changes in family structure. To the extent that a particular risk is not completely insured by other means, disability insurance potentially insures or exacerbates that risk, depending on selection into receiving disability benefits. For example, disability benefits in states of the world without a health shock but with limited earning prospects following job layoffs partially insure that risk. More generally, although mismatches with respect to health necessarily reduce the extent to which disability insurance insures health risk, the presence of non-health risk means that such mismatches might not reduce—and might even increase—the overall value of disability insurance. The extent to which mismatches insure non-health risk is therefore a critical determinant of the value of potential reforms. For example, unless benefits to less-severe recipients provide highly valuable insurance against non-health risk, reforms to limit such benefits, even at the expense of decreasing benefits to more-severe recipients, could produce large gains. But if benefits to less-severe recipients provide valuable insurance against non-health risk, such reforms could reduce the value of disability insurance.[3]

The extent to which disability programs and their mismatches with respect to health insure or exacerbate different risks is ultimately an empirical question, one that we investigate using complementary positive and normative analyses. In the positive analysis, we use a combination of survey and administrative data to establish new facts about the targeting of disability benefits on the basis of non-health factors. We find that less-severe disability recipients are on average much worse off than less-severe non-recipients, and by many non-health measures are even worse off than more-severe recipients. For example, we find in administrative data that prior to receiving disability benefits, less-severe recipients are 40% more likely to have experienced a mass layoff than more-severe recipients, 19% more likely to have experienced a foreclosure, and 23% more likely to have experienced an eviction. From survey data, less-severe recipients have similarly low consumption levels to more-severe recipients just prior to receiving disability benefits ($17,000 and $16,400, respectively). Conversely, more-severe non-recipients are on average *better* off than either more-severe

[3]Although it is natural to think that a risk would optimally be insured by "its own" program, that is not necessarily true when risks are richer than feasible policy instruments. For example, unemployment insurance typically does not insure any of the unemployment-related risk *within* unemployment spells of a given length, such as risk in re-employment wages or resources with which to smooth consumption during the spell.

or less-severe recipients. The strong associations between disability receipt and non-health shocks conditional on health means that USDP mismatches with respect to health are highly selected, in ways likely to increase the value of USDP, by both increasing insurance of non-health risk and decreasing distortion costs.

In the normative analysis, we determine how the targeting properties of U.S. disability programs, including their mismatches with respect to health, translate into their ex ante value, taking other programs as given. We use a flexible, sufficient statistics-type approach that highlights the importance, in a broad class of models, of two key factors in determining the value of disability benefits: the marginal utility of income and "counterfactual earnings"—earnings if not receiving disability benefits. We model marginal utility using PSID consumption measures and a wide range of assumptions about the functional form of utility. We model counterfactual earnings using quasi-experimental evidence from the literature on the effects of being awarded disability benefits. Our findings suggest that U.S. disability programs provide valuable insurance and increase welfare significantly. Under baseline assumptions, the value of disability benefits exceeds that of cost-equivalent tax cuts by 64%, creating a surplus worth $8,700 of government revenue per recipient per year. Moreover, we find that the high value of USDP is in part *because of*, not despite, mismatches with respect to health. We estimate that benefits to less-severe recipients create a value (insurance benefit less distortion cost) over cost-equivalent tax cuts of $7,700 per recipient per year, about three-fourths that of benefits to more-severe recipients ($9,900). As a result, benefits to less-severe recipients do not decrease the value of USDP—they increase it considerably, accounting for about half of the total value. The *lack* of benefits to more-severe non-recipients also appears to create value, as providing such benefits would have a negative value (−$2,200 per hypothetical recipient per year). USDP mismatches with respect to health are so favorably selected that U.S. disability programs are more valuable than a hypothetical, infeasible program that would target perfectly on severity. In other words, even somehow eliminating all mismatches at no administrative cost would significantly *decrease* the value of USDP. The estimates and conclusions are similar considering the cash benefits alone or the cash benefits combined with health insurance. Including health insurance modestly decreases the ratio of per-recipient surplus of benefits to less-severe relative to more-severe recipients, from 0.78 to 0.74–0.77 depending on the exact specification.

The high values of U.S. disability programs and their mismatches with respect to

health are driven by favorable selection on both marginal utility and counterfactual earnings. Much of the favorable selection into disability receipt is on the basis of non-health shocks. Using a decomposition analysis, we find that about one-half of the insurance value of USDP comes from insuring risk *within* observable health groups. This suggests that much, or even most, of the value of USDP comes from insuring non-health risk. USDP insurance of non-health risk in turn depends crucially on the strong selection into application, rather than on selective awards conditional on application. We estimate that a hypothetical disability program with similar adjudication standards but no systematic selection into application would have *negative* value, far below the large positive value of USDP. Although we lack direct evidence on what drives selection into application, additional simulation results suggest that earnings restrictions on recipients could play a key role.

We emphasize several points relevant for interpreting our results. First, our analysis, like most welfare analyses, takes other policies as given. That U.S. disability programs provide valuable insurance against a wide variety of risks may in part reflect gaps in the existing social safety net and does not imply that these programs are the most efficient way to insure these risks. Expanding other insurance programs like unemployment insurance or creating new programs like wage insurance or temporary disability insurance would potentially reduce the value of USDP, a point we revisit in Section 5.5. Second, our normative analysis focuses on insurance and welfare effects and does not explicitly incorporate non-welfarist considerations.[4] Third, our results do not speak to the effects of changes to particular features of USDP, such as the earnings limit or the weight on health in the adjudication process, since such changes would induce application responses, evidence on which is sparse. Nor does our analysis speak to the alignment of U.S. disability programs with their program

[4]Such considerations might include the view that disability benefits should be received only by individuals with certain health conditions or the view that certain individuals are more deserving of benefits than others. Because quantitative evidence on the strength of such views is limited, rather than incorporate such considerations in our main analysis, we instead quantify how strong such considerations would have to be, in terms of their effects on generalized social marginal welfare weights (Saez and Stantcheva, 2016), in order to overturn our key conclusions. We find that a non-welfarist planner who discounted the value of disability benefits to less-severe recipients (viewing each dollar of such benefits as being worth \$(1-d), where d is the discount) would get positive surplus from the actual disability benefits to less-severe recipients as long as the discount did not exceed about 35% ($d = 0.35$). That same planner would get positive surplus from disability benefits as a whole as long as the discount on benefits to less-severe recipients did not exceed about 75% ($d = 0.75$).

rules, since mismatches with respect to health are distinct from programmatic errors with respect to their rules, which include both health and non-health considerations. Fourth, the results do not imply that expanding disability receipt to new groups would increase social welfare—in fact, our analysis highlights that the value of USDP depends crucially on selective receipt. Finally, our normative analysis relies on several assumptions, including about behavioral responses to disability benefits, about which relatively little is known. Gaining a better understanding of such behavioral responses and the benefits and costs of reforms to other policies is a high priority for future research.

Acknowledging these caveats, our results have implications for the current debate over disability programs. The most contentious question in the debate is whether benefits to individuals without more-severe health conditions are diluting the value of disability programs. Our results suggest that in the case of U.S. disability programs, such benefits are not diluting value but enhancing it. Whereas eliminating such benefits would significantly reduce the value of USDP, "doubling down" on them by increasing benefit levels—which, among other things, further increases transfers from more-severe non-recipients to less-severe recipients—would increase it.

The primary contribution of this paper is to quantify the extent to which U.S. disability programs and their mismatches with respect to health insure non-health as well as health risk and to determine the welfare implications of this insurance. Our findings help reconcile two seemingly contradictory strands of the literature on disability programs. The first is a positive, empirical strand that has raised serious concerns about disability programs. The key findings of this literature are that mismatches with respect to health are prevalent, negative labor market shocks increase disability enrollment, and disability programs reduce the labor supply of marginal recipients.[5] These findings suggest that disability programs may provide benefits to many individuals without severe health conditions and that they discourage work. Our paper extends this literature by going "beyond health" to characterize more-severe and less-severe disability recipients and non-recipients on a wide variety of dimensions and to quantify the implications for the value of disability programs. The second, normative strand of literature on disability programs quantifies the welfare

[5]Mismatches: Benitez-Silva, Buchinsky and Rust (2004); Von Wachter, Song and Manchester (2011). Labor market shocks: Autor and Duggan (2003, 2006); Black, Daniel and Sanders (2002); Charles, Li and Stephens Jr (2018); Rege, Telle and Votruba (2009). Labor supply: Maestas, Mullen and Strand (2013); French and Song (2014).

effects of various reforms. This strand typically finds that expanding disability programs would increase welfare.[6] Our paper extends this literature by quantifying, with a flexible approach, the value of disability benefits in different types of states of the world and by illuminating the mechanisms underlying the welfare effects of disability programs, particularly the role of non-health risk. Methodologically, our decompositions of insurance value into underlying mechanisms help address what is arguably the main limitation of sufficient statistics-type analyses: their black-box nature.[7]

How could expanding U.S. disability programs increase welfare despite prevalent mismatches with respect to health, non-health drivers, and reductions in work? Our results help reconcile this tension. The strong selection into disability receipt on the basis of non-health shocks means that mismatches with respect to health do not greatly decrease—and in fact appear to increase—the value of USDP. Non-health drivers of enrollment do not just drive up the costs of USDP, they also drive up the benefits in terms of insuring non-health risk. With respect to effects on labor supply, we do not provide new evidence on this question but instead quantify the implications of existing evidence for the value of USDP. Our analysis highlights the importance of a less-discussed aspect of the existing evidence: that marginal disability recipients, despite their non-negligible counterfactual labor force *participation*, have extremely low counterfactual *earnings*.[8]

[6]Bound et al. (2004) find that the typical worker would gain from a government budget-neutral marginal expansion of SSDI. Chandra and Samwick (2009) and Meyer and Mok (2019) find that disability programs, together with other programs, provide valuable but incomplete insurance against health risk. Low and Pistaferri (2015) estimate a structural life-cycle model and find that expanding USDP would increase welfare. Cabral and Cullen (2019) use the demand for private disability insurance to bound the value of public disability benefits and find a high lower bound. The finding by Gelber, Moore and Strand (2018) that larger SSDI payments reduce mortality is in keeping with these welfare results. While existing research on the overall value of disability programs implicitly includes any value from insurance of non-health risk, to the best of our knowledge no other paper quantifies it or investigates underlying drivers.

[7]This could be useful in other settings as well. For example, decomposing the value of unemployment insurance into value from its support during shorter versus longer unemployment spells or for individuals with more versus less liquid assets would be informative about underlying determinants of its value and suggestive of potential ways in which the program might be altered to increase its value.

[8]For example, using the quasi-random assignment of applicants to adjudicators, Maestas, Mullen and Strand (2013) find that SSDI reduces the labor force participation of applicants on the margin of program entry by 28 percentage points, but that such applicants have average counterfactual earnings of just \$6,000–\$9,000 per year. Using a similar strategy, French and Song (2014) find that in the ten years after the SSDI/SSI decision, counterfactual earnings actually decline in later years after reaching a maximum of about \$5,000 per year around year five. Of course, these estimates do not reveal *why* counterfactual earnings are so low—whether because of health, skills, work ethic,

This paper also contributes to the broader literature on the take-up and targeting of government programs. The main focus of this literature has been incomplete take-up (see Currie, 2006, for a review), with more recent work shedding light on the normative implications of incomplete take-up by studying the characteristics of those who do versus do not take up benefits.[9] We find, in the context of USDP, that incomplete take-up among more-severe individuals significantly increases the ex ante value of the program. Improper take-up, by contrast, has received much less attention in economic research, despite being central to public debates, from that in the 1970s about "welfare queens" to those more recently about disability awards and EITC overpayments.[10] Our paper finds that benefits to less-severe individuals, which can be thought of as improper take-up relative to a specific benchmark, account for nearly one-half of the high value of USDP. Given that the risk of experiencing a more-severe work-limiting health condition might be the biggest risk that people face, our finding that the value of USDP is driven to such a large extent by mismatches with respect to health—i.e., by selective take-up with respect to *other* risks—suggests that selective take-up with respect to risks beyond a program's primary aim might be important in other applications as well.

2 Theory: Non-health risk, mismatches with respect to health, and the value of disability insurance

Our goal is to go "beyond health" to determine the value of disability programs and their mismatches with respect to health accounting for non-health risk. Consider a simple model in which an individual faces two kinds of risk: health and non-health, where h is the severity of the health shock and n is the seriousness of the non-health

etc.—which is relevant for certain non-welfarist perspectives.

[9]See, e.g., Bhargava and Manoli (2015); Alatas et al. (2016); Deshpande and Li (2019); Finkelstein and Notowidigdo (2019); Lieber and Lockwood (2019); Homonoff and Somerville (2020).

[10]The main economic research on improper take-up is earlier work on disability programs (e.g., Parsons (1991); Diamond and Sheshinski (1995); Low and Pistaferri (2015)). Beyond disability, recent work on the EITC by Jones and Ziliak (2019) finds that many of the households that receive EITC benefits despite being ineligible have low incomes. Kleven and Kopczuk (2011) analyze in theory how efforts to establish eligibility for a program trade off inclusion and exclusion errors. They find that optimal programs feature errors of both kinds.

shock. Consider a "health-based perspective" in which the objective of disability insurance is to provide benefits to the individual if and only if they have suffered a health shock of severity $\bar{h}$ or greater. This perspective divides the state space into two sets: states in which the individual has versus has not experienced a qualifying health shock: $h \gtrless \bar{h}$. We label states in which $h \geq \bar{h}$ "health-shock" and states in which $h < \bar{h}$ "non-health-shock." There are two types of mismatches relative to this perspective: awarding benefits in a non-health-shock state (an inclusion mismatch) and failing to award benefits in a health-shock state (an exclusion mismatch).

Now consider an alternative, "value-maximizing" perspective in which the goal is to maximize the ex ante value of disability insurance by providing benefits if and only if the insurance benefit exceeds the distortion cost—i.e., if and only if doing so produces positive (ex ante) surplus value, $s(h, n) > 0$. As we derive formally in Section 5, the insurance benefit depends mainly on the marginal utility of income, and the distortion cost depends mainly on what earnings would have been in the absence of disability benefits ("counterfactual earnings"), reflecting the reduction in labor supply from receiving disability benefits. This perspective divides the state space into two sets: states in which the value of receiving disability benefits is greater or less than zero: $s(h, n) \gtrless 0$. There are two types of mismatches relative to this perspective: awarding benefits when doing so reduces the value of disability insurance, and failing to award benefits when doing so would have increased the value of disability insurance.

Figure 1 shows how these two perspectives divide the state space into four mutually exclusive, exhaustive sets. It depicts a hypothetical joint $(h, s(h, n))$ distribution in the natural case in which health is a strong but noisy indicator of surplus. Health is likely a strong indicator of the value of receiving disability benefits because health shocks increase health spending and limit earning opportunities, and so tend to increase marginal utility and decrease counterfactual earnings. These two perspectives agree on providing benefits in the "Northeast" quadrant, states in which the individual suffers a health shock and receiving disability benefits produces positive surplus. They also agree on *not* providing benefits in the "Southwest" quadrant (likely the vast majority of states), states in which the individual does not suffer a health shock and receiving disability benefits produces negative surplus.

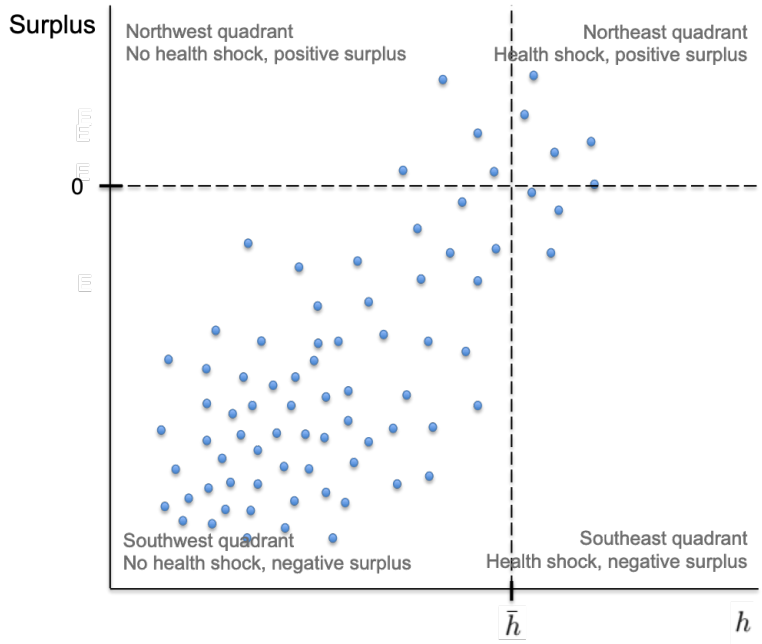
Yet the value of receiving disability benefits in a particular state is not entirely determined by the health shock. Insurance value is likely increasing in the seriousness

of the non-health shock since transfers into states in which marginal utility is high provide valuable insurance, regardless of why marginal utility is high. In terms of distortions, benefits in states in which counterfactual earnings are low tend to have low efficiency costs regardless of why counterfactual earnings are low. In Figure 1, while $s(h, n)$ is increasing in h on average, variation in n causes it to vary around that average. As a result, the two perspectives occasionally disagree on whether benefits should be provided. They disagree about states in the "Southeast" quadrant, in which the individual suffers a health shock but the non-health shock is sufficiently favorable (e.g., high spousal earnings) to make surplus negative. The perspectives also disagree about states in the "Northwest" quadrant, in which the individual does not suffer a health shock but the non-health shock (e.g., productivity shock not insured by UI) is sufficiently unfavorable to make surplus positive.

In settings with substantial non-health risk, the value of disability insurance and the costs of its mismatches with respect to health depend crucially on the extent to which they insure or exacerbate non-health risk, which in turn depends on selection into disability receipt on the basis of non-health shocks. Mismatches with respect to health necessarily reduce the extent to which disability insurance insures health risk—in this context, the risk of experiencing a health shock ($h \geq \bar{h}$)—by reducing the extent to which disability insurance transfers from non-health-shock to health-shock states. But mismatches could either insure or exacerbate non-health risk, making their net effect on the insurance value of disability insurance theoretically ambiguous. If individuals with more-serious non-health shocks are more (less) likely to receive disability benefits conditional on health, then mismatches with respect to health insure (exacerbate) non-health risk.[11] Assuming that health is a strong indicator of the value of receiving disability benefits, a representative exclusion mismatch—the failure to provide benefits in a state with surplus $E(s(h, n)|M)$—would forego cost-effective insurance. Similarly, a representative inclusion mismatch—providing benefits in a state with surplus $E(s(h, n)|L)$—would both exacerbate risk (by transferring to

[11]More precisely, exclusion mismatches insure risk within the set of states with health shocks—which we will mostly refer to as "non-health risk," though it could reflect a mix of health and non-health risk as defined by h and n—if and only if mean marginal utility in health-shock recipient states exceeds that in health-shock non-recipient states, and inclusion mismatches insure risk within the set of states without health shocks (also "non-health risk") if and only if mean marginal utility in non-health-shock recipient states exceeds that in non-health-shock non-recipient states. The less variable is health within health-shock and non-health-shock states, the greater the extent to which within-health-shock and within-non-health risk reflects non-health risk as defined by n.

Figure 1: Health-based vs. value-maximizing classifications of states



Notes: Figure shows how health-based and value-maximizing perspectives classify states of the world into those in which disability benefit receipt is versus is not desired. The x-axis measures the severity of the health shock, h. The y-axis measures the ex ante surplus of receiving disability benefits. From a health-based perspective, disability benefit receipt is desired if and only if $h \geq \bar{h}$. From a value-maximizing perspective, disability benefit receipt is desired if and only if $s(h, n) > 0$, where n is the seriousness of the non-health shock. These two perspectives both desire disability benefit receipt in states in the Northeast quadrant and non-receipt in states in the Southwest quadrant, but differ in their assessment of states in the Northwest and Southeast quadrants. $E(s|M)$ denotes mean surplus for $h \geq \bar{h}$ states, and $E(s|L)$ denotes mean surplus for $h < \bar{h}$ states. The blue dots depict a hypothetical joint $(h, s(h, n))$ distribution in the natural case in which health is a strong but noisy indicator of surplus.

a state with below-average marginal utility) and involve substantial efficiency costs (by creating a large labor supply distortion).

But the actual inclusion and exclusion mismatches of a particular disability program may not be representative if the program's tags and screens lead to systematic selection into applying for and being awarded disability insurance. The U.S. SSDI and SSI programs, for example, use a vocational grid that takes into consideration certain non-health characteristics ("tags") in award decisions, such as an applicant's age, education, and work experience. SSDI and SSI also impose restrictions on applicants and recipients ("screens"), such as earnings restrictions and a complicated

application process.[12] Systematic selection could either increase or decrease the cost of actual mismatches relative to representative ones. For example, Deshpande and Li (2019) find that higher application costs disproportionately discourage individuals with relatively severe conditions and limited resources from applying for SSDI and SSI benefits. In Figure 1, such an unfavorable exclusion mismatch would be in the "Northeast" quadrant (health shock and high surplus). On the other hand, strict earnings limits might discourage applications in states in which the individual can find suitable employment despite a health shock. In Figure 1, such a favorable exclusion mismatch would be in the "Southeast" quadrant (health shock but low surplus). Whether a disability program's tags and screens create value-enhancing or -reducing selection is ultimately an empirical question, one we aim to answer for U.S. disability programs in the following positive and normative analyses.

3 Institutional setting and data

Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI) (collectively "USDP") are intended to provide assistance to Americans who cannot work because of a severe health condition. SSDI and SSI use the same definition of disability.[13] Non-health factors, such as age and education, factor in through the medical-vocational grids.[14] For SSDI, eligibility is limited to individuals with qualifying earnings histories. The average annual SSDI cash benefit is around $15,000, and beneficiaries qualify for Medicare after a two-year waiting period.[15] Disabled workers are not permitted to earn more than the substantial gainful activity (SGA)

[12]Some screens are chosen intentionally to try to improve targeting, such as the restrictions on the earnings of applicants and recipients. Others, such as the complicated application process and the uncertainty in award decisions, are unintentional byproducts of program administration but can still be crucial to a program's targeting properties (Kleven and Kopczuk, 2011).

[13]Under Social Security rules, someone is considered disabled if (i) "[they] cannot do work that [they] did before," (ii) "[SSA] decide[s] that [they] cannot adjust to other work because of [their] medical condition(s)," *and* (iii) "[their] disability has lasted or is expected to last for at least one year or to result in death" (see `https://www.ssa.gov/benefits/disability/qualify.html`).

[14]This is a key difference between SSDI and SSI's eligibility rules, which account for certain non-health factors, and our classification of individuals on the basis of health alone. As a result, some correct awards under SSDI and SSI's rules are mismatches with respect to health by our definition, and some incorrect awards under SSDI and SSI's rules are not mismatches with respect to health by our definition. Our goal is not to judge SSDI and SSI's alignment with their rules but to characterize the health and non-health characteristics of recipients and non-recipients and to derive the implications for the value of benefits to different groups.

[15]Annual Statistical Report on the Social Security Disability Insurance Program, 2018.

level, which was $1,260 per month ($15,120 per year) in 2020. SSI is a means-tested program that has the same medical eligibility requirements and SSA adjudication process as SSDI but does not require a work history. The maximum federal SSI benefit for an individual was $783 per month ($9,396 per year) in 2020, and the SSI benefit is reduced by 50 cents for each dollar of earnings. SSDI and SSI recipients receive periodic medical reviews in which the adjudicator considers the recipient's medical condition and work activity. To provide a comprehensive picture of U.S. disability programs, we consider the SSDI and SSI programs together in our main analysis. We provide estimates for SSDI alone in Appendix J.

In Section 4, we use a combination of survey data and administrative data to study the characteristics of different groups of working-age households by their USDP status and severity. The first data source is the Panel Study of Income Dynamics (PSID), which we use to estimate differences by severity and USDP status in consumption, adverse non-health events (e.g., job loss), and resources (e.g., marriage). The main advantages of the PSID are its measures of consumption and adverse events and its long panel structure. We also use Social Security Administration data to measure adverse non-health events for more- and less-severe recipients. In particular, we measure mass layoffs as well as bankruptcy, eviction, and foreclosure from the data compiled by Deshpande, Gross and Su (2021). In Section 5, we use PSID data on consumption and earnings, together with quasi-experimental estimates from French and Song (2014), to estimate the surplus from disability benefits to different groups.

Following the literature (e.g., Low and Pistaferri, 2015), we classify households in the PSID 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, and "less-severe" otherwise. In the SSA administrative data, 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. Our results are robust to alternative measures (see Appendix A for more details). Appendix A includes more detail on all data sources, and Appendix B discusses the robustness of the main results to measurement error in severity and disability receipt. Our findings suggest that measurement error usually, though not always, works against our main conclusions.

4 Positive analysis: Non-health risk and selection into U.S. disability programs

Our goal is to understand selection into U.S. disability programs on the basis of non-health shocks conditional on health. To that end, we characterize mismatch groups—less-severe recipients (L-DI) and more-severe non-recipients (M-DI)—on non-health dimensions. We focus on non-health characteristics likely to be predictive of the key determinants of the value of receiving disability benefits: the marginal utility of income and counterfactual earnings.[16]

Fact 1: On average, more-severe households are worse off than less-severe households by all measures, health and non-health, but there is substantial variation. Figure 2a shows that average consumption for households with more-severe conditions is just one-half that of households with less-severe conditions. This pattern is consistent across all of the measures that we observe: Having a more-severe health condition is strongly associated with lower living standards, lower earnings, and more adverse non-health events, even after controlling for proxies for ability group. Yet these strong average tendencies mask substantial variation within each severity group. For example, Figure 2a also shows that the bottom decile of households with less-severe conditions have consumption significantly below the mean among households with more-severe conditions, while the top decile of households with more-severe conditions have consumption above the mean among households with less-severe conditions. More generally, across different measures, we find that although health is strongly associated with better circumstances, it is far from a perfect indicator.

Fact 2: Less-severe recipients (L-DI) are much worse off than less-severe non-recipients (L-NDI), and by many non-health measures are even worse

[16]For simplicity, in the main text we present raw comparisons without controls or error bars. Results with different sets of controls for ability and with confidence intervals are presented in Appendix Figure F1. Within-ability comparisons are relevant for the insurance and welfare values of disability insurance, since across-ability insurance or redistribution can be accomplished (or offset) by the income tax and transfer system. Controlling for measures of ability tends to reduce the differences between recipients and non-recipients, but the differences remain large and the key conclusions are unchanged. Appendix Figure G1 compares education level, our main measure of ability group, across USDP-by-severity groups. Appendix Figures J1 and J2 show the main results for SSDI only (excluding SSI). The results for SSDI are similar to the main results, with the exception that more-severe non-(SSDI)recipients look worse off because they include a substantial number of SSI recipients. Sample sizes for SSI-only are too small to generate precise estimates.

off than more-severe recipients (M-DI). Figure 2a shows that average consumption among L-DI is just over one-half that of L-NDI. This pattern is consistent across all of the measures that we observe: L-DI are substantially worse off than L-NDI, with substantial differences even after controlling for proxies for ability group (Appendix Figure F1). In fact, L-DI are so strongly selected that on non-health dimensions they are comparable to, and even sometimes worse off than, M-DI. Figure 2a shows that L-DI and M-DI have similar current consumption levels. To avoid measuring the causal effect of receiving disability benefits, we also go back to the year before each recipient started receiving benefits and find that L-DI and M-DI are equally likely to be at the bottom of the consumption distribution prior to receiving benefits (see Appendix Figure G2).[17]

In SSA administrative data, we find that L-DI experience adverse non-health events at rates that are similar to or higher than those of M-DI. Figure 2b shows that L-DI are more likely than M-DI to have experienced a mass layoff prior to receiving disability benefits. The mass layoff rate is 10.2% for L-DI and 7.3% for M-DI in the three years before entry, and this difference remains even after conditioning on working.[18] Figure 2b, on adverse financial events in SSA administrative data linked to nationwide financial records, shows a similar pattern. In the years prior to receiving disability benefits, L-DI are more likely than M-DI to have experienced bankruptcy (3.5% vs. 2.6%), foreclosure (3.8% vs. 3.2%), and eviction (4.7% vs. 3.8%).[19] PSID measures of adverse non-health events reinforce the results from the SSA administrative data: L-DI and M-DI in the PSID experience similarly high rates of involuntary job loss, distressed moves, and divorce in the four years prior to receiving disability benefits (Appendix Figure G3). Overall, the likelihood of experiencing any of these events in the four years prior to receiving disability benefits is 57% for L-DI and 50%

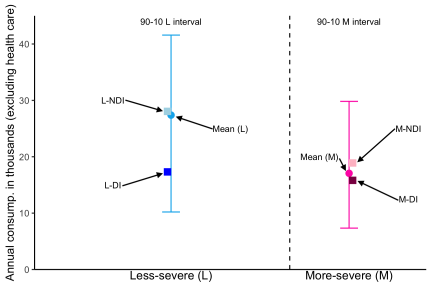
[17] To be sure, consumption and the other measures we study are determined by both health and non-health shocks. Since L-DI are by definition better off on health than M-DI, the fact that L-DI look similar to or worse than M-DI on these measures suggests that non-health shocks are worse for L-DI.

[18] Overall, mass layoff rates for disability recipients at entry are about twice those of non-recipients at the same ages (see Appendix Figure G4).

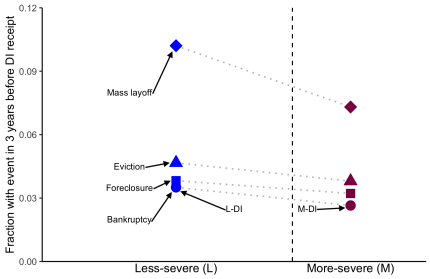
[19] Appendix Figure G5 breaks down the rates of adverse financial events into finer groups: initial allowance at Step 3 (meets medical listings), initial allowance at Step 5 (based on vocational grid), allowed on appeal, and never allowed.

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

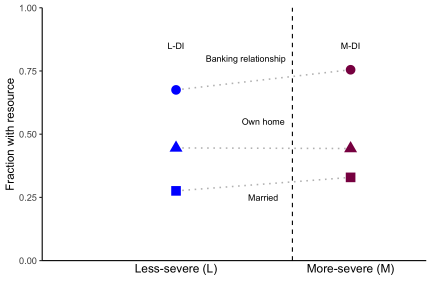
(a) PSID: Consumption



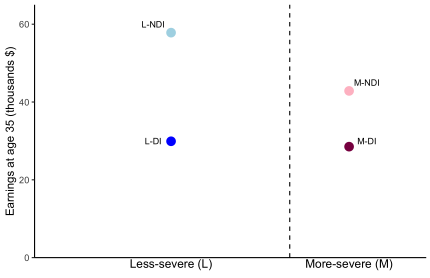
(b) Admin data: Prior to USDP receipt



(c) PSID: Outside resources



(d) PSID: Earnings at age 35

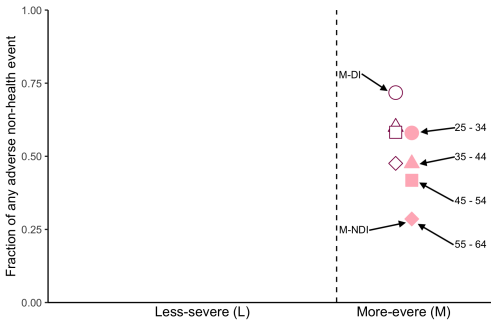


Notes: Figures 2a, 2c, and 2d present data from the PSID. Figure 2a presents statistics on consumption in the 2017 PSID for less-severe non-USDP-recipients (6,312 L-NDI), more-severe non-recipients (250 M-NDI), less-severe recipients (443 L-DI), and more-severe recipients (316 M-DI). Household consumption excludes health care and is divided by the square root of household size. Figure 2c 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 USDP benefits before age 36 and can be tracked back to age 35: 1,923 L-NDI, 90 M-NDI, 93 M-DI, and 152 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 2b presents rates of adverse financial events in the three years before receiving USDP for individuals who receive USDP benefits in SSA administrative data. The mass layoff sample (510,000 L-DI, 1.4M M-DI) is USDP recipients in the 831 records between 1990 and 2016 that appear in the Continuous Work History (CWHS). The bankruptcy sample is USDP recipients in the 831 files between 1995 and 2009 (4.9M L-DI, 8.4M M-DI). The foreclosure sample is approved USDP applicants in the 831 files between 2005 and 2014 who are homeowners (700,000 L-DI, 1.4M M-DI). The eviction sample is approved USDP applicants in the 831 files between 2005 and 2014 who are not homeowners (590,000 L-DI, 1.5M M-DI). "More-severe" in Figure 2b indicates recipients allowed at the initial state DDS level, and "less-severe" indicates recipients allowed upon appeal.

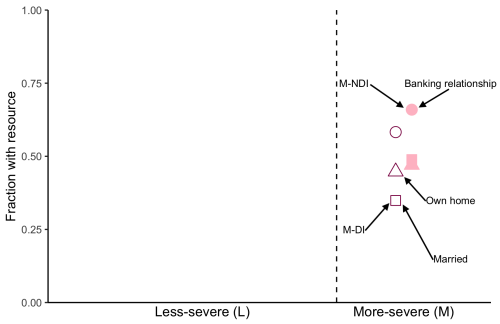
for M-DI.[20]

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

(a) PSID: Any adverse life event



(b) PSID: Resources at more-severe onset



Notes: Figure 3a 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, 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 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 USDP at that age. The sample sizes for age groups 25–34, 35–44, 45–54, and 55–64 are 896, 1199, 1054, and 848 for M-NDI; and 97, 211, 243, and 348 for M-DI. 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). 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 (see Appendix Table I2). Figure 3b presents rates of marriage, banking relationship, and homeownership one year prior to the year of more-severe onset for M-NDI (207) and M-DI (240) (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.

In addition to adverse non-health events, we also analyze how L-DI are selected on resources available to cope with adverse events. As shown in Figure 2c, we find that L-DI are less likely than M-DI to be married (28% vs. 33%) or to have a banking relationship (68% vs. 75%), and are similarly likely to own a home (45% vs. 44%). To account for the possible causal effect of receiving disability benefits on resources, we go back to the year before disability receipt and find a similar pattern (Appendix

[20]These likelihoods far exceed the baseline level among non-recipients: at entry, disability recipients are 70% more likely than non-recipients of the same age to experience these adverse events (see Appendix Figure G6).

Figure G8a). Prior to disability receipt, L-DI are also less likely to have either private or public health insurance (67% vs. 75%, from Appendix Figure G8b), less likely to receive public transfers (37% vs. 40% for food stamps, from G8c), and less likely to receive help from relatives (23% vs. 25%, from G8d).[21]

These results suggest that L-DI are highly selected on one key determinant of the value of receiving disability benefits: marginal utility. What about the other key determinant, counterfactual earnings? Although we do not have a direct measure of counterfactual earnings, Figure 2d shows average earnings of the household head at age 35, well before most individuals start receiving disability benefits but after at least some health and non-health shocks are realized. Households who later become L-DI have head earnings at age 35 similar to the age-35 head earnings of households who later become M-DI.[22]

These results suggest that L-DI benefits are less costly than benefits to representative less-severe individuals would be. How do L-DI come to receive disability benefits in the first place? From the PSID, about 25% had a more-severe condition when they applied, even though they no longer have a more-severe condition. Another 15% received UI immediately before applying for disability benefits, suggesting that they may have turned to U.S. disability programs as a form of permanent insurance following a qualifying job loss. Another 40% experienced some observable shock (job loss, distressed move, or divorce) before applying for disability benefits but never received UI. The remaining 20% have no observable shock prior to applying for disability benefits but have low resources on average.[23]

Fact 3: More-severe non-recipients (M-NDI) are better off on non-health measures than both more-severe and less-severe recipients. Figure 2a shows that M-NDI consumption is on average slightly greater than M-DI consumption, though still well below L-NDI consumption. Figure 3a shows a similar pattern in the

[21]Relative to the baseline severity definition, the results are even stronger when using the broader "a lot + somewhat" definition and are similar when using the "physical + mental health" definition (see Appendix Figure G7).

[22]Because disability receipt tends to last many years, the value of USDP depends not only (or even primarily) on counterfactual earnings at the time of entry but on counterfactual earnings for many years thereafter. We discuss quasi-experimental evidence on this key issue in the normative analysis (Section 5). Research suggests that events like mass layoffs, which are more common for L-DI than M-DI, are associated with long-term earnings losses (e.g., Jacobson, LaLonde and Sullivan, 1993).

[23]From the 2014 SIPP linked to SSA data, among L-DI, 30% are allowed due to being deemed to meet medical eligibility criteria (Step 3 of the disability determination process), 41% due to being deemed to have no capacity for any work based on the vocational grid (Step 5), and 25% on appeal.

likelihood of experiencing an adverse non-health event (specifically, job loss, distressed move, or divorce) at different ages in the PSID. For each age range (25–34, 35–44, 45–54, 55–64), we plot on the "more-severe" side the rate of adverse non-health events for M-NDI at that age and for M-DI who begin receiving disability benefits in that age range at that age. We find that M-NDI are 10–30% less likely in every age range to have experienced an adverse non-health event.[24] Turning to outside resources, we find in Figure 3b that M-NDI are more advantaged than M-DI in the year of disability onset. M-NDI are more likely than M-DI to be married (49% vs. 35%), to have a banking relationship (66% vs. 58%), and to own a home (47% vs. 45%). With respect to counterfactual earnings, Figure 2d shows that the age-35 head earnings of households that later become M-NDI are almost 50% greater than those of households that later become M-DI, though still substantially lower than those of households that later become L-NDI.

These results suggest that were M-NDI to receive disability benefits, such benefits would be less valuable than benefits to representative more-severe individuals. Why do M-NDI not receive disability benefits? From the SIPP-SSA linked data, we find that about 45% of M-NDI can be explained by mistakes in the decision process: individuals with more-severe conditions who applied but were rejected by SSA on medical grounds.[25] The remaining 55% of M-NDI do not apply. This self-selection out of applying for disability benefits appears likely to insure non-health risk: Relative to M-NDI who apply, M-NDI who do not apply have higher educational achievement (3.0% vs. 11% college degrees), higher rates of marriage and homeownership (46% vs. 55%, and 52% vs. 61%), and lower rates of "very low" food security (23% vs. 17%). Still, a nontrivial share of M-NDI who do not apply are disadvantaged.

Interpretation and robustness. It is possible that selection into disability receipt within less-severe or more-severe individuals is at least partly based on health, rather than only non-health factors. In Appendix Table F1, we estimate differences in consumption, resources, and age-35 earnings between L-DI and L-NDI, and between M-DI and M-NDI, controlling for fine health measures. The results do not change

[24] Appendix Figures G9 and G10 show results for head and spouse job loss, distressed move, and divorce separately.

[25] From the 2014 SIPP linked to SSA data, among M-NDI who apply and are rejected, 12% are rejected due to being deemed to have no severe impairment (Step 2 of the disability determination process), 13% due to being deemed to have capacity for previous work (Step 4), and 67% due to being deemed to have capacity for any work based on the vocational grid (Step 5).

qualitatively; in some cases, they get stronger. In Appendix B, we also address potential concerns about the L-DI vs. M-DI comparison, including that households that experience serious non-health shocks might be more likely to incorrectly report having a more-severe health condition, and that disability recipients might be more likely to incorrectly report having a more-severe health condition. Our findings suggest that such potential issues likely, though not necessarily, work against our main conclusions.

5 Normative analysis: Estimating the value of U.S. disability programs and their mismatches

5.1 Theory: Value of receiving disability benefits in different states

Our goal is to estimate the ex ante net value of disability benefits in different states of the world: the extent to which their ex ante value exceeds their cost, including cost from the distortions created by disability benefits receipt.[26] Like all welfare analyses, ours relies on assumptions for mapping observable characteristics into unobservable welfare. In our case, the key assumptions are that individuals optimize and that there are no externalities beyond "fiscal externalities": effects of behavioral responses on net government spending. In the empirical implementation, discussed in Section 5.2, we also make assumptions about marginal utility that determine the value of transferring resources from some states to others and so the insurance value of disability benefits.

Ex ante value of disability benefits in states Ω_b: $EAWTP_\theta(\Omega_b)$. Ex ante "risk type" θ faces a given set of possible ex post states of the world, Ω_θ. We wish to assess the ex ante value to θ of disability benefits in states $\Omega_b \subseteq \Omega_\theta$ (e.g., states in which the individual is a non-recipient with a health shock). To first order, the *ex ante* value to type θ, in terms of income in states Ω_τ, of disability benefits in states Ω_b is the

[26] To simplify the exposition, we refer to "states of the world" ω rather than "state-times" ω_t, but the distributions, probabilities, and conditional means of interest are across state-times: across both lifetime states of the world and time periods within each lifetime state of the world, as in Chetty (2006).

$EAWTP_\theta(\Omega_b)$ that solves

$$\underbrace{p(\Omega_b|\theta)E(\lambda_\omega \times WTP_\omega|\theta,\Omega_b)}_{\text{MB of disability benefits in states } \Omega_b} = \underbrace{p(\Omega_\tau|\theta)E(\lambda_\omega|\theta,\Omega_\tau)EAWTP_\theta(\Omega_b)}_{\text{MC of giving up } EAWTP_\theta(\Omega_b) \text{ in states } \Omega_\tau}, \tag{1}$$

where $p(\cdot)$ is the probability operator, λ_ω is the marginal utility of income in state ω, WTP_ω is the *ex post* value of disability benefits in state ω, Ω_b refers to the event that the realized state ω is in the set of benefit-receiving states ($\omega \in \Omega_b$), and Ω_τ refers to the event that the realized state ω is in the set of benefit-receiving states ($\omega \in \Omega_\tau$). The left-hand-side is the ex ante marginal benefit of disability benefits in Ω_b states. The right-hand-side is the ex ante marginal cost of giving up $EAWTP_\theta(\Omega_b)$ worth of income in Ω_τ states. Solving for $EAWTP_\theta(\Omega_b)$ and rearranging yields

$$EAWTP_\theta(\Omega_b) = \underbrace{\frac{p(\Omega_b|\theta)}{p(\Omega_\tau|\theta)}}_{\pi} E\left[\underbrace{\left(\frac{\lambda_\omega}{E(\lambda_\omega|\theta,\Omega_\tau)}\right)}_{(1+M_\omega)} WTP_\omega \middle| \theta, \Omega_b\right]. \tag{2}$$

The first term, π, is the no moral hazard actuarially fair price of income in Ω_b states in terms of income in Ω_τ states. The term in the expectation that multiplies the ex post value of disability insurance, WTP_ω, is one plus the "markup" on transfers from Ω_τ to ω:

$$M_\omega \equiv \frac{\lambda_\omega - E(\lambda|\theta,\Omega_\tau)}{E(\lambda|\theta,\Omega_\tau)}. \tag{3}$$

The markup is the amount by which marginal utility in ω exceeds mean marginal utility in the states used to value disability insurance (Ω_τ), as a share of the latter. One dollar worth of expected income in ω is worth $\$(1+M_\omega)$ of expected income in Ω_τ states, so M_ω is the marginal insurance value of transferring resources across these states per dollar of such transfers.

Ex ante cost to the government of providing disability benefits in states Ω_b: $EAG_\theta(\Omega_b)$. The *ex post* cost to the government of the individual receiving disability benefits in state ω, G_ω, includes not only the cost of the disability benefit itself but also fiscal externalities on other aspects of the government budget from any induced behavioral responses, such as reductions in the net taxes paid by a recipient's household due to labor supply responses to receiving disability benefits. The *ex ante* expected cost to the government of type θ receiving disability benefits in states Ω_b,

per state of the world in which type θ pays associated taxes, is

$$EAG_\theta(\Omega_b) = \pi E(G_\omega|\theta, \Omega_b). \tag{4}$$

Ex ante surplus from disability benefits for type θ in states Ω_b: $s_\theta(\Omega_b)$. Ex ante surplus from disability benefits for type θ in states Ω_b, per state of the world in which type θ pays associated taxes, is

$$s_\theta(\Omega_b) = g_\theta \times EAWTP_\theta(\Omega_b) - EAG_\theta(\Omega_b), \tag{5}$$

where g_θ is the marginal cost to the government of providing \$1 worth of surplus to θ through the income tax and transfer system (the "inverse-optimum weight" in Hendren, 2020).[27] The first term, $g_\theta \times EAWTP_\theta(\Omega_b)$, is how much it would cost the government to provide $EAWTP_\theta(\Omega_b)$ of surplus to θ via income tax cuts in Ω_τ states. The second term, $EAG_\theta(\Omega_b)$, is how much it would cost the government to provide $EAWTP_\theta(\Omega_b)$ of surplus to θ via disability benefits in Ω_b states. The difference, $s_\theta(\Omega_b)$, therefore measures the surplus from disability benefits for type θ in states Ω_b in terms of net government revenue. It answers: How much net revenue would the government raise from the combination of (i) providing disability benefits to type θ in states Ω_b and (ii) raising taxes on θ exactly enough to leave θ indifferent between this hypothetical combined disability-insurance-and-tax policy and the status quo? The surplus is positive if and only if it is cheaper for the government to help θ by providing disability benefits in Ω_b states than by cutting θ's income taxes in Ω_τ states. The surplus is greater when: the value of disability benefits in Ω_b states is greater (greater $EAWTP$), the efficiency cost of disability benefits in Ω_b states is smaller (smaller EAG), and the efficiency gain from cutting income taxes in Ω_τ states is smaller (greater g_θ).[28]

Overall mean surplus to all types: $s(\{\Omega_b(\theta)\}_{\theta\in\Theta})$. The overall mean social surplus from each type θ receiving disability benefits in states $\{\Omega_b(\theta)\}_{\theta\in\Theta}$, in terms

[27] The inverse-optimum weight, g_θ, is the marginal rate of transformation from surplus to θ to net revenue to the government through the income tax system. An income tax cut that makes θ \$1 better off costs the government $\$g_\theta$ of net revenue, and an income tax hike that makes θ \$1 worse off brings the government $\$g_\theta$ of net revenue. Inverse-optimum weights are the social planning weights that would rationalize the status quo tax schedule as optimal.

[28] The ratio of the value of disability benefits to those of a cost-equivalent tax cut is $EAWTP_\theta(\Omega_b)/(MVPF_\theta(\text{tax cut}) \times EAG_\theta(\Omega_b)) = g_\theta \times EAWTP_\theta(\Omega_b)/EAG_\theta(\Omega_b) = (s_\theta(\Omega_b) + EAG_\theta(\Omega_b))/EAG_\theta(\Omega_b)$.

of net government revenue per taxpaying state, is the mean surplus across types,

$$s(\{\Omega_b(\theta)\}_{\theta\in\Theta}) = E_\theta[s_\theta(\Omega_b(\theta))]. \tag{6}$$

5.2 Empirical implementation

Ex post value and cost. Our aim is to quantify, for different risk types θ, the ex ante surplus from disability benefits, $s_\theta(\Omega_b)$ from equation (5), in each of four sets of states of the world: $\Omega_b = \Omega_{h\text{-}di} \in \{$more-severe recipient (M-DI), less-severe recipient (L-DI), more-severe non-recipient (M-NDI), less-severe non-recipient (L-NDI)$\}$. The key building blocks are the ex post value and ex post cost of disability benefits in different states.

We assume that the ex post value of disability benefits in state ω is

$$WTP_\omega = \mathbb{1}(z_\omega^{DI=0} \leq \bar{z}) \times b, \tag{7}$$

where $z_\omega^{DI=0}$ is counterfactual earnings in state ω, $\bar{z}$ is the USDP earnings limit, and b is the USDP benefit. The indicator function indicates whether earnings in state ω would exceed the earnings limit if, counterfactually, the individual did not receive disability benefits and so was not subject to the earnings limit. This is a conservative first-order approximation that tends to understate the value of a disability benefit-like expansion of the budget constraint.[29] The earnings limit, $\bar{z}$, is the annualized value of the Social Security Administration's earnings limit for SSDI recipients ($1,260 per month in 2020, or $15,120 per year).

As a baseline, the benefit, b, is $13,000, roughly the average annual cash benefit of disability recipients in recent years. This follows the literature on the welfare effects of USDP in excluding the health insurance component of disability benefits as a baseline given the likelihood that for many recipients it mainly displaces other forms

[29]For simplicity and because of measurement issues with determining SSDI and SSI eligibility for non-recipients in the PSID, we quantify the value of a SSDI-like expansion of the budget constraint rather than modeling SSI and SSDI separately. Equation (7) is a conservative approximation to this value in that it assumes that the ex post value of disability benefits is zero in states in which counterfactual earnings exceed the earnings limit $\bar{z}$, no matter how small the excess, whereas in reality someone whose counterfactual earnings were up to $(\bar{z} + b)$ could work less *and* consume more with disability benefits—and so potentially gain significantly. This assumption means that extended disability receipt in "lifetime states of the world" in which counterfactual earnings exceed the earnings limit is extremely socially costly on net, since it is all cost and no benefit, year after year.

of subsidized health care cost sharing they otherwise would have received (see, e.g., Liebman, 2015). In Appendix Section C, we consider a wide range of assumptions about health insurance, including ones in which the health insurance component is more valuable in states in which the individual has a more-severe health condition. We find in those cases that the health insurance component tends to modestly reduce the value of L-DI benefits relative to M-DI benefits (see Appendix Table F3).[30]

To be sure, this first-order approximation to the value of disability *benefits* does not account for any associated hassle or stigma costs. Moreover, this analysis of the value of disability benefits in different states of the world does not speak directly to specific policy reforms, since policies are limited by information and incentive compatibility constraints in the extent to which they can target particular states. This analysis does not capture, for example, application responses to policy reforms, which are potentially important for the costs of such reforms. In Appendix Section D, we consider three frequently-discussed policy reforms taking into account the limited evidence on application responses: limiting receipt by less-severe individuals, decreasing benefit levels, and decreasing the allowance rate.

The ex post cost to the government of the individual receiving disability benefits in state ω is

$$G_\omega = b + \Delta_\omega, \tag{8}$$

where Δ_ω is the increase in non-disability net government spending due to behavioral responses to receiving disability benefits in state ω, such as losses in net tax revenue from earnings responses and changes in transfers from other programs.[31] We assume

[30]Health insurance is a much smaller driver of the net value of USDP than of gross government expenditures on USDP in part because it 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. 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). For this reason and others, most papers on the value of USDP exclude health insurance, including Bound et al. (2004), Chandra and Samwick (2009), Low and Pistaferri (2015), Cabral and Cullen (2019), and Meyer and Mok (2019). We find that under plausible assumptions, health insurance decreases the ratio of the per-recipient surplus of L-DI to M-DI benefits from its baseline value of 0.78 to 0.74–0.77 (see Appendix Table F3).

[31]Together, these definitions of ex post value (equation (7)) and ex post cost (equation (8)) imply that the fiscal externality FE—the increase in government costs due to behavioral responses as a share of the cost were there no behavioral responses (the "mechanical effect" ME)—is $FE(\Omega_b) = \frac{EAG_\theta(\Omega_b)}{ME_\theta(\Omega_b)} - 1 = \frac{[1-p(z_\omega^{DI=0} \leq \bar{z}|\theta,\Omega_b)]b + E(\Delta_\omega|\theta,\Omega_b)}{p(z_\omega^{DI=0} \leq \bar{z}|\theta,\Omega_b)b} - 1$.

that Δ_ω is 20% of the causal reduction in earnings from receiving disability benefits in state ω, $\Delta_\omega = 0.2 \times (z_\omega^{DI=0} - z_\omega^{DI=1})$. This is an estimate of the reduction in income tax revenue due to labor supply responses to receiving disability benefits, based on Hendren and Sprung-Keyser's (2020) estimate of the average marginal net tax and transfer rate on the reduction in earnings from receiving disability benefits of 20%.[32]

Key determinants. The key determinants of both the ex post value and the ex post cost of disability benefits are counterfactual earnings with and without disability benefits ($z_\omega^{DI=1}$ and $z_\omega^{DI=0}$). To construct counterfactual no-benefit earnings ($z_\omega^{DI=0}$) for individuals who actually receive disability benefits, we use French and Song's (2014) quasi-experimental estimates of the effects of disability benefits on the earnings of individuals with different types of health conditions three years after the disability decision. We assume that counterfactual no-benefit earnings for individuals who receive disability benefits are observed (with-benefit) earnings from the PSID plus French and Song's (2014) estimated mean reduction in earnings due to disability benefits by detailed health condition. We assume that counterfactual *with-benefit* earnings ($z^{DI=1}$) for individuals who do not receive disability benefits are the smaller of predicted counterfactual earnings based on French and Song (2014) and the SSDI earnings limit. It could be that these estimates of the effect of *being awarded* disability benefits understate the full causal effect of USDP on earnings—for example, if the long application process reduces the later earnings of individuals who are rejected (Parsons, 1991; Maestas, Mullen and Strand, 2013). It also could be that the long-term effects of disability benefits are larger than the short-term effects, especially if the non-health shocks experienced by less-severe recipients are temporary in nature. For these reasons, we test the robustness of our conclusions to much larger effects of disability benefits on earnings, for both more-severe and less-severe recipients, and find that our key conclusions are robust (see Appendix Table F4).[33]

[32] In principle, Δ_ω includes changes in transfers from other programs as well, such as SNAP and TANF. In practice, Bound et al. (2004) find that the change in other program expenditures from marginal disability applications is less than 1% of the change in tax revenue from the earnings response.

[33] One consideration with using these estimates is heterogeneity in the earnings responses to disability benefits. French and Song (2014) estimate the (within-detailed health category) local average treatment effects of being awarded disability benefits on earnings among applicants on the margin of program entry. Applying these estimates to all states likely leads us to overstate earnings responses in states in which the individual receives disability benefits (since in many the individual is inframarginal), which tends to decrease the implied value of USDP. Another consideration is the time horizon. Receipt of disability benefits often lasts many years, so they are a costly way to insure

The key determinant of the insurance value of disability insurance is marginal utility (λ_ω). We assume that marginal utility is a power function of non-health consumption, $\lambda = c^{-\gamma}$ (i.e., constant relative risk aversion utility), where γ is the coefficient of relative risk aversion. We use PSID data on total household consumption excluding health care, since including health care consumption could artificially inflate the consumption level of households with more-severe health conditions. We adjust for household size by dividing household consumption by the square root of household size. We also impose an annual consumption floor at the 10th percentile of the consumption distribution. Using observed consumption tends to overstate the counterfactual no-benefit consumption of disability recipients (since it reflects disability benefits), and so to understate their no-benefit marginal utility and the ex ante value of disability benefits in that state. We measure the value of disability benefits to the household in terms of income in states in which they do not receive benefits, $\Omega_\tau = \Omega_{NDI}$.[34]

Our baseline specification uses a coefficient of relative risk aversion of two and state-independent marginal utility. The functional form of marginal utility plays a crucial role in determining the value of transferring resources from some states to others and so in the insurance value of disability insurance. The assumed level of risk aversion is critical because it determines the magnitude of the gap in marginal utilities, and so the insurance value of transfers, for any given gap in consumption. For this reason, we test robustness to different levels of risk aversion and to state-dependent utility and find that our key conclusions are robust (see Appendix Table F5).[35]

transitory shocks. French and Song (2014) estimate counterfactual earnings for ten years after the disability award decision and find that earnings actually decline in later years after reaching a maximum of about $5,000 per year around year five. This trajectory is in keeping with evidence from Jacobson, LaLonde and Sullivan (1993) and Sullivan and Von Wachter (2009) that displaced workers experience long-term earnings losses. Still, we probe the robustness of our estimates and find a positive surplus from disability benefits in less-severe recipient (L-DI) states for earnings reductions up to four times those estimated by French and Song (2014). Our results would be qualitatively similar using estimates from Maestas, Mullen and Strand (2013), but they estimate only employment effects (not earnings effects) by subgroup.

[34]The assumption that utility is a state-independent, constant relative risk aversion function of measured consumption expenditures is standard in the literature (e.g., Meyer and Mok, 2019). Dividing household consumption by the square root of household size provides a measure of the household's living standard that reflects economies of scale in consumption. Imposing a consumption floor protects against mis-measured low consumption values unduly affecting the results given the sensitivity of marginal utility to low consumption.

[35]As Chetty and Looney (2006) emphasize, a small consumption gap is consistent with a large marginal insurance value if the consumption gap is small because risk aversion is large (or vice

Risk types. As a baseline, we assume that an individual's risk type is determined by their education: $\theta \in \{$high school dropout, high school graduate, some college, college plus$\}$. The conceptual experiment is that someone ex ante, knowing only what education group they will belong to, draws their state of the world from the cross-sectional distribution of households with a working-age head and with that education level. In other words, we take the cross-sectional distribution of households of a given education level to represent the distribution of possible ex post states of the world that someone with that education level faces ex ante. Ideally, the risk type would both isolate risk from heterogeneity and separate the ex ante population into groups with different ex ante earning ability, since the income tax and transfer system can redistribute across groups with different earnings levels. We consider other approaches for defining risk types and find that the key results are robust to plausible alternatives (see Appendix Table H2). Each risk type θ's inverse optimum weight, g_θ, is the within-θ average of Hendren's (2020) baseline estimates of inverse optimum weights by household income centile.

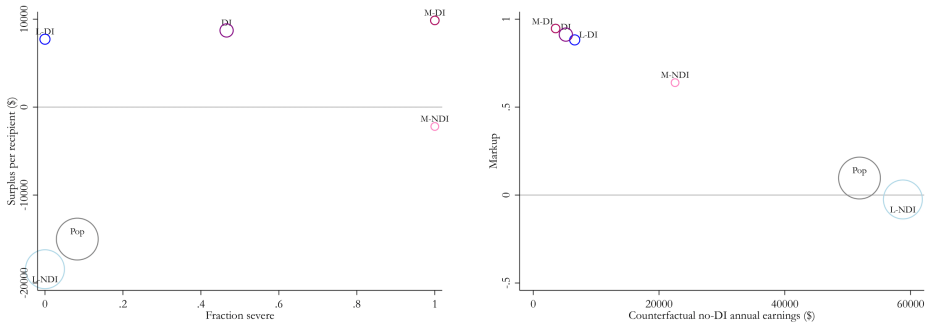
5.3 Empirical estimates of the value of disability benefits in different states

Figure 4a shows the empirical analogue of the theoretical Figure 1, and Table 1 reports related estimates. The results reveal two key findings. First, U.S. disability programs generate substantial ex ante surplus. We estimate that disability benefits as a whole are 64% more valuable than tax cuts with the same cost to the government would be. As a result, the annual ex ante surplus from USDP, measured in terms of government revenue, is $920 per household or $81 billion in aggregate (based on there being about 88 million working-age households in the U.S.). This means that were the government to abolish USDP, the tax cuts necessary to leave individuals unharmed relative to the status quo would cost the government $920 per household per year more than what it would save from abolishing USDP. Second, USDP mismatches with respect to health are not costly—they are highly valuable. We estimate that mismatches produce an

versa). The evidence on state-dependence of marginal utility of consumption is mixed. For example, Finkelstein, Luttmer and Notowidigdo (2013) conclude that marginal utility is lower when health is worse, whereas Lillard and Weiss (1997) conclude that it is higher and Brown, Goda and McGarry (2016) find substantial heterogeneity across ages and disability types. We consider both directions in our robustness calculations.

Figure 4: Surplus and its determinants: markup and counterfactual earnings

(a) Surplus per recipient household (b) Markup and counterfactual earnings



Notes: Figure 4a, the empirical analogue of Figure 1, plots surplus per recipient household, $p(\Omega_\tau)s/p(\Omega_b)$, against the fraction more-severe for the following sets of states of the world, Ω_b: all states (Pop); states in which the household receives USDP benefits (DI), including when more-severe (M-DI) and less-severe (L-DI); and states in which the household does not receive USDP benefits (NDI), including when more-severe (M-NDI) and less-severe (L-NDI). Surplus s, defined in equation (6), is in units of government revenue per Ω_τ state, so $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient household. Figure 4b plots the mean markup on transfers from non-recipient states (NDI) to different sets of states Ω_b, $E(M_\omega|\Omega_b)$, against mean counterfactual no-benefit annual earnings in those states. 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 USDP awards on earnings for DI states, as described in the text. In both panels, marker areas are proportional to population shares and monetary amounts are in 2016 dollars per year. Sample sizes: 313 M-DI, 438 L-DI, 248 M-NDI, 6259 L-NDI.

annual ex ante surplus, in terms of government revenue, of \$510 per household or \$45 billion in aggregate. This is mostly attributable to benefits in less-severe states: Such benefits are worth 57% more than cost-equivalent tax cuts and produce ex ante surplus of \$440 per household. In fact, benefits in less-severe states (L-DI) appear to be nearly as valuable as benefits in more-severe states (M-DI).

That USDP mismatches are valuable rather than costly reflects strong selection into disability receipt conditional on health. On average, health is a strong indicator of the value of receiving disability benefits: Receiving disability benefits in the average more-severe state generates substantial surplus (\$5,000 on average), whereas receiving disability benefits in the average less-severe state reduces surplus significantly (−\$16,800 on average). But USDP mismatches are highly-selected subsets of their respective severity groups. The ex post states of the world in which an individual would receive disability benefits despite not having a more-severe condition

Table 1: Value of disability benefits in different states

	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$)
DI	0.91	5,191	924	8,718	1.64
M-DI	0.95	3,557	488	9,858	1.73
L-DI	0.88	6,623	436	7,719	1.57
NDI	0.00	57,476	-15,930	-17,819	0.19
M-NDI	0.64	22,591	-74	-2,203	0.86
L-NDI	-0.02	58,838	-15,856	-18,429	0.17
All states	0.10	51,932	-15,005	-15,005	0.29
M	0.82	11,252	414	4,982	1.34
L	0.03	55,620	-15,420	-16,818	0.23

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 benefits (DI), including when more-severe (M-DI) and less-severe (L-DI); those in which the household does not receive USDP benefits (NDI), including when more-severe (M-NDI) and less-severe (L-NDI); and all states (All states), including when more-severe (M) and less-severe (L). 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 USDP awards on earnings for DI states, as described in the text. Surplus s, defined in equation (6), is in units of government revenue per Ω_τ 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 28. Monetary amounts are in 2016 dollars per year. Sample sizes: 313 M-DI, 438 L-DI, 248 M-NDI, 6259 L-NDI, 561 M, 6697 L.

(L-DI) have a large surplus from disability benefits ($7,700 on average), while the ex post states of the world in which an individual would not receive disability benefits despite having a more-severe condition (M-NDI) have a negative surplus from disability benefits (−$2,200 on average). That USDP mismatches are so favorably selected on the value of receiving disability benefits is in keeping with the findings from the positive analysis that USDP mismatches are favorably selected on a wide variety of non-health factors.

Figure 4b decomposes the selection on value into its two key determinants: marginal utility and counterfactual earnings. USDP recipients as a whole, and less-severe recipients (L-DI) in particular, are highly selected in terms of both marginal utility and counterfactual earnings, and so in terms of both the insurance benefit and distortion cost of receiving disability benefits. The markup on transfers from non-recipients to L-DI states is 0.88, similar to that on transfers to M-DI states of 0.95. Mean counterfactual annual earnings in L-DI states is $6,600, versus $3,600 in M-DI states

and $57,500 in NDI states. Hence, the large surplus from benefits in L-DI states is driven not only by low consumption but also by low counterfactual earnings in those states. Strong selection on counterfactual earnings is crucial because of the earnings limit, which involves large private and fiscal externality costs in high-earning states. Selection into USDP receipt on both marginal utility and counterfactual earnings is so strong that awards to both M-DI and L-DI would generate positive surplus even if markups were just 25 percent of their estimated values (Appendix Table F5) or earnings responses were four times French and Song's (2014) estimates (Appendix Table F4).

The main conclusions that U.S. disability programs and their mismatches with respect to health are highly valuable hold not only overall, averaging across the different risk (education) types, but also for each risk type individually (see Appendix Table F2). They also hold under a wide range of assumptions about the key ingredients of the calculations, including the value and cost of the health insurance component (Appendix Table F3); earnings responses to receiving benefits (Appendix Table F4); risk aversion and state-dependence of utility (Appendix Table F5); the definition of severity (Appendix Table H1); and the definition of risk types (Appendix Table H2). For example, state-dependent utility such that the marginal utility of a given level of consumption is 25% lower when more-severe (based roughly on Finkelstein, Luttmer and Notowidigdo's (2013) estimates) decreases the surplus from M-DI benefits by 56% and increases the surplus from L-DI benefits by 5%. State-dependent utility in the opposite direction (marginal utility 25% higher when more-severe) increases the surplus from M-DI benefits by 54%, decreases the surplus from L-DI benefits by 5%, and causes exclusion errors to become modestly costly (counterfactual surplus per recipient of $940).[36] L-DI benefits remain highly valuable even excluding those L-DI who had a more-severe condition when first receiving DI (which reduces the aggregate surplus from L-DI benefits by 26%) or at any time up to and including when first receiving DI (conditional on observing the individual prior to DI receipt; this reduces the aggregate surplus by 40%). Including the health insurance component tends to increase the surplus from both M-DI and L-DI benefits, somewhat more for M-DI than L-DI given their greater consumption of health care. As a result, the ratio of

[36]The marginal utility of more-severe individuals (modestly) affects the surplus from benefits to *less*-severe individuals by affecting the marginal utility of non-recipients (because of the presence of more-severe non-recipients) and so the markup on transfers from non-recipients to L-DI.

the per-recipient surplus of L-DI to M-DI benefits falls from its baseline value of 0.78 to 0.74–0.77 when using our main assumptions about the health insurance component and to 0.54–0.71 when using extreme, ratio-minimizing assumptions across the board (see Appendix Section C for more details).

5.4 Key drivers: Non-health risk and selective applications

Non-health risk as a driver: About one-half of the value of USDP comes from insuring risks other than that of having a more-severe health condition. We investigate the connections between mismatches with respect to health and insurance against health and non-health risk with a decomposition analysis. 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 ω into the sum of (i) the mean benefit received in states in the same health category as ω, $h_\omega \in \{\text{L, M}\}$, 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 Ω_b):

$$b_\omega = \underbrace{E(b|h = h_\omega)}_{\text{Health-contingent benefit}} + \underbrace{[b_\omega - E(b|h = h_\omega)]}_{\text{Within-health transfer}}. \tag{9}$$

With this decomposition, the markup on transfers from all of θ's states (Ω_θ) to states $\Omega_b \subseteq \Omega_\theta$ can be decomposed as

$$M_\Omega(\Omega_b) = \underbrace{Cov_h\left[\frac{p(\Omega_b|h)}{p(\Omega_b)}, E\left(\widehat{\lambda}|h\right)\right]}_{\text{Insurance against health risk}} + \underbrace{E_h\left\{\frac{p(\Omega_b|h)}{p(\Omega_b)}[1 - p(\Omega_b|h)]\left[E\left(\widehat{\lambda}|\Omega_b, h\right) - E\left(\widehat{\lambda}|\sim\Omega_b, h\right)\right]\right\}}_{\text{Insurance against non-health risk}}, \tag{10}$$

where $\widehat{\lambda} \equiv \frac{\lambda}{E(\lambda)}$ is marginal utility normalized to have unit mean in the population, and the conditioning on θ has been suppressed for notational ease. Details of the derivation are in Appendix Section E.[37]

The overall markup, $M_\Omega(\Omega_b)$, is the sum of two components. The "insurance against health risk" component, from the health-contingent benefit in equation (9), is

[37] We calculate the markups on transfers within each education group and take the mean across education groups. These markups measure value in terms of income in all states of the world, $M_\Omega(\Omega_b;\theta) \equiv \frac{E(\lambda|\theta,\Omega_b) - E(\lambda|\theta)}{E(\lambda|\theta)}$, rather than in terms of income in NDI states. We do this because it simplifies equation (10) and facilitates comparison across different groups of benefits. It also leads to somewhat smaller overall markups.

from insuring the risk of having a more-severe health condition ("health risk"). The more concentrated are benefits in states in which the individual has a more-severe health condition (greater $p(\Omega_b|M)$ relative to $p(\Omega_b|L)$), the greater the insurance of health risk. This component is necessarily reduced by mismatches with respect to health, which by definition reduce the targeting of more-severe states. The "insurance against non-health risk" component, from the within-health transfer in equation (9), is from insuring "non-health risk," i.e., risk *within* health categories.[38] The greater the marginal utility gap between recipient and non-recipient states *within* more-severe and less-severe states (i.e., greater $E\left(\widehat{\lambda}|\Omega_b, h\right) - E\left(\widehat{\lambda}| \sim \Omega_b, h\right)$), the more valuable the insurance of non-health risk. This component is zero in the absence of mismatches. Mismatches make this component positive (negative) if appropriately-weighted selection into receiving disability benefits on marginal utility conditional on health is positive (negative).

Table 2 shows the results. The insurance that USDP provide against the risk of having a more-severe health condition is valuable, contributing 28 percentage points to the total markup. But the insurance that USDP provide against other risks contributes even more: 48 percentage points, 63% of the total markup. Even when we use additional health and disability measures to define finer health categories in equation (10), we still find that at least one-half of the insurance value of USDP comes from insuring risks *within* states with the same realizations of measured health (see Appendix Table F6).[39] These results, which are in keeping with the earlier result that roughly half of the ex ante surplus from USDP comes from mismatches, suggest that a majority of the overall insurance value of USDP comes from insuring risks beyond that of having a more-severe health condition.

Mismatches with respect to health drive this insurance of non-health risk, albeit at the expense of necessarily reducing insurance against the risk of having a more-severe health condition. Compared to a hypothetical, infeasible program that would

[38]We term this component "non-health risk" partly as a shorthand for "within-health category risk" and partly because the evidence suggests that much of this component is driven by factors that would not ordinarily be called "health," such as labor market shocks. But it is important to keep in mind that this component is technically within-health category risk, which both includes some health risk (e.g., some states in a given health category have higher out-of-pocket health spending than others) and excludes some non-health risk (to the extent that a risk is positively associated with having a more-severe health condition, part of it is included in the "health risk" component).

[39]For example, using four or six severity categories rather than two decreases the within-health share from 63% to 56%. Using the 30 categories defined by interacting six severity categories with five self-reported health categories decreases the within-health share to 51%.

Table 2: Decomposition of the markup on transfers into different states

	Surplus per recip ($)	Markup,	Contrib from insuring:		Share from insuring:	
	$(p(\Omega_\tau)s/p(\Omega_b))$	$E(M_\Omega)$	Health	Non-health	Health	Non-health
DI	8,718	0.75	0.28	0.48	37%	63%
M-DI	9,858	0.78	0.66	0.12	85%	15%
L-DI	7,719	0.73	-0.06	0.79	-8%	108%
M-NDI	-2,203	0.48	0.66	-0.18	136%	-36%
Health-DI	4,982	0.66	0.66	0.00	100%	0%
Random-App-DI	-6,143	0.22	0.22	0.01	98%	2%
Earnings-Test-DI	5,857	0.62	0.34	0.28	55%	45%

Notes: Table presents social surplus per recipient, overall markup, and markup components and shares from insurance of health and non-health risk associated with disability benefits in each of several sets of states of the world, defined by the row. Contributions of insurance of health and non-health risk to overall markup are defined in equation (10). Shares of markup from insurance of each type of risk is the contribution of insurance of that type of risk to the overall markup as a share of the overall markup. The overall markup, $E(M_\Omega)$, is the across-risk group (education) average of the average within-risk group markup on transfers from all states (not just NDI states, as in the main analysis) to the states in the table row. Surplus s (from equation (6)) is in units of government revenue per Ω_τ state per year, so $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient per year. "Health-DI" is a hypothetical program that pays benefits if and only if the individual has a severe health condition. "Random-App-DI" is a hypothetical program whose award probabilities are a function of severity and age, as estimated by Low and Pistaferri (2015), and with no systematic selection into application, rather than the actual selective applications to DI. That Random-App-DI provides a small amount of insurance against non-health risk arises from the dependence of its award probabilities on age. "Earnings-Test-DI" is a hypothetical program whose award probabilities are a function of severity and age, as estimated by Low and Pistaferri (2015), and with no systematic selection into application (like Random-App-DI) *except* that the individual applies only in states in which counterfactual no-DI earnings are below two times the SSDI earnings test threshold. This is a simple model of the type of selection that an earnings test might produce. Sample sizes: 313 M-DI, 438 L-DI, 248 M-NDI, 561 Health-DI, 1060 Random-App-DI, 675 Earnings-Test-DI.

perfectly target states with more-severe health conditions ("Health-DI"), the across-health component of USDP is less than one-half as large (0.28 for USDP versus 0.66 for Health-DI). Yet USDP more than makes up for their lesser insurance of the risk of having a more-severe health condition by providing highly valuable insurance against other risks. Insurance of other risks contributes 0.48 to the USDP markup (whereas by definition Health-DI does not insure risks within health categories), elevating the overall markup and surplus per recipient of USDP above those of Health-DI (0.75 versus 0.66 and $8,700 versus $5,000, respectively).

Although both types of mismatches insure non-health risk, the vast majority of such insurance comes from L-DI benefits. The "L-DI" row of the table decomposes

the markup on transfers from all states to L-DI states, using the same approach used to decompose the overall markup (substituting "L-DI" for "DI" in equation (10)). L-DI benefits exacerbate health risk, by transferring from all states to lower-marginal utility L states, but provide highly valuable insurance against non-health risk, by transferring from L-NDI states to much higher-marginal utility L-DI states. The insurance against non-health risk provided by L-DI benefits accounts for over half of the overall USDP markup and 85% of the "insurance of non-health risk" markup component.

Selective applications as a driver: the insurance of non-health risk and overall value of USDP are driven primarily by strong selection into application (as opposed to selective awards given applications). We investigate the importance of selective applications by comparing USDP to a hypothetical disability program with no systematic selection into application ("Random-App-DI").[40] The results, shown in the second-to-last row of Table 2, suggest that selection into application plays a crucial role in driving the overall value of USDP. Whereas USDP generate $8,700 per recipient worth of annual surplus, Random-App-DI would generate negative surplus (−$6,100)—i.e., individuals would prefer cost-equivalent tax cuts to Random-App-DI. The difference is driven almost entirely by USDP insuring risks beyond that of having a more-severe health condition. The USDP markup exceeds the Random-App-DI markup by 53 percentage points, 47 percentage points (89%) of which are from within-health transfers. Selection into application is responsible for the vast majority of the insurance against non-health risk provided by USDP, which itself is highly important to the overall value of USDP. Of course, selective award decisions likely play a major role in shaping selection into application, so the importance of selection into application does not imply that selective award decisions are unimportant (e.g., it may be that many would-be applicants are deterred from applying by the belief that their application would not be approved). Rather, it means that the net effect of USDP award decisions and other features—including restrictions on applicants and recipients, awareness, hassle costs, and stigma—is strongly favorable

[40]To simulate the distribution of benefits under Random-App-DI, we use Low and Pistaferri's (2015) estimates of USDP award success probabilities, which are a function of an applicant's work limitation status and age, and assume that each applicant applies for Random-App-DI up to two times. Because actual USDP awards account for factors beyond those in Low and Pistaferri's (2015) model (e.g., through the vocational grid), this analysis does not isolate the effect of selective applications alone but the combined effect with richer award decisions as well.

selection into application that greatly increases the value of USDP.

Which of these features in particular explains the strong, value-enhancing selection into application for disability benefits? Unfortunately there is little in the way of compelling variation that would facilitate a direct empirical investigation of this important issue. But indirect evidence suggests that the stringent earnings limit could potentially play a key role. We simulate the effects of the type of selection that an earnings limit might produce by considering a hypothetical program ("Earnings-Test-DI") that is equivalent to Random-App-DI except that the individual applies only in states in which counterfactual no-benefit earnings are below two times the earnings limit.[41] The last row of Table 2 shows the results. Whereas Random-App-DI would be quite costly (surplus per recipient of −$6,100), Earnings-Test-DI would be quite valuable (surplus per recipient of $5,900). The difference arises because the assumed self-selection out of application by individuals in states with high counterfactual no-benefit earnings avoids the high-cost, low-value benefits in those states, and it greatly increases the extent to which disability benefits insure non-health risk. To be clear, these results provide suggestive evidence on the *potential* screening effects of an earnings limit alone and so do not isolate the *marginal* effect of the earnings limit on top of the many other tags and screens of U.S. disability programs.

5.5 Policy implications

Together, the findings from the positive and normative analyses paint a picture of U.S. disability programs that target well on the key determinants of the value of receiving disability benefits. Subject to the caveats that we have discussed, our results have important implications for some of the biggest questions in the debate over these programs.

[41] The results are qualitatively similar if the threshold for applying is one or three times the earnings limit instead. As with Random-App-DI, we report the results of a simulation in which individuals rejected once re-apply a second time, but the results are similar if individuals apply only once or up to three times. This simple model of selection may understate the screening effects of the earnings limit by ignoring dynamics. Given the substantial fixed costs to individuals with non-negligible earning opportunities of applying for disability benefits (especially the long application process during which earnings must be low), the earnings limit likely deters applications even in many states in which current no-benefit earnings are low but anticipated future earnings are higher. Strong selection into application on the basis of future no-benefit earnings is consistent with French and Song's (2014) finding that rejected applicants on the margin of program entry have low earnings throughout the ten-year period following their application.

Is the provision of benefits to individuals with less-severe conditions diluting the value of U.S. disability programs? Our results suggest that the benefits received by less-severe recipients are not diluting the value of USDP, at least not in terms of insurance and welfare, but rather providing highly valuable insurance against a variety of major risks. To the extent that more-severe recipients are considered deserving of disability benefits on the basis of the ex ante value of such benefits, less-severe recipients appear nearly as deserving. To be sure, our results do not imply that USDP should reduce efforts to target individuals who have severe health conditions or increase efforts to target based on non-health factors. The value of such changes to tags or screens depends on the induced application responses, the evidence on which is unfortunately sparse. It could be that efforts to target individuals with severe health conditions are an important driver of the strong selection into application on non-health risk that we find to be so crucial to the value of USDP.

Would proposed reforms to U.S. disability programs increase or decrease social welfare? We quantify the value of proposed reforms to U.S. disability programs, with details in Appendix Section D and results in Appendix Table F7. The results suggest that proposed reforms to decrease benefit levels, allowance rates, or awards to individuals with less-severe health conditions would each significantly decrease social surplus. Even a costless elimination of awards to individuals with less-severe health conditions would decrease social surplus. That is not to say, however, that *increasing* eligibility would increase social surplus, since we find that extreme selectivity is crucial to the value of USDP.[42] Of course, a distinct question, and one that this paper is not equipped to answer, is whether U.S. disability programs *should* be expanded. Such a normative judgment depends on considerations beyond insurance and welfare, including non-welfarist considerations about the appropriate role of disability programs or about the extent to which different individuals or states are deserving of benefits. Our results suggest that such considerations would have

[42] An additional consideration for the effects of disability insurance on social welfare is possible redistribution value. Throughout the paper, we have focused on efficiency—insurance of risk *within* risk types, setting aside any redistribution *across* risk types—since redistribution across risk types can be accomplished by the income tax and transfer system. But without offsetting changes in other policies, expanding USDP would redistribute from higher- to lower-ability groups, since lower-ability groups both are more likely to receive disability benefits and pay lower absolute amounts of the payroll taxes that fund USDP. Such redistribution would increase the overall social welfare value of USDP under social welfare functions that value redistribution more than the income tax and transfer implicitly does. We find that accounting for redistribution increases the value of USDP to someone behind the veil by about 30%.

to be fairly strong, in terms of their effects on generalized social marginal welfare weights (Saez and Stantcheva, 2016), in order to overturn our main conclusions. To overturn the conclusion that disability benefits to less-severe recipients are valuable would require that each dollar of such benefits be viewed as being worth less than 65 cents. To overturn the conclusion that disability benefits as a whole are valuable would require that each dollar of benefits to less-severe recipients be viewed as being worth less than 25 cents.

Are disability programs the *best* way to insure non-health risk? Our finding that U.S. disability programs provide cost-effective insurance against non-health risk does not necessarily mean that they are the *most* cost-effective way to do so. The earnings limit in particular makes disability benefits a costly way to insure risks that do not involve prolonged low earnings prospects. For individuals with high earnings capacity, insuring them through other means—perhaps by expanding UI or creating a new wage insurance program—could be valuable. But the available evidence suggests that the vast majority of disability recipients, including those with less-severe health conditions, would earn little even if not receiving disability benefits.[43] Moreover, expansions of alternative programs would have efficiency costs of their own and may lack the advantageous targeting properties that we find for USDP.[44]

Setting aside overall welfare effects, how might changes in other programs affect the value of U.S. disability programs? We provide suggestive evidence on this question in two ways. First, we estimate how the value of DI benefits varies with the generosity of other parts of the social safety net. As a rough proxy for the generosity of the rest of the safety net in each state in the U.S., we use the ratio of families receiving TANF to families living in poverty in that state, as calculated by Meyer and Floyd (2020).

[43] Only about 20–30% of SSDI recipients on the margin of allowance would earn at least the "substantial gainful activity" threshold (about $16,000 in 2021) if they did not receive SSDI (Maestas, Mullen and Strand, 2013; French and Song, 2014). This share would presumably be lower for inframarginal recipients and at higher earnings thresholds, though we are unaware of direct evidence on either.

[44] Acknowledging the substantial uncertainty about the fiscal externality costs of changes in USDP, our analysis of potential changes in USDP policy, reported in Appendix Table F7, suggests that USDP provide more cost-effective insurance than UI does. Our estimates of the ex ante MVPFs of each of three key dimensions of disability policy—paying benefits to individuals with less-severe conditions, changing benefit levels, and changing the award threshold—are about 1.42. This significantly exceeds Hendren and Sprung-Keyser's (2020) estimates of the ex ante MVPF of changes in UI (main estimate of 0.61 with range 0.53–0.74). The higher MVPFs of changes in USDP arise from reinforcing effects of higher insurance value markups (0.88–0.91 for USDP versus 0.20 for UI) and lower fiscal externality costs (0.34 for USDP versus 0.95 for UI).

Appendix Table F8 presents the value of DI and L-DI benefits for states in each quartile of this measure. In more generous states, the per-capita surplus from DI and L-DI benefits is lower, recipiency rates of DI and L-DI are lower, markups on DI and L-DI benefits are lower, and the counterfactual earnings of DI and L-DI recipients are higher. Although these differences could be driven by other factors, the results are in keeping with the idea that a more robust non-DI safety net would reduce the value of DI in general and of L-DI benefits in particular.

Second, we estimate how the value of disability benefits to less-severe individuals changes as certain subsets of less-severe individuals are excluded. If individuals with less-severe conditions who received UI prior to receiving disability benefits did not switch to disability—for example, due to a major UI extension—the estimated surplus from L-DI benefits would fall by 33%. Excluding individuals with less-severe conditions who experienced an observable job loss prior to receiving disability benefits but did not receive UI—as might occur under a major expansion of UI eligibility—would reduce the estimated surplus from L-DI benefits by an additional 20 percentage points. While only suggestive, these results indicate that USDP might continue to provide substantial insurance against non-health risk even if there were large expansions of programs designed to insure non-health risk.[45]

6 Conclusion

The public debate over disability insurance has centered on concerns about individuals with less-severe health conditions receiving benefits. In this paper, we go beyond health risk alone to quantify the insurance provided by U.S. disability programs against all risks, non-health as well as health. We find that not only are these programs highly valuable, but their mismatches with respect to health, including benefits to individuals with less-severe health conditions, actually increase their value. This is because disability recipients with less-severe conditions are a highly selected group: Compared to non-recipients, less-severe recipients are more likely to have ex-

[45] Of course, the changes in other programs required to have such an impact on disability enrollment might be large given the long-lasting nature of disability receipt, the persistently low earnings among individuals who apply for disability benefits, and the evidence that changes in UI have little effect on disability receipt (Mueller, Rothstein and Von Wachter, 2016). Note also that expansions of other programs would not necessarily decrease the value of USDP. They might even increase it by improving USDP targeting further still, as Low and Pistaferri (2015) find in their structural model.

perienced adverse non-health events, are less likely to have the resources to insure those events, and are more likely to have low earnings even if not for receiving disability benefits. In fact, recipients with less-severe conditions are so highly selected that they look similar to or worse off than recipients with more-severe conditions on these dimensions.

As a result of this strong selection, disability benefits to both more-severe and less-severe recipients produce insurance value substantially greater than distortion cost, and so are significantly more valuable ex ante than cost-equivalent tax cuts. Our results suggest that about one-half of the value of U.S. disability programs comes from insuring non-health risk. For this reason, reforms to increase the emphasis of these programs on health could significantly reduce their ex ante value.

The importance of non-health risk for the value of U.S. disability programs may be just one example of a broader phenomenon. No program exists in a vacuum: Its effects reflect the diversity of risks in the economy, how well insured those risks are by other programs and institutions, and how its tags and screens select on those risks. We find that U.S. disability programs insure risks well beyond health, and that this "incidental" role is central to their overall value. Other programs might be similar in having their costs and benefits driven in large part by factors outside of their core aims. The extent to which they do is an empirical question, one that future research could investigate using similar methods.

References

Alatas, Vivi, Ririn Purnamasari, Matthew Wai-Poi, Abhijit Banerjee, Benjamin A Olken, and Rema Hanna. 2016. "Self-Targeting: Evidence from A Field Experiment in Indonesia." *Journal of Political Economy*, 124(2): 371–427.

Autor, David H, and Mark G Duggan. 2003. "The Rise in the Disability Rolls and the Decline in Unemployment." *The Quarterly Journal of Economics*, 118(1): 157–206.

Autor, David H., and Mark G. Duggan. 2006. "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding." *Journal of Economic Perspectives*, 20(3): 71–96.

Benitez-Silva, Hugo, Moshe Buchinsky, and John Rust. 2004. "How Large Are the Classification Errors in the Social Security Disability Award Process?" National Bureau of Economic Research.

Bhargava, Saurabh, and Dayanand Manoli. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from An IRS Field Experiment." *American Economic Review*, 105(11): 3489–3529.

Black, Dan, Kermit Daniel, and Seth Sanders. 2002. "The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust." *American Economic Review*, 92(1): 27–50.

Bound, John, Julie Berry Cullen, Austin Nichols, and Lucie Schmidt. 2004. "The Welfare Implications of Increasing Disability Insurance Benefit Generosity." *Journal of Public Economics*, 88(12): 2487–2514.

Brown, Jeffrey R, Gopi Shah Goda, and Kathleen McGarry. 2016. "Heterogeneity in State-Dependent Utility: Evidence from Strategic Surveys." *Economic Inquiry*, 54(2): 847–861.

Cabral, Marika, and Mark R Cullen. 2019. "Estimating the Value of Public Insurance Using Complementary Private Insurance." *American Economic Journal: Economic Policy*, 11(3): 88–129.

Chandra, Amitabh, and Andrew A Samwick. 2009. "Disability Risk and the Value of Disability Insurance." In *Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly*. 295–336. University of Chicago Press.

Charles, Kerwin Kofi, Yiming Li, and Melvin Stephens Jr. 2018. "Disability Benefit Take-Up and Local Labor Market Conditions." *Review of Economics and Statistics*, 100(3): 416–423.

Chetty, Raj. 2006. "A General Formula for the Optimal Level of Social Insurance." *Journal of Public Economics*, 90(10–11): 1879–1901.

Chetty, Raj, and Adam Looney. 2006. "Consumption Smoothing and the Welfare Consequences of Social Insurance in Developing Economies." *Journal of Public Economics*, 90(12): 2351–2356.

Currie, Janet. 2006. "The Take-Up of Social Benefits." In *Public Policy and the Income Distribution*. , ed. Alan J. Auerbach, David Card and John M. Quigley, 80–148. New York, NY: Russell Sage Foundation Publications.

Deshpande, Manasi, and Yue Li. 2019. "Who is Screened Out? Application Costs and the Targeting of Disability programs." *American Economic Journal: Economic Policy*, 11(4): 213–48.

Deshpande, Manasi, Tal Gross, and Yalun Su. 2021. "Disability and Distress: The Effect of Disability Programs on Financial Outcomes." *American Economic Journal: Applied Economics*, 13(2): 151–78.

Diamond, Peter, and Eytan Sheshinski. 1995. "Economic Aspects of Optimal Disability Benefits." *Journal of Public Economics*, 57(1): 1–23.

Finkelstein, Amy, and Matthew J Notowidigdo. 2019. "Take-Up and Targeting: Experimental Evidence from SNAP." *The Quarterly Journal of Economics*, 134(3): 1505–1556.

Finkelstein, Amy, Erzo F. P. Luttmer, and Matthew J. Notowidigdo. 2013. "What Good Is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption." *Journal of the European Economic Association*, 11: 221–258.

Finkelstein, Amy, Nathaniel Hendren, and Erzo F. P. Luttmer. 2019. "The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment." *Journal of Political Economy*, 127(6): 2836–2874.

Finkelstein, Amy, Neale Mahoney, and Matthew J. Notowidigdo. 2018. "What Does (Formal) Health Insurance Do, and for Whom?" *Annual Review of Economics*, 10(1): 261–286.

French, Eric, and Jae Song. 2014. "The Effect of Disability Insurance Receipt on Labor Supply." *American Economic Journal: Economic Policy*, 6(2): 291–337.

Gelber, Alexander, Timothy Moore, and Alexander Strand. 2018. "Disability Insurance Income Saves Lives." Unpublished Manuscript.

Greszler, Rachel, Robert E Moffit, and Nina Owcharenko Schaefer. 2019. "Social Security and Health Care Entitlement Reform." *Blueprint for Balance: A Federal Budget for Fiscal Year 2020*.

Hendren, Nathaniel. 2020. "Measuring Economic Efficiency Using Inverse-Optimum Weights." *Journal of Public Economics*, 187: 104198.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies." *The Quarterly Journal of Economics*, 135(3): 1209–1318.

Homonoff, Tatiana, and Jason Somerville. 2020. "Program Recertification Costs: Evidence from SNAP." National Bureau of Economic Research.

Jacobson, Louis S, Robert J LaLonde, and Daniel G Sullivan. 1993. "Earnings Losses of Displaced Workers." *The American Economic Review*, 685–709.

Jones, Maggie R., and James P. Ziliak. 2019. "The Antipoverty Impact of the EITC: New Estimates from Survey and Administrative Tax Records." Center for Economic Studies, U.S. Census Bureau Working Papers 19–14.

Kleven, Henrik Jacobsen, and Wojciech Kopczuk. 2011. "Transfer Program Complexity and the Take-Up of Social Benefits." *American Economic Journal: Economic Policy*, 3(1): 54–90.

Lieber, Ethan MJ, and Lee M Lockwood. 2019. "Targeting with In-Kind Transfers: Evidence from Medicaid Home Care." *American Economic Review*, 109(4): 1461–85.

Liebman, Jeffrey B. 2015. "Understanding the Increase in Disability Insurance Benefit Receipt in the United States." *Journal of Economic Perspectives*, 29(2): 123–50.

Lillard, Lee A, and Yoram Weiss. 1997. "Uncertain Health and Survival: Effects on End-of-Life Consumption." *Journal of Business & Economic Statistics*, 15(2): 254–268.

Livermore, Gina, David Stapleton, Henry Claypool, et al. 2009. *Health insurance and health care access before and after SSDI entry.* Commonwealth Fund.

Low, Hamish, and Luigi Pistaferri. 2015. "Disability Insurance and the Dynamics of the Incentive Insurance Trade-off." *American Economic Review*, 105(10): 2986–3029.

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

McCrery, Jim, and Earl Pomeroy. 2016. "SSDI Solutions: Ideas to Strengthen the Social Security Disability Insurance Program."

Meyer, Bruce D, and Wallace KC Mok. 2019. "Disability, Earnings, Income and Consumption." *Journal of Public Economics*, 171: 51–69.

Meyer, Bruce D, Wallace KC Mok, and James X Sullivan. 2015. "Household Surveys in Crisis." *Journal of Economic Perspectives*, 29(4): 199–226.

Meyer, Laura, and Ife Floyd. 2020. "Cash Assistance Should Reach Millions More Families to Lessen Hardship." Report, Center on Budget and Policy Priorities.

Mueller, Andreas I, Jesse Rothstein, and Till M Von Wachter. 2016. "Unemployment Insurance and Disability Insurance in the Great Recession." *Journal of Labor Economics*, 34(S1): S445–S475.

Parsons, Donald O. 1991. "The Health and Earnings of Rejected Disability Insurance Applicants: Comment." *The American Economic Review*, 81(5): 1419–1426.

Rege, Mari, Kjetil Telle, and Mark Votruba. 2009. "The Effect of Plant Downsizing on Disability Pension Utilization." *Journal of the European Economic Association*, 7(4): 754–785.

Riley, Gerald F. 2006. "Health insurance and access to care among Social Security Disability Insurance beneficiaries during the Medicare waiting period." *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 43(3): 222–230.

Saez, Emmanuel, and Stefanie Stantcheva. 2016. "Generalized Social Marginal Welfare Weights for Optimal Tax Theory." *American Economic Review*, 106(1): 24–45.

Short, Pamela Farley, and France Marie Weaver. 2008. "Transitioning To Medicare Before Age Sixty-Five: The question is whether to eliminate, partially or entirely, the two-year waiting period for Medicare." *Health Affairs*, 27(Suppl1): w175–w184.

Sullivan, Daniel, and Till Von Wachter. 2009. "Job Displacement and Mortality: An Analysis Using Administrative Data." *The Quarterly Journal of Economics*, 124(3): 1265–1306.

United States Office of Management and Budget. 2020. *Fiscal Year 2021: A Budget for America's Future.* U.S. Government Publishing Office.

Von Wachter, Till, Jae Song, and Joyce Manchester. 2011. "Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program." *American Economic Review*, 101(7): 3308–29.

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."[46]

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-

[46] 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).[47]

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

[47]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 **?** 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.[48]

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 and SSI applications. In the SIPP-SSA linked data, we use self-reported work-limiting

[48]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.

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).[49] 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.

- Robustness of Fact 3: With respect to the PSID severity measure, non-recipients

[49] 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.

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.[50]

- 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 49 and 50 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 50 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, 2015; Cabral and Cullen, 2019; Meyer and Mok, 2019), since for many recipients the

[50] 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 49, also appear to work against the M-NDI vs. M-DI result, since true M-DI appear more disadvantaged than self-reported M-DI.

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 F4.

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).[51] 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, since more-severe individuals tend to benefit more from health insurance.[52,53]

[51]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.

[52]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.

[53]Note that health insurance, because of its moral hazard cost and insurance value, is potentially

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.[54] 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.[55] 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.[56]

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

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.

[54]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.

[55]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."

[56]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.

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.[57] 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 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

[57] 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.

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 providing 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.[58]

Results. We present results for the different health insurance scenarios in Appendix Table F3. 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.

[58]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 F7 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 ω of a Δb decrease in the disability benefit is

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

where DI_ω equals one if the individual receives disability benefits in state ω under status quo policies (i.e., before the benefit decrease). Recall from equation (7) that the ex post marginal value in state ω 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 ω, $\bar{z}$ 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 ω into the sum of (i) the mean benefit received in states in the same health category as ω, $h_\omega \in \{\text{L, M}\}$, 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 Ω_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 θ. The equations should be understood to apply to a particular risk type.

Health-contingent benefits, $E(b|h = h_\omega)$. State ω's health-contingent benefit is the mean benefit received in states in the same health category as ω, $h_\omega \in \{\text{L, M}\}$, $E(b|h = h_\omega) = p(\Omega_b|h)b$. 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 Ω (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 Ω_b). State ω's within-health category transfer is $(b - E(b|h = h_\omega)) \geq 0$ if the individual receives a benefit in ω and $-E(b|h = h_\omega) \leq 0$ if the individual does

not receive a benefit in ω. 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.[59]

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 Ω_b, $\sum_h \left\{MV_h^{across} + MV_h^{within}\right\}$ is

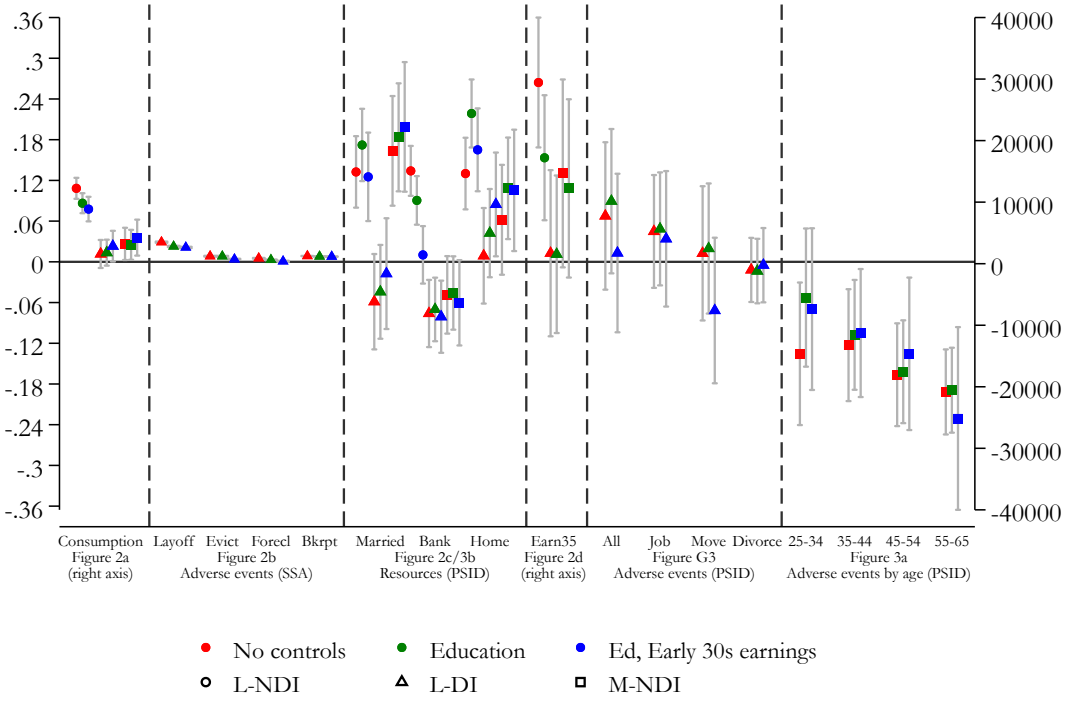
$$= p(\Omega_b)\left\{1 + \underbrace{Cov_h\left[\frac{p(\Omega_b|h)}{p(\Omega_b)}, E\left(\widehat{\lambda}|h\right)\right] + E_h\left\{\frac{p(\Omega_b|h)}{p(\Omega_b)}[1-p(\Omega_b|h)]\left[E\left(\widehat{\lambda}|h,\Omega_b\right) - E\left(\widehat{\lambda}|h,\sim\Omega_b\right)\right]\right\}}_{M_\Omega(\Omega_b)}\right\}.$$

The first term, $p(\Omega_b)$, is the mechanical effect of an increase in benefits in Ω_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.

[59] 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: Selection into USDP within severity, controlling for fine health measures

	Within-L difference (Fact 2)		Within-M difference (Fact 3)	
	No controls	Health controls	No controls	Health controls
Consumption	-10,667***	-6,659***	-3,219***	-4,452***
	(839.9)	(993.0)	(883.3)	(954.4)
Married	-0.191***	-0.201***	-0.164***	-0.220***
	(0.0253)	(0.0296)	(0.0418)	(0.0457)
Homeownership	-0.121***	-0.168***	-0.0620	-0.0989**
	(0.0253)	(0.0283)	(0.0432)	(0.0478)
Banking relationship	-0.210***	-0.161***	0.0487	0.0605
	(0.0174)	(0.0206)	(0.0394)	(0.0432)
Age-35 earnings	-27,744***	-16,797***	-14,672***	-15,160**
	(4,769)	(6,166)	(4,569)	(5,957)
Head age		X		X
Diagnosis		X		X
Self-reported health (prior)		X		X
Self-reported health (current)		X		X
ADLs/IADLs		X		X

Notes: Table presents the coefficient estimate on an indicator of disability receipt from a regression of the outcome (row name) on the indicator alone ("No controls" columns) or on the indicator plus health controls ("Health controls" columns). The first two columns show results for less-severe households (sample size 6,666; 2,034 for age-35 earnings). Columns (3) and (4) show results for more-severe households (sample size 556; 180 for age-35 earnings). The health controls are age (linear control), diagnosis (dummies for arthritis, asthma, blood disease, cancer, diabetes, heart attack, heart disease, lung disease, stroke, psychiatric condition, learning disability, memory problems, depression, and other), previous and current self-reported health (dummies for current severe condition, current moderate condition, current excellent/very good/good/fair poor health, current work limitation a lot/somewhat/just a little/not at all, ever severe condition, ever moderate condition, ever fair health, and ever poor health), and activities of daily living (dummies for the number of ADLs and the number of IADLs).

Table F2: 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 Ω_τ 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 28. 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 F3: 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 α 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 α 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 F4: 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 Ω_τ 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 28. Monetary amounts are in 2016 dollars per year. Sample sizes: 313 M-DI, 438 L-DI, 248 M-NDI.

Table F5: 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. 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 Ω_τ 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 28. Monetary amounts are in 2016 dollars per year. Sample sizes: 313 M-DI, 438 L-DI, 248 M-NDI.

Table F6: Decomposition of the markup on DI based on different categorizations of health

Health categories	Number of categories	Share of DI markup from insuring: Health	Non-health
More- vs. less-severe (baseline)	2	37%	63%
Ever vs. never severe	2	43%	57%
More- vs. less-severe at first receipt	2	41%	59%
Severe/moderate/minor/none	4	44%	56%
6 severity categories	6	44%	56%
Self-reported health	5	24%	76%
Severity x self-reported health	30	49%	51%

Notes: Table presents shares of the overall DI markup from insurance of health and non-health risk based on different categorizations of health, defined by the row. Contributions of insurance of health and non-health risk to overall markup are defined in equation (10). Shares of markup from insurance of each type of risk is the contribution of insurance of that type of risk to the overall markup as a share of the overall markup. "More- vs. less-severe (baseline)" repeats the baseline decomposition (shown in the first row of Table 2), which uses two health categories: more-severe (if the household reports that the head or spouse "can do nothing" as a result of a health condition or that a health condition limits "a lot" the amount of work that the head or spouse can do) and less-severe (otherwise). "Ever vs. never severe" uses two health categories: ever more-severe (in any of the PSID waves from 1985–2016 in which the household appears) and never more-severe. "More- vs. less-severe at first receipt" uses two health categories: more-severe at first receipt of DI and not. "Severe/moderate/minor/none" uses four health categories: more-severe, moderate severity (health limits work "somewhat"), minor severity (health limits work "just a little"), and none ("no health condition limits work" or health limits work "not at all"). "6 severity categories" uses six health categories: because of health, head or spouse "can do nothing"; health limits work "a lot"; health limits work "somewhat"; health limits work "just a little"; health limits work "not at all"; head and spouse do not have "any physical or nervous condition that limits the type of work or the amount of work." "Self-reported health" uses the five health categories: "excellent," "very good," "good," "fair," and "poor." "Severity x self-reported health" uses the 30 categories defined by interacting the six severity categories from the "6 severity categories" specification with the five self-reported health categories. Sample size: 751.

Table F7: 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 Ω_τ 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.

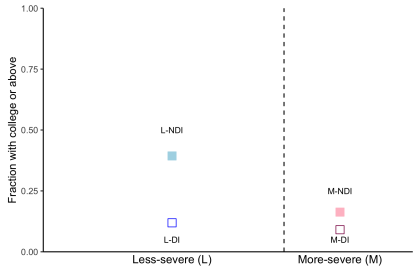
Table F8: Value of L-DI and DI benefits by safety net generosity in different U.S. states

	Surplus per household ($)		Recipiency rate		Markup		Cf no-benefit earnings ($)		Sample size	
State quartile	L-DI	DI	L-DI	DI	L-DI	DI	L-DI	DI	L-DI	DI
1 (least generous)	929	1,630	0.073	0.130	1.25	1.19	5,169	4,466	173	292
2	305	660	0.062	0.114	0.719	0.696	6,137	4,727	113	199
3	286	890	0.052	0.103	0.626	0.862	5,712	4,619	64	123
4 (most generous)	193	432	0.040	0.081	0.678	0.698	10,899	7,590	83	130

Notes: Table presents the surplus per household ($p(\Omega_\tau)s$), recipiency rate, mean markup on benefits ($E(M_\omega)$), and counterfactual no-benefit earnings for L-DI and DI recipients for each quartile of the distribution of the generosity of the non-DI safety net in different U.S. states. U.S. states are categorized into generosity quartiles according to the state's ratio of families receiving TANF to families living in poverty in 2017, as calculated in Appendix Table 1 of Meyer and Floyd (2020). Surplus per household is in units of government revenue per household. Monetary amounts are in 2016 dollars per year. See Table 1 for details on the methods and data.

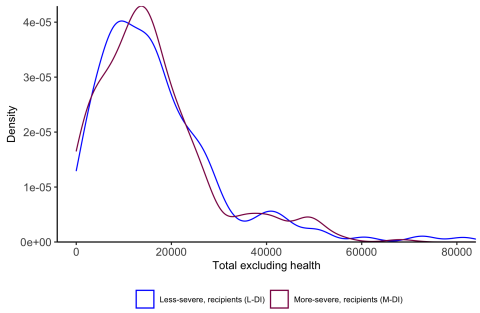
G Supplemental Materials (not for publication): Additional Figures

Figure G1: PSID: College or above



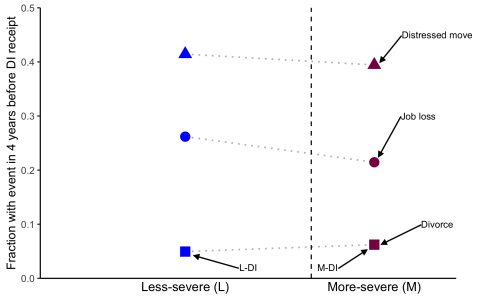
Notes: Figure presents rates of college degree or above in the 2017 PSID for each of the four severity-by-USDP receipt groups: less-severe non-DI-recipients (6,312 L-NDI), more-severe non-recipients (250 M-NDI), less-severe recipients (443 L-DI), and more-severe recipients (316 M-DI). "More-severe" is 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



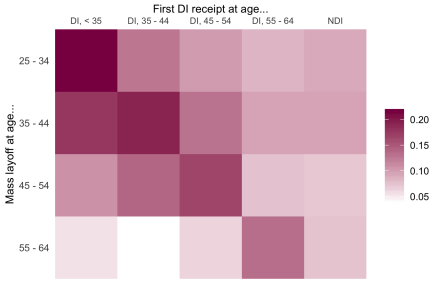
Notes: Figure presents the distribution of consumption (excluding health care expenditures) one year prior to the year of initial USDP receipt for each of the two severity-by-DI-receipt groups in the 2017 PSID: more-severe DI-recipients (213 M-DI), and less-severe recipients (286 L-DI). Household consumption 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



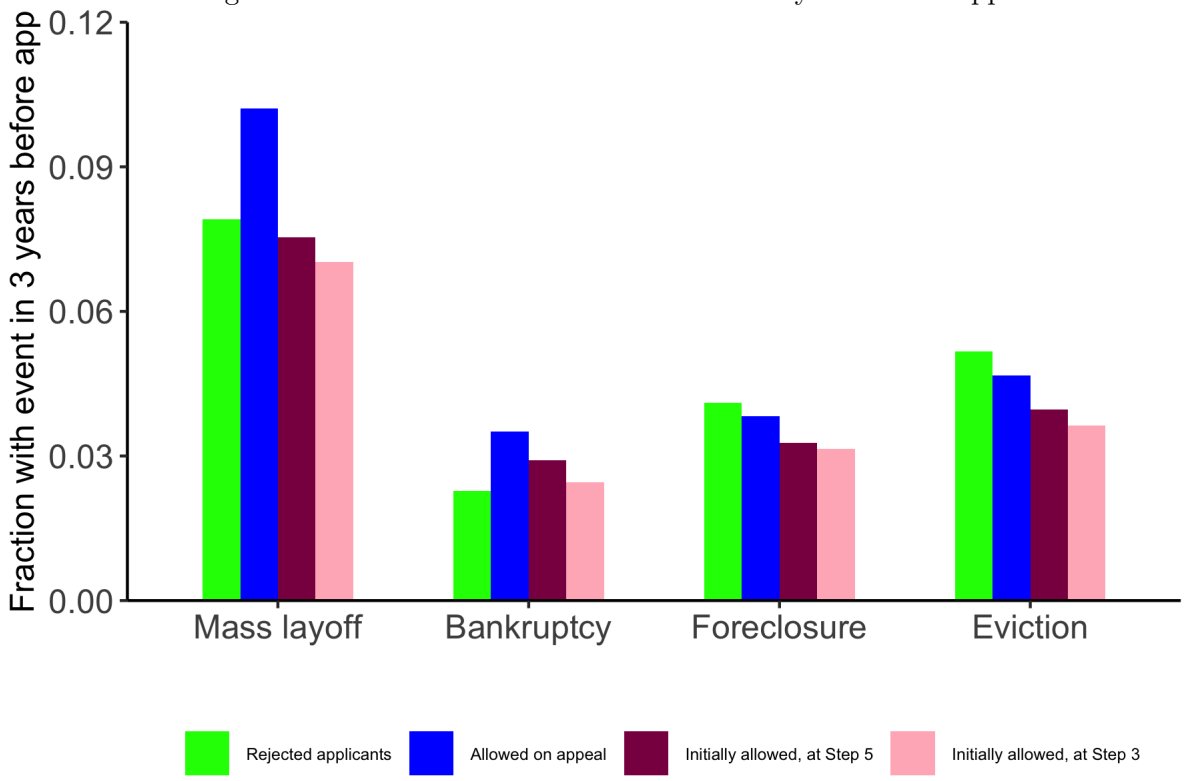
Notes: Figure presents rates of various adverse life events in the four years before receiving USDP for households receiving USDP in the 2017 PSID, by more-severe and less-severe. Job loss includes "involuntary" reasons for separation: strike/lockout, laid off/fired, or company going out of business 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



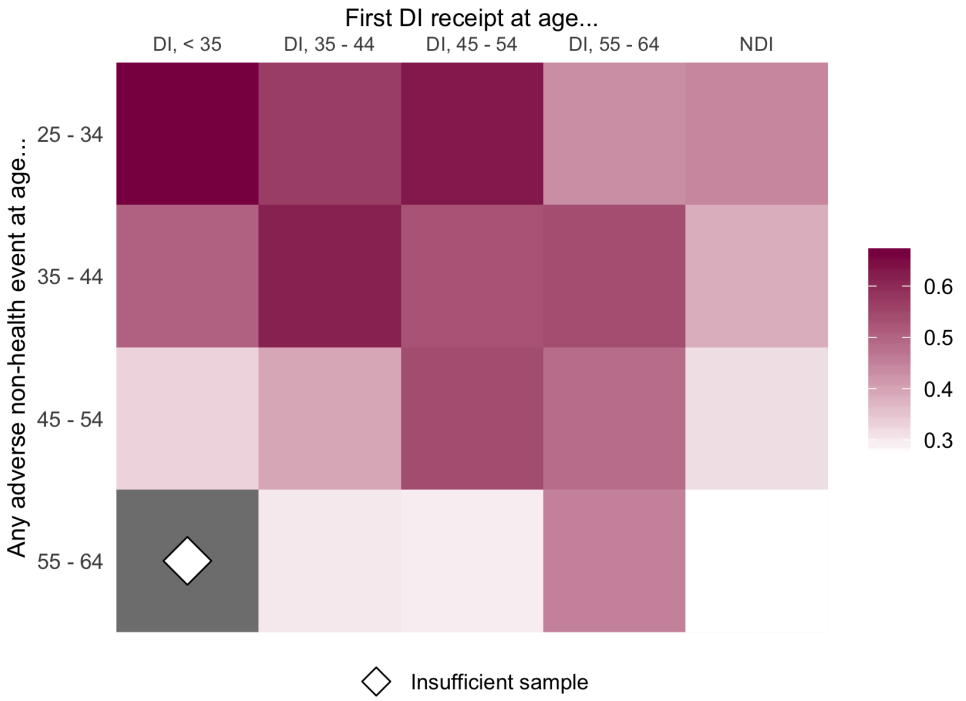
Notes: Figure represents rates of mass layoff at various ages for individuals in the CWHS, by USDP status, and age at USDP entry. Individuals are first categorized by their USDP status: whether they ever received USDP (DI) or never received USDP (NDI). The DI group is further categorized by age at USDP 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



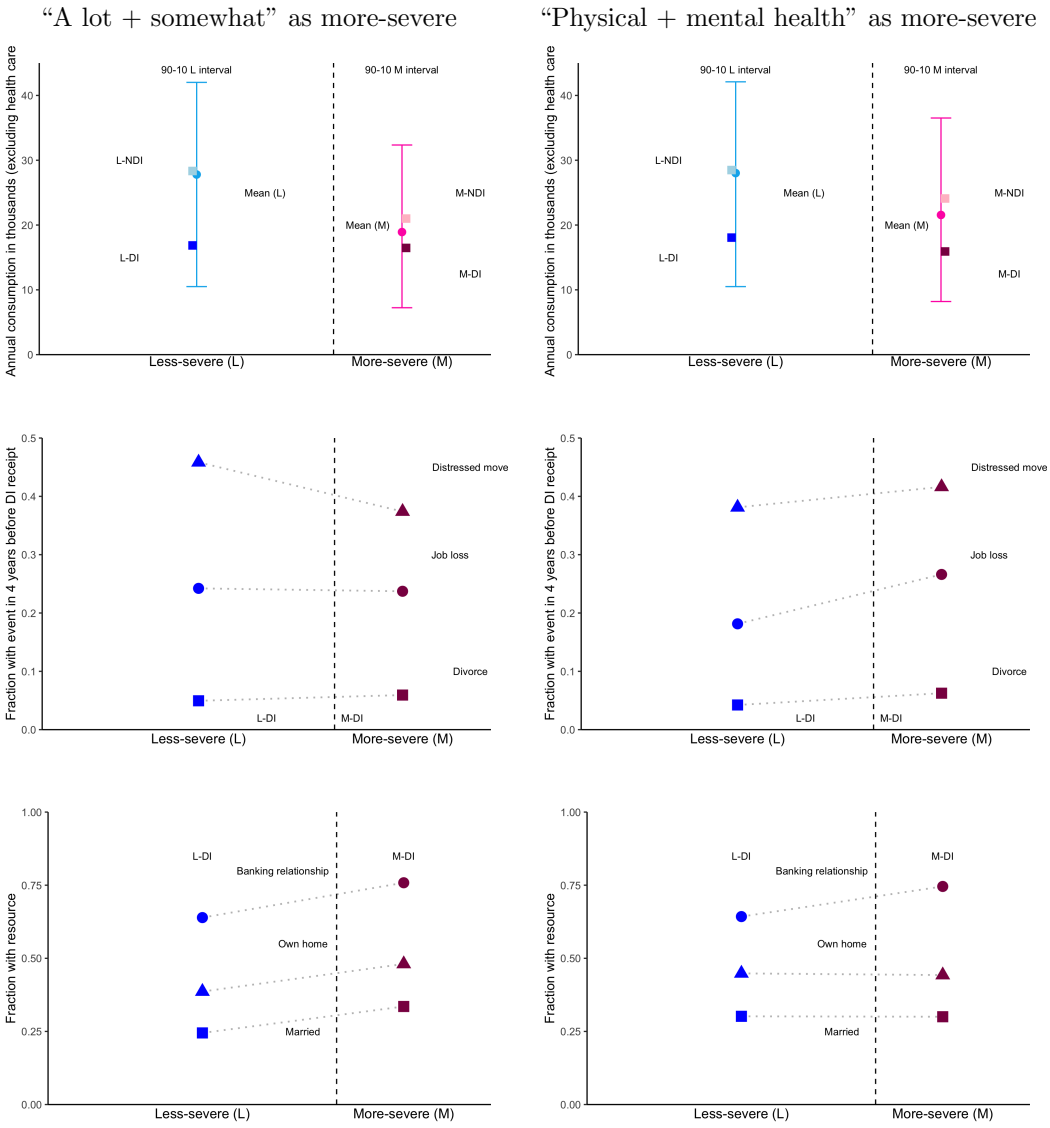
Notes: Figure presents rates of adverse financial events in the three years before receiving USDP (SSDI and SSI) for individuals who receive USDP benefits in SSA administrative data. "Initially allowed, at Step 3" indicates that recipient was allowed at initial state DDS level at the earliest possible 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 **?**. 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.

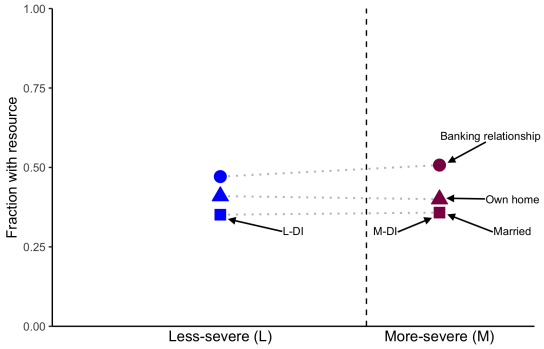
Figure G7: PSID: Positive results using alternative definitions of more-severe



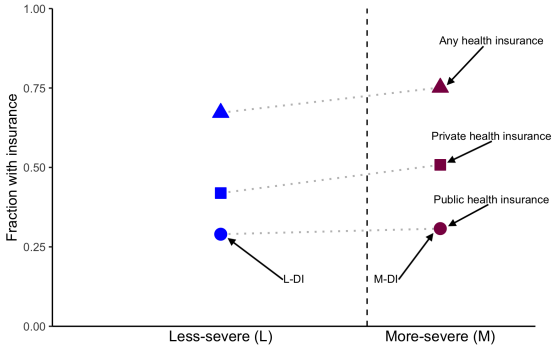
Notes: These figures replicate Figures 2a, G3, and 2c using two alternative definitions of more-severe. In the left column, we use a broader definition of more-severe that includes the "somewhat" response (in addition to the baseline "a lot" response) to the question of how much a health condition limits 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

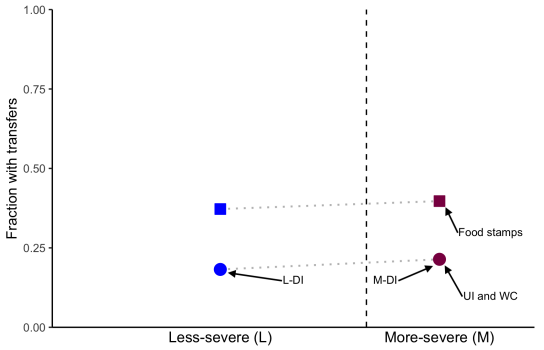
(a) Outside resources at application



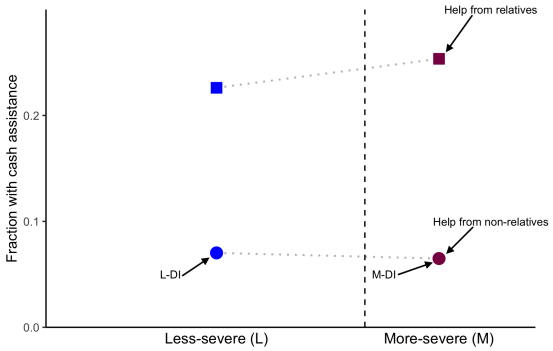
(b) Health insurance coverage (head) at application



(c) Transfers at application



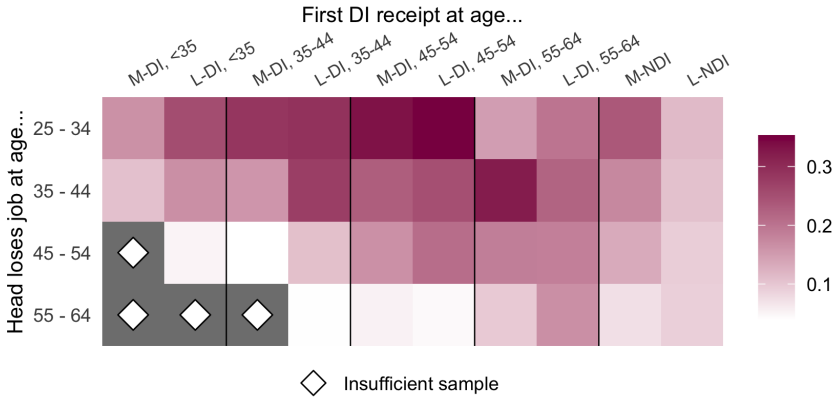
(d) Cash assistance at application



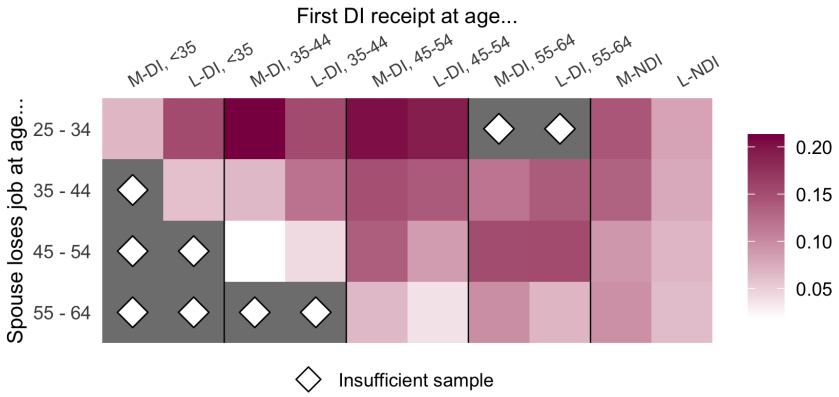
Notes: Figures present rates of outside resources one year prior to the year of initial USDP receipt for each of two severity-by-USDP-receipt groups in the 2017 PSID: more-severe USDP-recipients (243 M-DI), and less-severe recipients (320 L-DI). Figure G8a presents rates of marriage, banking relationship, 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



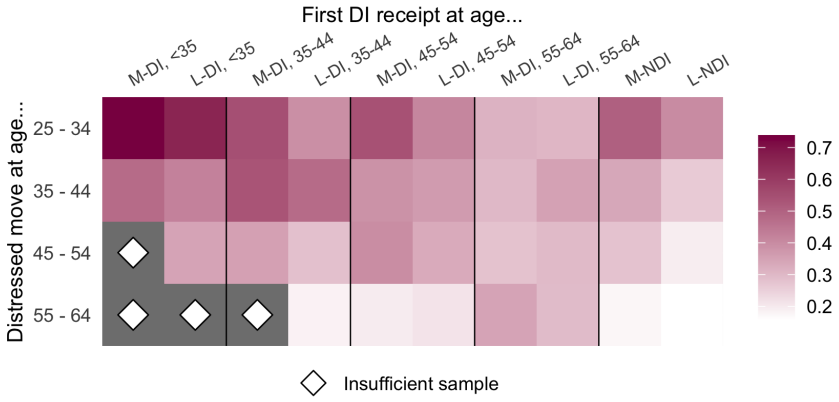
(b) Spouse job loss by USDP and health status



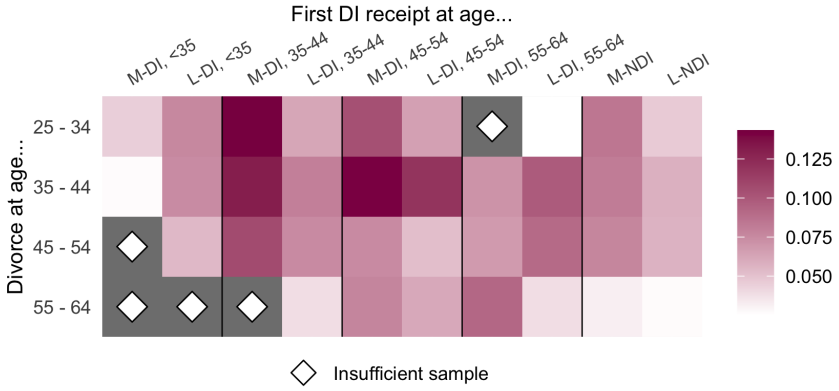
Notes: Figures present rates of head and spouse job loss at various ages for households in the PSID, by USDP status, health status, and age at USDP entry. Households are first categorized by their USDP status: whether they ever received USDP (DI) or never received USDP (NDI). DI are further classified 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 Ω_τ 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 28. 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 Ω_τ 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 28. 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? - Yes - No - 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? - A lot - Somewhat - Just a little - 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? - Yes - No
(2)	Does ... health or condition prevent ... from working at a job or business? - Yes - 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? 1. Owns or is buying home, either fully or jointly; mobile home owners who rent lots are included here 2. Pays rent 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?

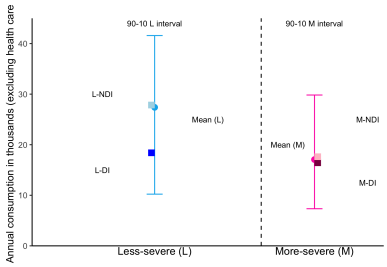
Table I5: Health insurance question

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

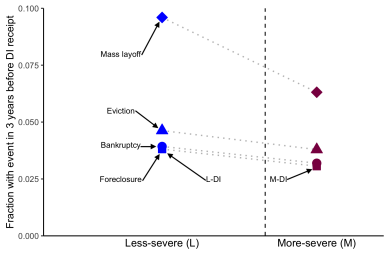
J Supplemental Materials (not for publication): SSDI-Only Figures/Tables

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

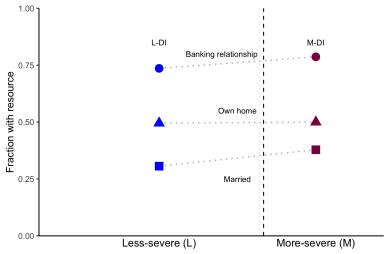
(a) PSID: Consumption



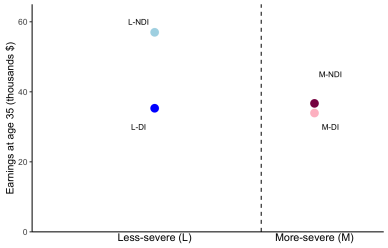
(b) Admin data: Prior to SSDI receipt



(c) PSID: Outside resources



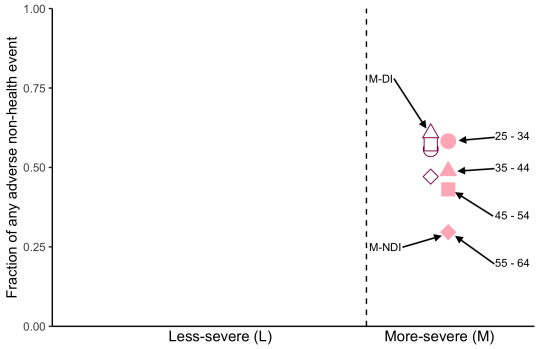
(d) PSID: Earnings at age 35



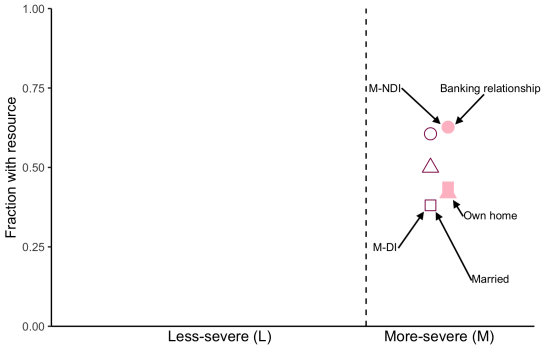
Notes: Figures J1a, J1c, and J1d present data from the PSID. Figure J1a presents statistics on consumption in the 2017 PSID for less-severe non-SSDI-recipients (6,442 L-NDI), more-severe non-recipients (316 M-NDI), less-severe recipients (312 L-DI), and more-severe recipients (250 M-DI). Household consumption 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

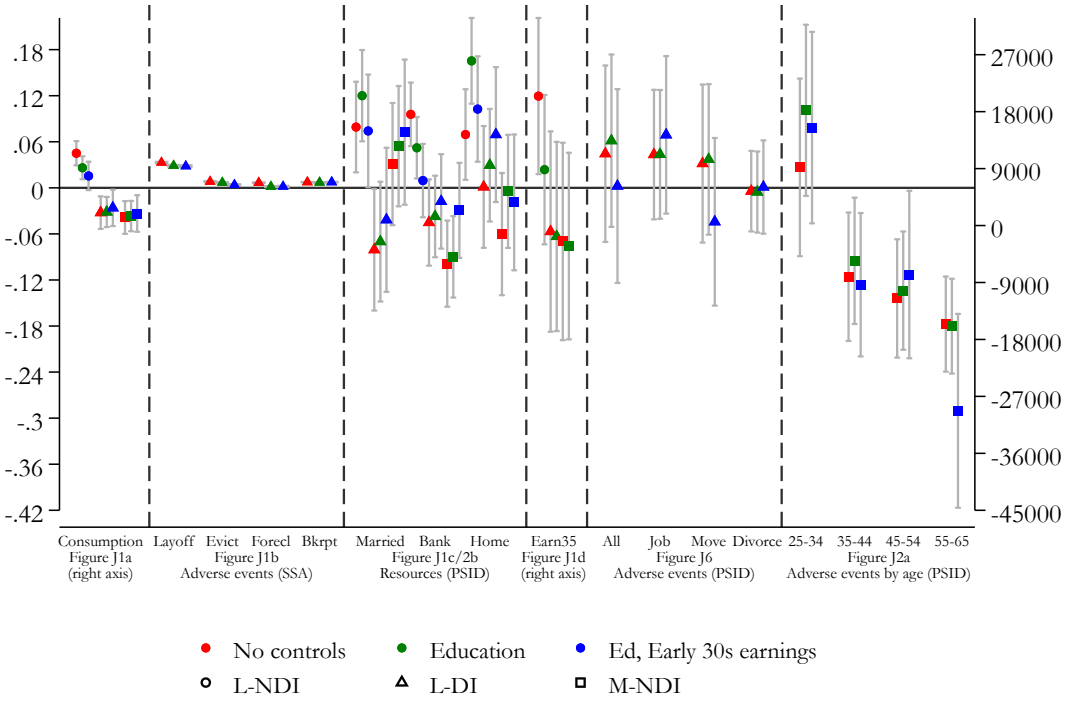


(b) PSID: Resources at more-severe onset



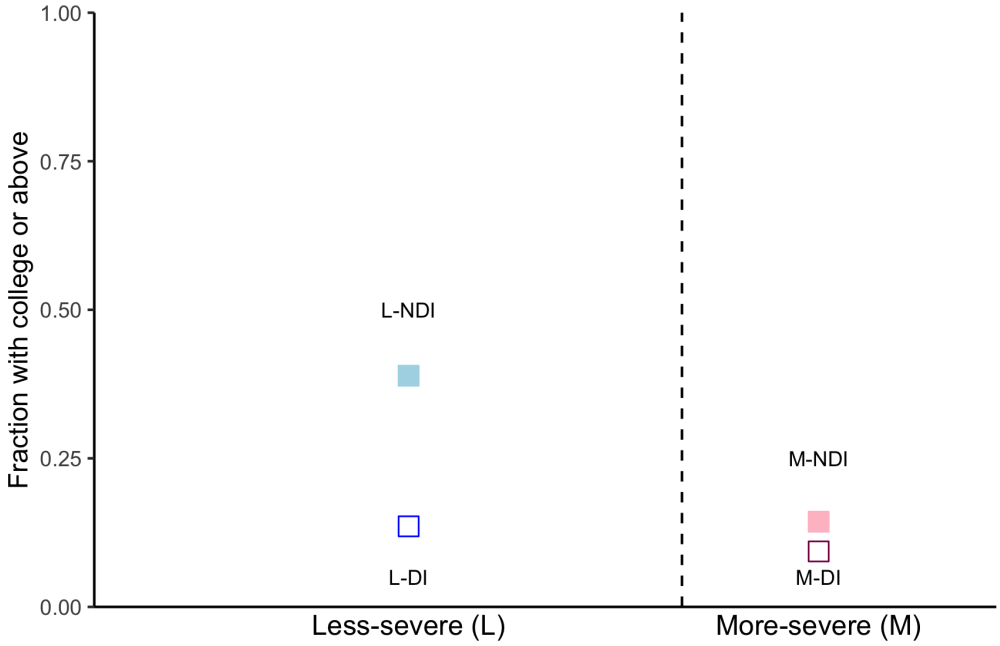
Notes: Figure J2a presents rates of experiencing an "adverse life event"—head or spouse job loss, involuntary move, or divorce—at various ages for households in the PSID, by SSDI status, health status, and age at SSDI entry. Households are first categorized by their SSDI status: whether they ever received 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



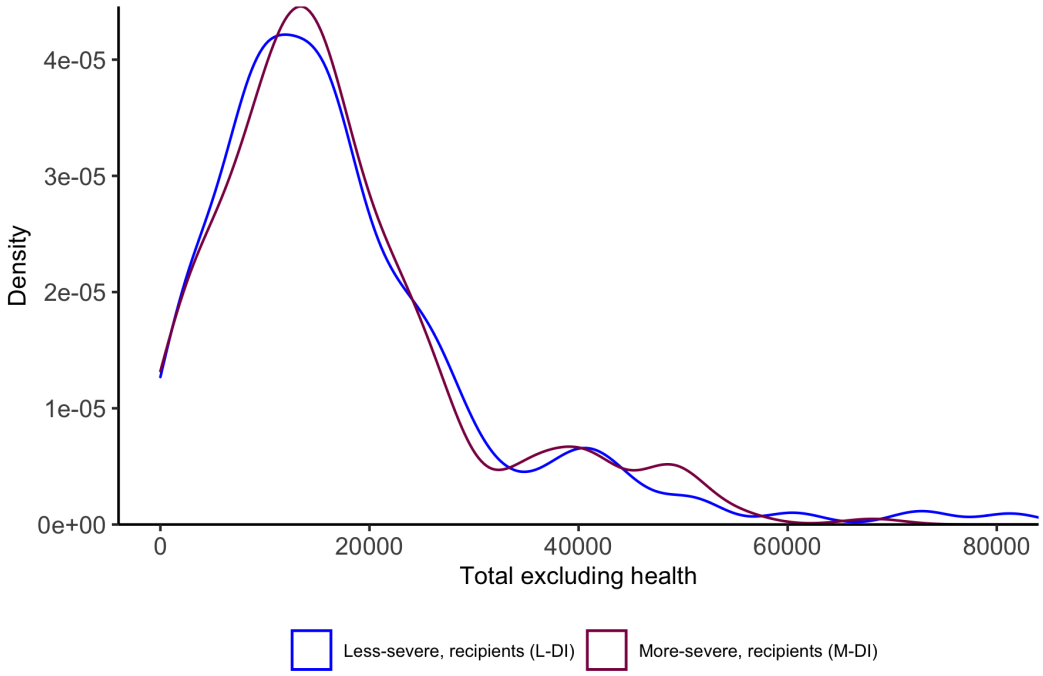
Notes: Figure presents all outcomes from Figures J1, J2, and J6 with controls for ability group. The figure plots the coefficients on indicators for each SSDI-by-severity group (L-NDI, L-DI, and M-NDI; M-DI is the excluded group) from a regression of the outcome on those indicators and ability controls. 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



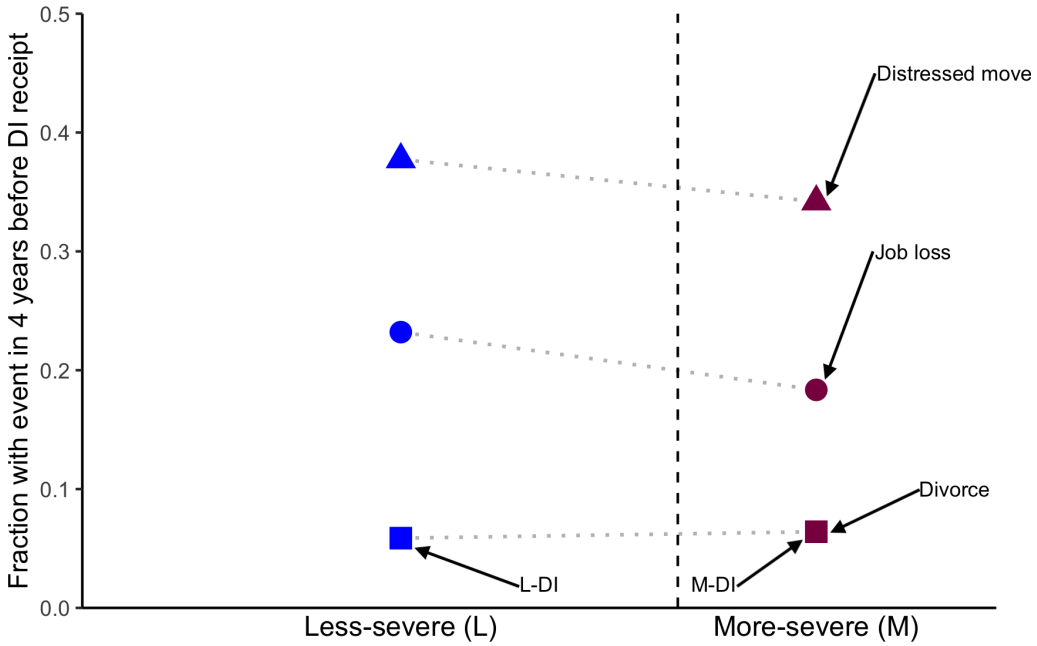
Notes: Figure presents rates of college degree or above in the 2017 PSID for each of the four severity-by-SSDI receipt groups: less-severe non-SSDI-recipients (6,442 L-NDI), more-severe non-recipients (316 M-NDI), less-severe recipients (312 L-DI), and more-severe recipients (250 M-DI). "More-severe" 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



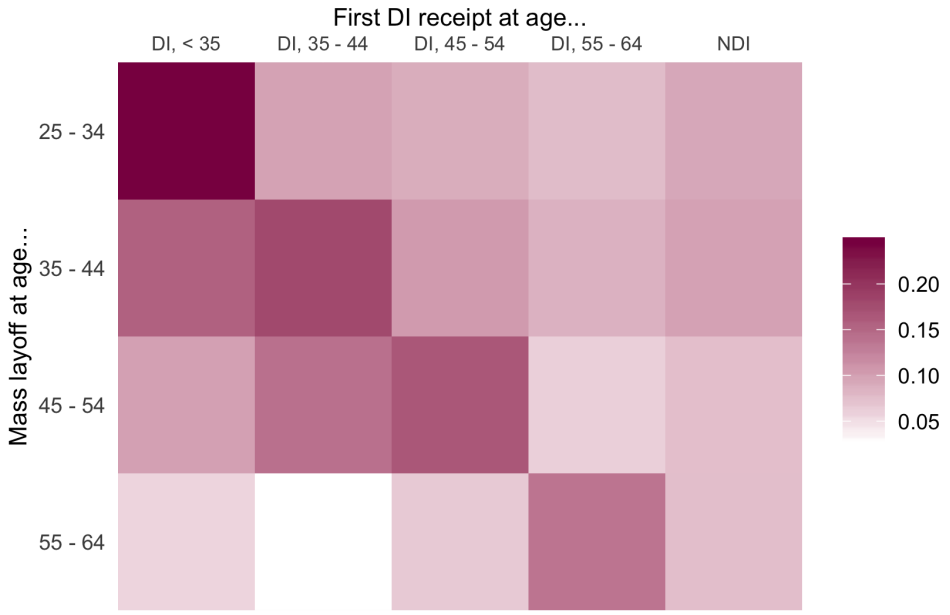
Notes: Figure presents the distribution of consumption (excluding health care expenditures) one year prior to the year of initial SSDI receipt for each of the two severity-by-SSDI-receipt groups in the 2017 PSID: more-severe SSDI-recipients (199 M-DI), and less-severe recipients (242 L-DI). Household 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



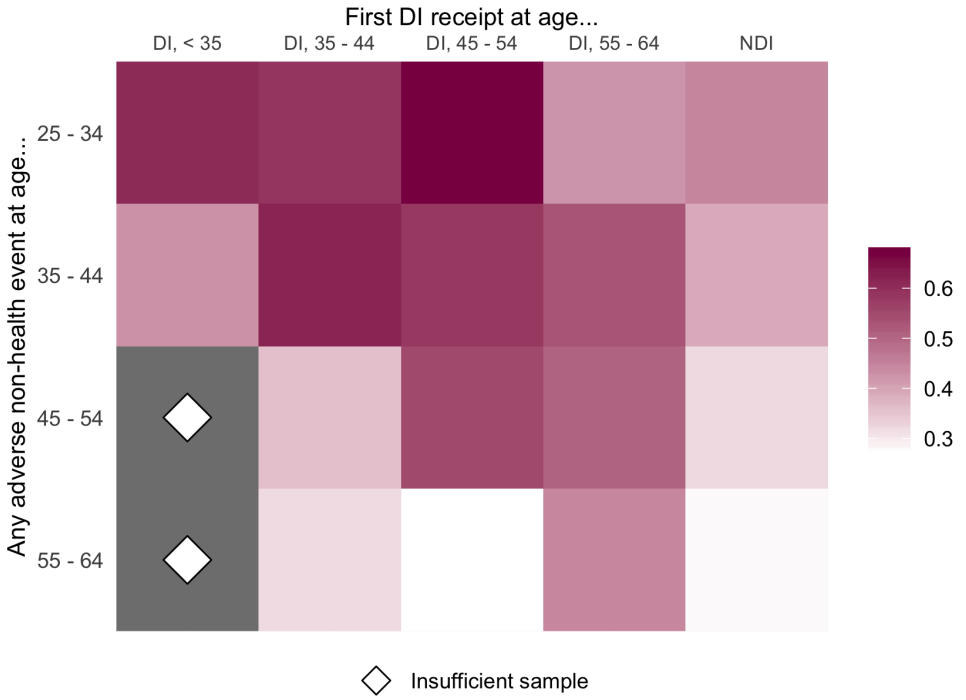
Notes: Figure presents rates of various adverse life events in the four years before receiving SSDI for households receiving SSDI in the 2017 PSID, by more-severe and less-severe. Job loss includes "involuntary" reasons for separation: strike/lockout, laid off/fired, or company going out of business 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



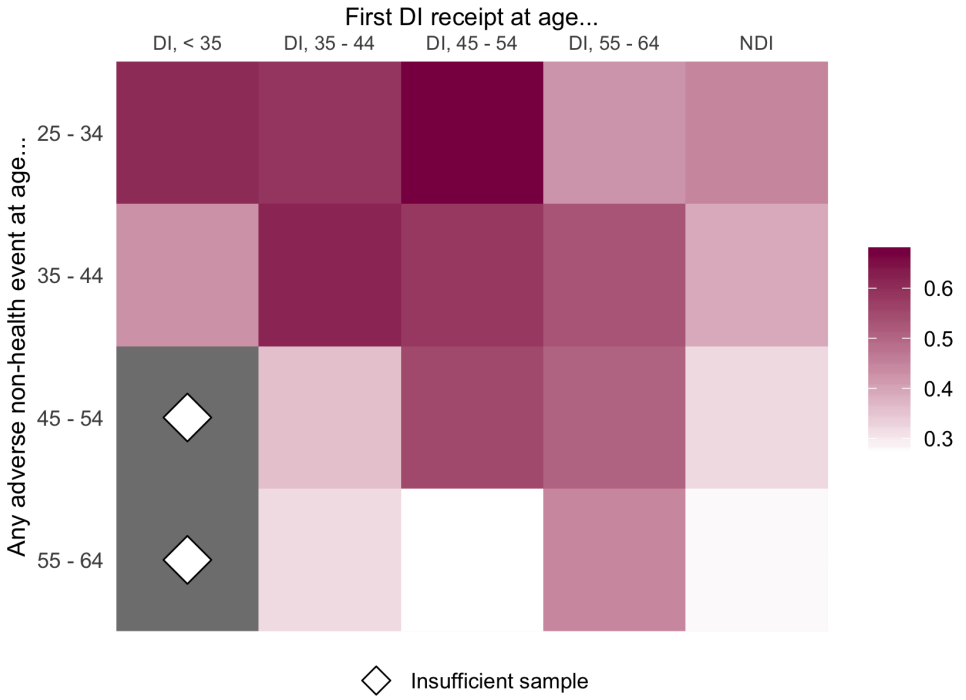
Notes: Figure represents rates of mass layoff at various ages for individuals in the CWHS, by SSDI status, and age at USP entry. Individuals are first categorized by their SSDI status: whether they ever received SSDI (DI) or never received SSDI (NDI). The SSDI group is further categorized by age at SSDI 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



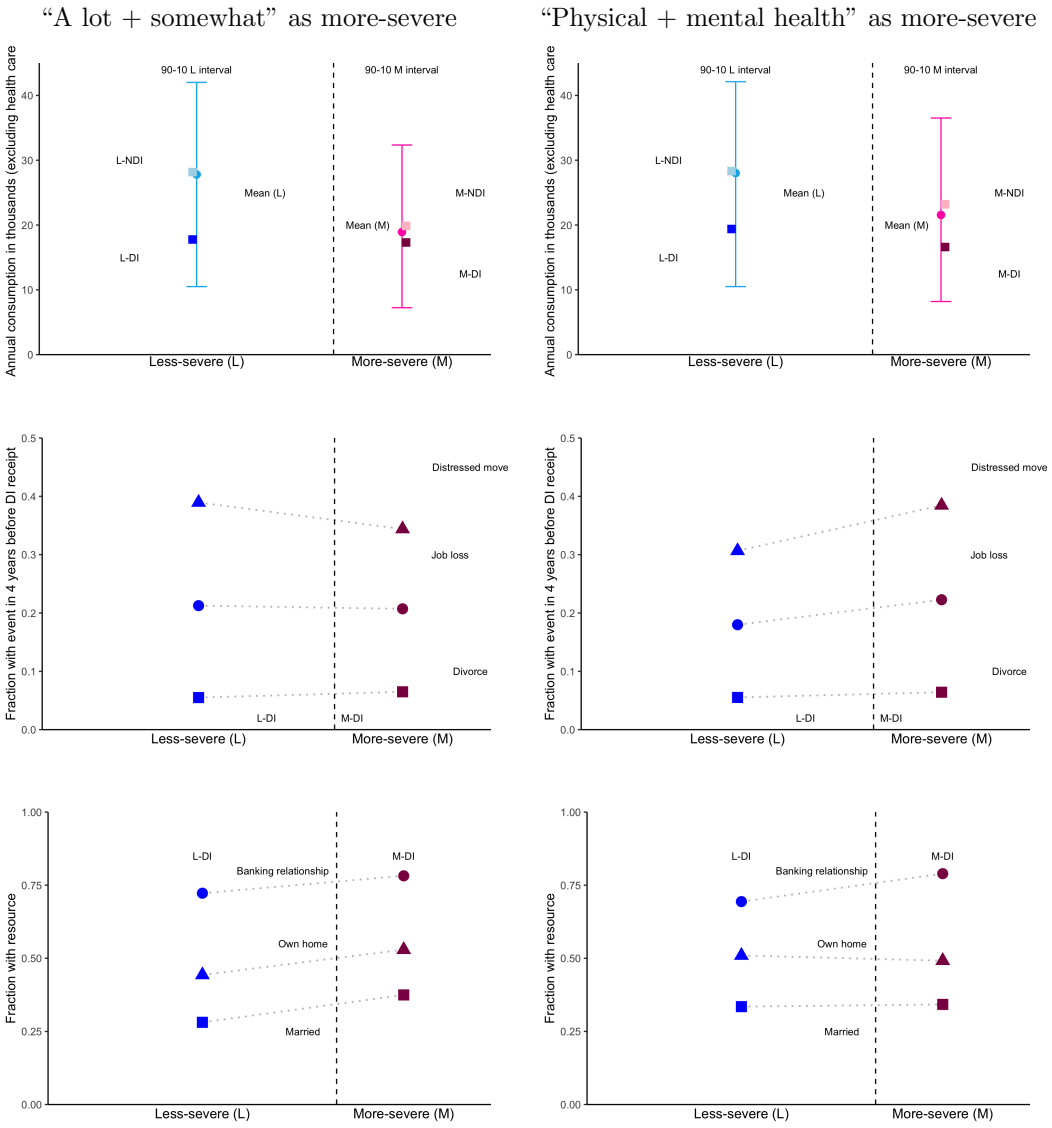
Notes: Figure presents rates of experiencing an "adverse life event"—head or spouse job loss, involuntary move, or divorce—at various ages for households in the PSID, by SSDI status, and age at SSDI entry. Households are first categorized by their SSDI status: whether they ever received SSDI (DI) or 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.

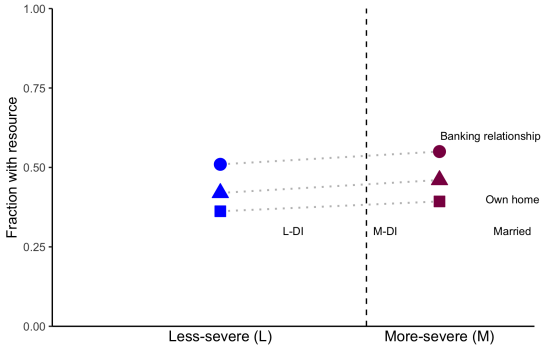
Figure J10: PSID: Positive results using alternative definitions of more-severe



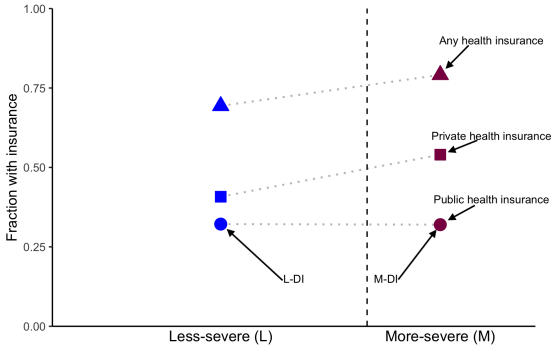
Notes: These figures replicate Figures J1a, J6, and J1c using two alternative definitions of more-severe. In the left column, we use a broader definition of more-severe that includes the "somewhat" response (in addition to the baseline "a lot" response) to the question of how much a health condition 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

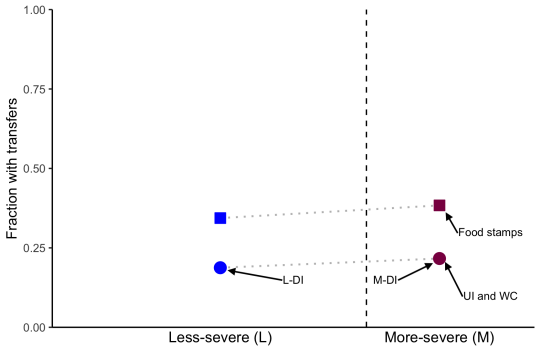
(a) Outside resources at application



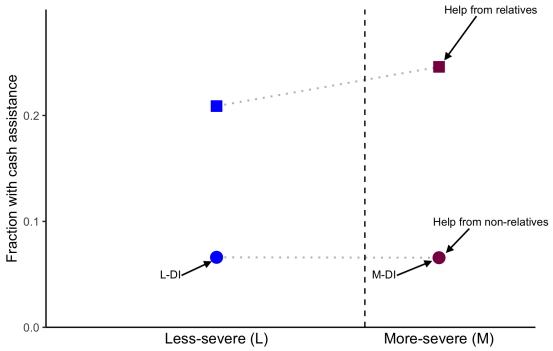
(b) Health insurance coverage (head) at application



(c) Transfers at application



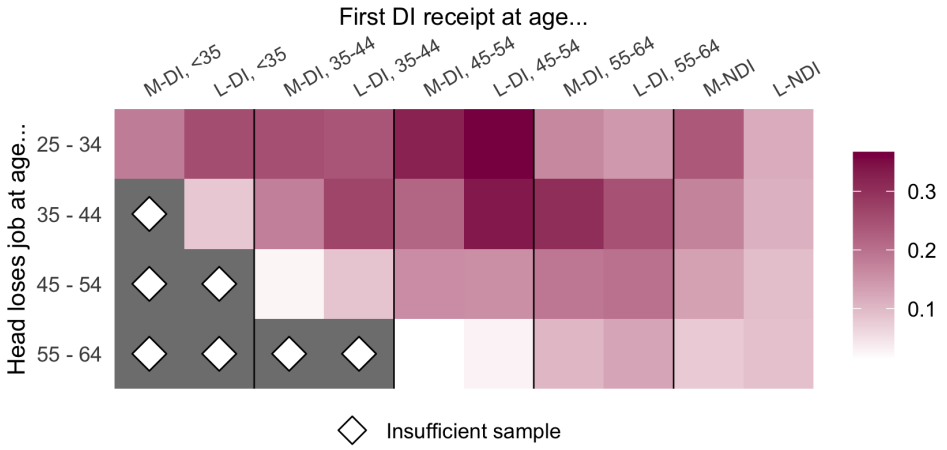
(d) Cash assistance at application



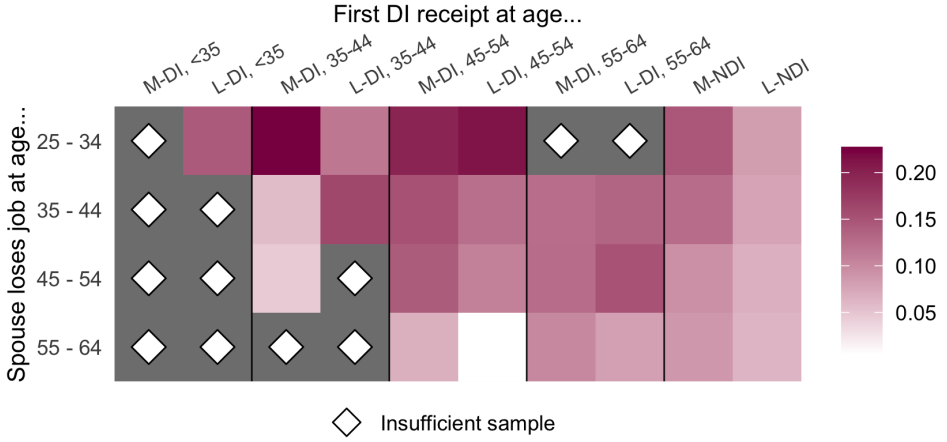
Notes: Figures present rate of outside resources one year prior to the year of initial SSDI receipt for each of the two-severity-by-SSDI-receipt groups in the 2017 PSID: more-severe SSDI-recipients (216 M-DI), and less-severe recipients (274 L-DI). Figure J11a presents rates of marriage, banking relationship, 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



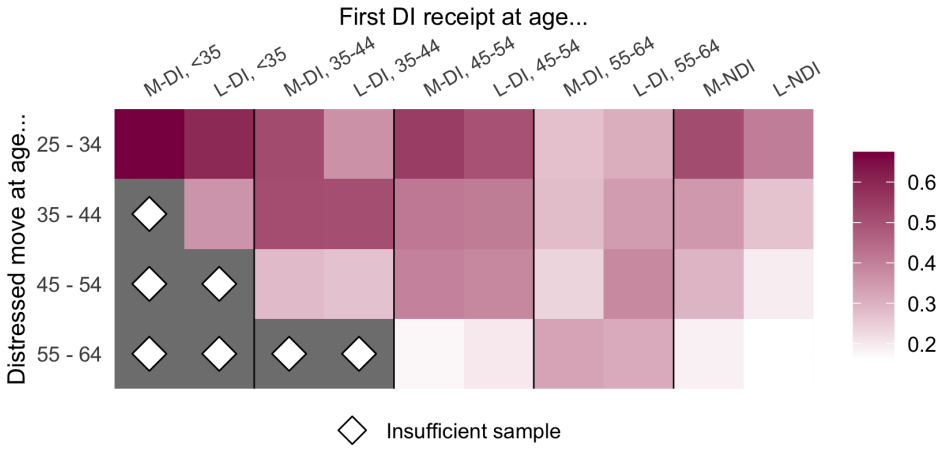
(b) Spouse job loss by SSDI and health status



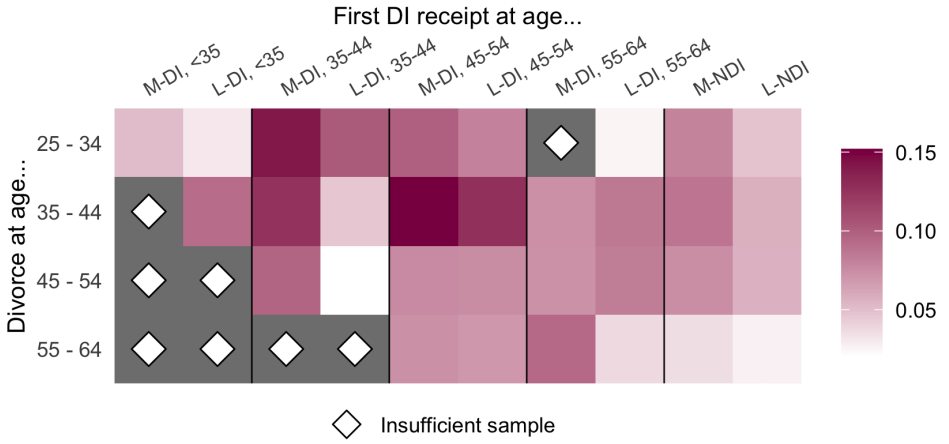
Notes: Figures present rates of experiencing head and spouse job loss at various ages for households in the PSID, by SSDI status, health status, and age at SSDI entry. Households are first categorized by their SSDI status: whether they ever received SSDI (DI) or never received SSDI (NDI). SSDI are further 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."