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

Manasi Deshpande Lee M. Lockwood

August 27, 2020

DRAFT

Abstract

The public debate over disability insurance (DI) has centered on concerns about individuals with less-severe health conditions receiving benefits. We go beyond health risk alone to quantify DI's overall insurance value, including value from insuring non-health risk. We find that DI 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 DI on the basis of non-health shocks is so strong among individuals with less-severe health conditions that by many measures less-severe DI recipients are worse off than severe DI recipients. As a result, under baseline assumptions, DI benefits to less-severe recipients have an annual value (insurance benefit less efficiency cost) of $7,700 per recipient, about three-fourths that of DI benefits to severe recipients ($9,900). Insurance against non-health risk accounts for about one-half of DI's value.

*Deshpande: University of Chicago and NBER, mdeshpande@uchicago.edu; Lockwood: University of Virginia and NBER, leelockwood@virginia.edu. We are grateful to David Autor, Katherine Baicker, Joseph Doyle, Arindrajit Dube, Mark Duggan, Liran Einav, Amy Finkelstein, Leora Friedberg, Michael Greenstone, Steven Haider, Jill Horwitz, Bill Johnson, Jonathan Kolstad, Magne Mogstad, Austin Nichols, Matthew Notowidigdo, Ed Olsen, Luigi Pistaferri, Jim Poterba, Adam Sacarny, Joe Shapiro, Isaac Sorkin, Melanie Wasserman, and seminar participants at Federal Reserve Board, Michigan State, MIT, Stanford University, University of Illinois at Chicago, and University of 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.

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 and Supplemental Security Income programs (which we collectively refer to as DI) together provide health insurance and $200 billion annually in cash benefits to nearly 13 million Americans. The statutory purpose of DI is to provide assistance to people who cannot work because of severe health conditions. Yet the expansion of DI 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 DI's incorrect awards with respect to health—especially individuals who receive DI despite not having a severe condition—are diluting DI's value.

In this paper, we go beyond health risk alone to quantify DI's overall insurance value, including value from insuring non-health risk. We quantify the extent to which DI insures different risks by comparing DI 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. We focus in particular on DI's "classification errors" with respect to health: individuals who receive benefits despite not having a severe condition ("non-severe recipients" or "inclusion errors") and individuals who do not receive benefits despite having a severe condition ("severe non-recipients" or "exclusion errors"). Building on these comparisons, we use a sufficient statistics approach to quantify the (ex ante) value of receiving DI 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.

Going "beyond health" is crucial for making informed judgments about DI and its classification errors. Given DI's statutory aim, it is natural that DI is generally evaluated on the basis of its provision of benefits to people with severe health conditions. Moreover, health is likely a strong indicator of the value of receiving DI in a particular state: Severe health conditions limit earnings, which both increases the insurance benefit of receiving DI transfers (by decreasing consumption and so, presumably, increasing marginal utility) and decreases the efficiency cost of behavioral responses to DI's earnings test and other features. But health is not a *perfect* indicator because there are many other, non-health determinants of marginal utility and of the efficiency cost of behavioral responses to DI. So from a broader social insurance perspective, DI's overall value also depends on the extent to which it insures non-health risk. Individuals face a wide variety of health and non-health risks, including job loss, productivity shocks, and changes in family structure. To the extent that a particular risk is not completely insured by other means, DI potentially insures or worsens that risk, depending on selection into DI, regardless of whether DI aims to insure non-health risk. DI insures (worsens) a risk to the extent that the likelihood of receiving DI benefits is posi-

DRAFT

tively (negatively) correlated with the size of the loss. For example, receiving DI in a state without a severe health condition but with limited earning prospects after a job layoff partially insures that risk and might thereby increase DI's value. Although DI's classification errors with respect to health necessarily reduce the extent to which DI insures health risk, the presence of non-health risk means that DI's errors might not reduce—and might even increase—DI's overall value.[1]

DI's value is ultimately an empirical question, one that we investigate using complementary positive and normative analyses. In the positive analysis, we establish new facts about selection into DI on the basis of non-health factors, using a combination of survey and administrative data. We find that non-severe DI recipients are on average much worse off than non-severe non-recipients, and by many non-health measures are even worse off than severe DI recipients. For example, we find in administrative data that prior to entering DI, non-severe recipients are 33% more likely to have experienced a mass layoff than severe recipients, 27% more likely to have experienced a foreclosure, and 16% more likely to have experienced an eviction. From survey data, non-severe recipients have similarly low consumption levels to severe recipients just prior to DI entry ($17,000 and $16,400, respectively). Conversely, severe non-recipients are on average *better* off than either severe or non-severe recipients. From survey data, severe non-recipients are 24% less likely to have experienced an adverse non-health event—job loss, distressed move, or divorce—than either severe or non-severe recipients. The strong associations between DI receipt and non-health shocks conditional on health means that DI's classification errors are highly selected, in ways likely to increase DI's value, mainly by increasing the extent to which DI insures non-health risk.

In the normative analysis, we determine how DI's targeting properties, including its classification errors with respect to health, translate into its value, taking other programs as given. Given the evidence from the positive analysis that DI insures a wide variety of risks, we use a flexible sufficient statistics approach that does not require modeling the full set of risks people face or insurance mechanisms they use. Our approach highlights the importance, in a broad class of models, of two key factors in determining the value of DI: the marginal utility of income and "counterfactual earnings"—how much one would have earned if not receiving DI. 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 DI. Our findings suggest that DI provides valuable insurance and increases welfare significantly: Under baseline assumptions, the value of DI benefits is $8,700 per recipient per year.

[1] Of course, one's overall normative judgment of the desirability of certain DI benefits might also depend on other, non-welfarist factors (beyond the effects on individuals' well-being) such as views about the appropriate role of DI. But for normative perspectives that put at least some weight on people's well-being, a benefit's value is an important determinant of its overall desirability.

DRAFT

Moreover, we find that DI's high value is in part *because of*, not despite, its classification errors with respect to health. We estimate that providing benefits to non-severe recipients creates a value (insurance benefit less distortion cost) of \$7,700 per recipient per year, about three-fourths that of providing benefits to severe recipients (\$9,900). As a result, DI's inclusion errors do not decrease DI's value—they increase it considerably, accounting for about half of DI's total value. DI's exclusion errors also appear to create value, since providing benefits to severe non-recipients would have a negative value (−\$2,200 per hypothetical recipient per year). DI's classification errors are so favorably selected that DI is more valuable than a hypothetical, infeasible program that targeted perfectly on severity. In other words, even somehow eliminating all of DI's classification errors at no administrative cost would significantly *decrease* DI's value.[2]

The high values of DI and its classification errors are driven by strongly favorable selection into DI receipt, overall and conditional on health. Selection is strong on both key determinants of the value of receiving DI in a particular state of the world: marginal utility and counterfactual earnings. Conditional on education and compared to non-recipients, under baseline assumptions DI recipients' mean marginal utility is 90% greater and their mean counterfactual earnings is less than one-tenth as large. Conditional on education and compared to severe recipients, non-severe recipients' mean marginal utility is just 3% lower and their mean counterfactual earnings is just \$3,100 higher.

Much of the favorable selection into DI receipt is on the basis of non-health shocks, which greatly increases DI's value by providing highly valuable insurance against non-health risk. Using a decomposition analysis, we find that about one-half of DI's insurance value comes from insuring risk *within* observable health groups. This suggests that much, or even most, of DI's value comes from insuring non-health risk. DI's insurance of non-health risk in turn depends crucially on the strong selection into application for DI, induced by DI's tags and screens. We estimate that a hypothetical DI program with the same adjudication standards but no systematic selection into application would have *negative* value, far below DI's large positive value.

We emphasize several points relevant for interpreting our results. First, our analysis takes other policies as given. Our finding that DI provides valuable insurance against a wide variety of risks in part reflects the gaps in the existing social safety net. It does not imply that DI is the most efficient way to insure these risks. Expanding other insurance programs—e.g., a large expansion of UI or a new wage insurance or temporary disability program—would likely reduce the value of DI's insurance against non-health risk. Second, our normative

[2]Our main analysis considers only insurance value relative to distortion cost. In Section 5.4, we additionally consider possible redistribution value from DI.

DRAFT

analysis focuses on DI's insurance and welfare effects and does not explicitly incorporate non-welfarist considerations, such as DI's alignment with its statutory aims or the underlying reasons someone has high marginal utility or low counterfactual earnings (e.g., whether due to labor market shocks or laziness). Third, our results do not speak to the effects of changes to particular DI design features, such as the earnings test, the five-month waiting period for benefits, or the weight on health in the adjudication process. Evaluating such changes would require evidence on the application responses to these changes, which is sparse. Finally, our normative analysis relies on several assumptions, including about behavioral responses to DI, about which relatively little is known. Better understanding 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 DI. The most contentious question in the debate is whether DI's inclusion errors—benefits to individuals without severe health conditions—are diluting the value of DI. Our results suggest that DI's inclusion errors are not diluting the value of DI but enhancing it. Whereas eliminating DI's inclusion errors would significantly reduce DI's value, "doubling down" on DI's classification errors by increasing benefit levels—which, among other things, further increases transfers from severe non-recipients to non-severe recipients—would increase DI's value. These conclusions are driven by our key positive finding that DI recipients without severe conditions are a highly selected group: They are much worse off than non-recipients, and by many non-health measures are worse off than recipients with severe conditions. As a result, DI benefits to non-severe recipients provide valuable insurance against a wide variety of risks at an efficiency cost that is small in comparison.[3]

The primary contribution of this paper is to quantify the extent to which DI and its classification errors insure non-health as well as health risk and to determine the welfare implications of this insurance. In doing so, we fill key holes in the evidence central to the debate over DI, including the non-health characteristics of less-severe DI recipients and the insurance benefit and distortion cost of these benefits. Our findings help reconcile two seemingly contradictory strands of the literature on DI. The first is a positive, empirical strand that has raised serious concerns about DI. The key findings of this literature are that (i) classification errors are prevalent (Benıtez-Silva et al., 1999; Von Wachter, Song and Manchester, 2011), (ii) labor market shocks increase DI enrollment (Autor and Duggan, 2003, 2006; Black, Daniel and

[3]Although we do not explicitly incorporate non-welfarist considerations in our main analysis, our results can be used to 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 different conclusions. For example, a non-welfarist planner who discounted the value of DI benefits to non-severe recipients (viewing each dollar of such benefits as being worth \$(1-d), where d is the discount) would get positive surplus from such benefits as long as the discount did not exceed about 35% ($d = 0.35$). That same planner would get positive surplus from DI as a whole as long as the discount on benefits to non-severe recipients did not exceed about 75% ($d = 0.75$).

DRAFT

Sanders, 2002; Charles, Li and Stephens, 2018; Rege, Telle and Votruba, 2009), and (iii) DI reduces the labor supply of marginal recipients (Maestas, Mullen and Strand, 2013; French and Song, 2014). These findings suggest that there are large differences between DI and its statutory aim—in particular, that DI provides benefits to many individuals without severe health conditions and discourages work. Our paper extends this literature by going "beyond health" to characterize severe and less-severe DI recipients and non-recipients on a wide variety of dimensions and quantify the implications for the value of DI.

The second, normative strand of literature on DI quantifies the welfare effects of certain reforms to DI. This strand typically finds that expanding DI would increase welfare. Bound et al. (2004) estimate the marginal welfare effect of a proportional increase in DI benefits and taxes and find that the typical worker would value the increase in benefits more than the cost of providing them. Chandra and Samwick (2009) and Meyer and Mok (2019) find that DI, together with other programs, provides valuable but incomplete insurance against health risk. Low and Pistaferri (2015) estimate a structural life-cycle model and find that expanding DI would increase welfare. Cabral and Cullen (2019) use the demand for private DI to bound the value of public DI and find a high lower bound. Our paper extends this literature by quantifying, with a flexible approach, the value of DI benefits in different types of states of the world and by illuminating the mechanisms underlying DI's welfare effects, particularly the role of non-health risk.

Our results on the strong selection into DI on the basis of non-health risk help reconcile the seemingly contradictory findings of the literature on DI: how expanding DI could increase welfare despite prevalent classification errors, non-health drivers, and reductions in work. The strong selection into DI means that DI's classification errors do not greatly decrease—and in fact appear to increase—DI's value. Non-health drivers of enrollment do not just drive up DI's costs, they also drive up DI's benefits in terms of insuring non-health risk. With respect to the effects of DI on labor supply, we do not provide new evidence on this question but instead quantify the implications of existing evidence for the value of DI. Our analysis highlights the key role of a less-discussed aspect of the existing evidence: that marginal DI recipients, despite their non-negligible counterfactual labor force *participation*, have extremely low counterfactual *earnings*.[4]

[4]The key evidence on the counterfactual labor supply of DI recipients comes from studies that exploit the quasi-random assignment of applicants to adjudicators (e.g., Maestas, Mullen and Strand, 2013; French and Song, 2014). This evidence is often interpreted as suggesting that awards to marginal recipients have high costs and low benefits, due to the large effect of DI on labor force participation (28 percentage points in Maestas, Mullen and Strand (2013), for example). Less emphasized but crucial for the costs and benefits of DI is the effect of DI on *earnings*, since that determines the fiscal externality cost and is informative about the private cost of the earnings test. Maestas, Mullen and Strand (2013) estimate that marginal recipients—likely the recipients with the highest counterfactual earnings—would have earned just $6,000–$9,000 per year had they not received DI and that being awarded DI reduces their average earnings by only about $4,000 per year. Of course, these estimates do not reveal *why* their counterfactual earnings are so low—whether

DRAFT

The paper proceeds as follows. In Section 2, we discuss the relationships between non-health risk, classification errors, and the value of DI, including the conditions under which classification errors would increase versus decrease DI's value. In Section 3, we provide institutional background and describe the survey and administrative data we use. In Section 4, we present the positive analysis of selection into DI receipt on the basis of a wide variety of health and non-health characteristics, focusing on DI's classification error groups. In Section 5, we estimate the insurance and welfare effects of receiving DI benefits in different types of states, investigate the underlying mechanisms, and discuss policy implications. Section 6 concludes.

2 Theory: Non-Health Risk, Classification Errors, and the Value of DI

DRAFT

Our goal is to go "beyond health" to determine the value of DI and its classification errors 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 severity of the non-health shock. From a health-based perspective, the objective of DI 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 suffered a qualifying health shock: $h \gtrless \bar{h}$. We label states in which $h \geq \bar{h}$ "severe" (S) and states in which $h < \bar{h}$ "non-severe" (NS). There are two types of errors relative to this perspective: awarding benefits in a non-severe state (an inclusion error) and failing to award benefits in a severe state (an exclusion error).

Now consider an alternative, "value-maximizing" perspective in which the goal is to maximize the ex ante value of DI 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 DI ("counterfactual earnings"), reflecting DI's discouragement of recipient labor supply. This perspective divides the state space into two sets: states in which the value of receiving DI is greater or less than zero: $s(h, n) \gtrless 0$. There are two types of errors relative to this perspective: awarding benefits when doing so reduces DI's value, and failing to award benefits when doing so would have increased DI's value.

The value of receiving DI in a particular state, $s(h, n)$, is likely strongly increasing in the

because of health, skills, work ethic, etc.—which is relevant for certain non-welfarist perspectives.

severity of the health shock, h. Severe health shocks increase health spending and worsen earning opportunities, and so tend to increase marginal utility and decrease counterfactual earnings. As a result, having a severe health condition is likely a strong indicator of the value of receiving DI: $E(s(h,n)|S) \gg E(s(h,n)|NS)$.[5] These two perspectives, health-based and value-maximizing, therefore likely "agree" on whether benefits should be provided in most states. Thus, even in terms of value alone, setting aside non-welfarist considerations about the appropriate role of DI or about which potential recipients are deserving, classification errors with respect to health are potentially a major concern.

Yet the value of receiving DI in a particular state, $s(h,n)$, is not entirely determined by the health shock: It is likely increasing in the severity of the non-health shock, n, as well. In terms of insurance, transfers into states in which marginal utility is high provide valuable insurance regardless of why marginal utility is high. It does not matter whether marginal utility is high due to a severe health condition or a severe non-health shock, such as to productivity, financial portfolios, or household composition. 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, whether due to health or non-health reasons.

Because of non-health risk, health is not a perfect indicator of the value of receiving DI in a particular state. The two perspectives occasionally "disagree" on whether benefits should be provided: Certain "correct" awards and non-awards reduce the value of DI and certain classification errors with respect to health increase the value of DI. Making transfers to a state of the world in which the individual, though healthy, has experienced a major non-health shock (e.g., productivity shock) can provide cost-effective insurance. Not making transfers to a state in which the individual, though unhealthy, has experienced advantageous realizations of non-health risk (e.g., high spousal earnings) can avoid providing cost-ineffective insurance.

Figure 1 shows how these two perspectives divide the state space into four mutually exclusive, exhaustive groups. It depicts a hypothetical joint $(h, s(h,n))$ distribution in the natural case

DRAFT

[5] In principle, having a severe health condition could be only weakly or even negatively correlated with surplus if non-health shocks were sufficiently negatively correlated with health shocks as to more than overcome the (partial) effect of health on surplus. Suppose surplus is linear in h and n: $s(h,n) = s_0 + \beta_h h + \beta_n n$, with $\beta_h, \beta_n > 0$. Then

$$E(s|S) - E(s|NS) = \beta_h[E(h|S) - E(h|NS)] + \beta_n[E(n|S) - E(n|NS)].$$

The value of receiving DI is greater in severe than in non-severe states on average as long as non-health shocks are not too strongly negatively related to severe health shocks, i.e., as long as

$$E(n|S) - E(n|NS) > -\frac{\beta_h}{\beta_n}[E(h|S) - E(h|NS)].$$

A positive association between non-health shocks and severe health shocks would further strengthen the association between severe health and surplus beyond the partial effect of health alone, since on average benefits in severe health states would partially insure non-health risk as well as health risk.

in which health is a strong but noisy indicator of surplus: $s(h, n)$ is strongly increasing in h on average but varies around that average because of variation in n. The two perspectives agree on providing benefits in Quadrant I, states in which the individual suffers a severe health shock and receiving DI produces positive surplus. They also agree on *not* providing benefits in Quadrant III (likely the vast majority of states), states in which the individual does not suffer a severe health shock and receiving DI produces negative surplus.

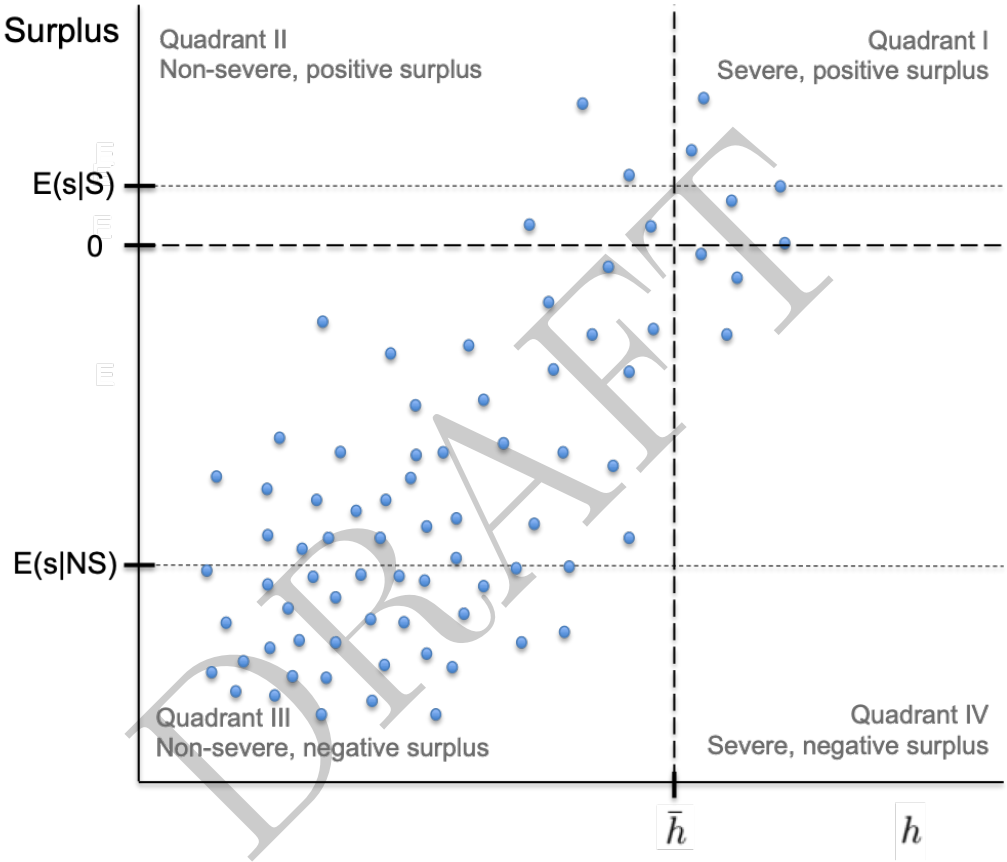
The two perspectives disagree, however, on whether to provide benefits in Quadrant IV, states in which the individual suffers a severe health shock but in which receiving DI produces negative surplus. In these states, the non-health shock is sufficiently favorable to make surplus negative despite the severe health shock. Awarding benefits in such a state is a correct award from the health-based perspective but decreases the value of DI. The perspectives also disagree on whether to provide benefits in Quadrant II, states in which the individual does not suffer a severe health shock but in which receiving DI produces positive surplus. In these states, the non-health shock is sufficiently unfavorable to make surplus positive despite the non-severe health shock. Awarding benefits in such a state is an error of inclusion from the health-based perspective but increases the value of DI.

In settings with substantial non-health risk, the value of DI and the costs of its classification errors depend crucially on the extent to which DI and its classification errors insure or worsen non-health risk, which in turn depends on selection into DI on the basis of non-health risk. Classification errors necessarily reduce the extent to which DI insures health risk—in this context, the risk of having a severe health condition—by reducing the extent to which DI transfers from non-severe to severe states. But classification errors could either insure or worsen non-health risk, making their net effect on DI's insurance value theoretically ambiguous. If individuals with more severe non-health shocks are more likely to select into DI conditional on health, then DI's classification errors insure (worsen) non-health risk.[6] Since classification errors reduce insurance against health risk, selection into DI on the basis of non-health risk conditional on health would have to be strongly favorable for classification errors to not reduce DI's overall insurance value (i.e., insurance against overall risk, including health and non-health). Assuming that health is a strong indicator of the value of receiving DI, random exclusion errors—the failure to provide benefits in states with mean surplus of $E(s(h, n)|S)$—would likely forego cost-effective insurance. Similarly, random inclusion errors—providing benefits in states with mean surplus of $E(s(h, n)|NS)$—would likely both

[6]More formally, DI's exclusion errors insure risk within the set of states with severe health conditions—risk that could reflect a mix of health and non-health risk as defined by h and n—if and only if mean marginal utility in severe recipient states exceeds that in severe non-recipient states, and DI's inclusion errors insure risk within the set of states with non-severe health conditions if and only if mean marginal utility in non-severe recipient states exceeds that in non-severe non-recipient states. The more uniform is health within severe and non-severe states, the greater the extent to which within-severe and within-non-severe risk reflects non-health risk n.

DRAFT

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



DRAFT

Notes: Figure shows how health-based and value-maximizing perspectives classify states of the world into those in which DI receipt is desired versus not-desired. The x-axis measures the severity of the health shock, h. From a health-based perspective, DI receipt is desired if and only if $h > \bar{h}$ ("severe"). The y-axis measures the ex ante surplus of receiving DI benefits. From the value-maximizing perspective, DI receipt is desired if and only if $s(h, n) > 0$. Both perspectives desire DI receipt in states in Quadrant I and do not desire receipt in states in Quadrant III, but differ in their assessment of states in Quadrants II and IV. This hypothetical figure depicts the natural case in which surplus is positively correlated with h across states.

worsen insurance (by transferring to states with below-average marginal utility) and involve substantial efficiency costs (by creating large labor supply distortions).

But DI's actual inclusion and exclusion errors may not be random: DI's tags and screens likely lead to systematic selection into applying for and being awarded DI. For example, the U.S. SSDI and SSI programs 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"). Some are chosen intentionally to try to improve targeting, such as the requirement that applicants be out of the labor force for at least five months before collecting benefits and that recipients cannot earn more than a low amount. Others are unintentional byproducts of program administration, such as the complicated application process and the uncertainty in award decisions. DI's award process also reflects mistakes from imperfect information—for example, an adjudicator awarding benefits to an applicant who appears to have a severe health condition but in fact does not.

The systematic selection created by DI's tags, screens, and mistakes could either increase or decrease the cost of DI's classification errors relative to representative errors. Classification errors could be more costly if DI's tags and screens disproportionately deter applications or reject applicants in states of the world in which the individual has a high marginal utility of income. For example, Deshpande and Li (2019) find that higher application costs disproportionately discourage individuals with severe conditions and limited resources from applying. In Figure 1, such an unfavorable exclusion error would be represented by the individual receiving DI in the severe states with the lowest surplus levels. Unfavorable inclusion errors would similarly be represented by the individual receiving DI in the non-severe states with the lowest surplus levels. On the other hand, classification errors could be less costly, even valuable, if DI's tags and screens instead deter applications or reject applicants in states of the world in which the individual has a low marginal utility of income. For example, DI's strict earnings limit might discourage the individual from applying in states with better earnings opportunities. In Figure 1, such a favorable exclusion error would be represented by the individual receiving DI in the severe states with the highest surplus (and similarly for favorable inclusion errors).

Whether DI's tags and screens create value-enhancing or -reducing selection is ultimately an empirical question, which we aim to answer in the following positive and normative analyses. In the positive analysis, we examine selection into DI on characteristics beyond health. In the normative analysis, we estimate the value of receiving DI in different types of states and quantify the effect of classification errors the value of DI.

DRAFT

3 Institutional Setting and Data

3.1 Background on U.S. disability programs

Like DI programs throughout the developed world, the two U.S. disability programs, Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI), are intended to provide assistance to people who cannot work because of a severe health condition. To qualify for SSDI, individuals must have worked for a certain number of quarters. Workers who become disabled between the ages of 30 and 42 years old must have 20 quarters of earnings to qualify for SSDI. This requirement increases by one quarter for every year above 42. Workers under 30 qualify with fewer credits. Individuals who do not meet the work history requirements for SSDI may be eligible for Supplemental Security Income (SSI), a means-tested program that has the same medical eligibility requirements as SSDI but does not require a work history.

The SSDI benefit is increasing in past earnings, at a decreasing rate. The average annual SSDI benefit is around $15,000, and beneficiaries qualify for Medicare after a two-year waiting period.[7] There is no household income or asset test, but disabled workers are not permitted to earn more than the substantial gainful activity (SGA) level, which was $1,260 per month in 2020.

Supplemental Security Income (SSI) has the same medical eligibility requirements and SSA adjudication process as SSDI, but different non-medical requirements. SSI requires low income and assets but has no work history requirements, while SSDI is not means-tested but requires a work history. The maximum monthly SSI benefit for an individual was $783 in 2020, and the SSI benefit is reduced by 50 cents for each dollar of earnings. Individuals who have a sufficient work history as well as low income and assets can qualify for both SSDI and SSI.

Since our paper estimates the effects of DI in the context of the broader social safety net, it is important to understand its interactions with other programs. In principle, unemployment insurance (UI) and SSDI are mutually exclusive since the former is conditioned on remaining in the labor force and the latter is conditioned on being out of the labor force. In practice it is possible that individuals receive both types of benefits, though this is uncommon.[8]

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

[8] Concurrent Receipt of Social Security Disability Insurance (SSDI) and Unemployment Insurance (UI): Background and Legislative Proposals (Morton, 2015).

DRAFT

3.2 Data

In Section 4, we use a combination of survey data and administrative data to study the characteristics of different groups by DI status and severity: consumption, adverse non-health events, and resources available to cope with shocks. The main survey data source for this section is the Panel Study of Income Dynamics (PSID), a panel survey of between 5,000 and 8,000 households. The PSID was annual from 1968 to 1997, has been administered every two years since then. The main advantage of the PSID for our context is the long panel structure, which allows us to go back in time to measure events and characteristics of households at the time of DI entry or disability onset. For nearly all measures in the positive analysis, we classify individuals by their DI status and health status in 2016 (measured in the 2017 wave), the latest available year, and then track them back to their entry onto DI and the onset of a severe health condition.

We use the PSID to measure the following characteristics, with the exact questions and response options reproduced in Appendix Tables A5 - A9:

- **Consumption.** We use the PSID's detailed consumption data, which includes food, housing and utilities, transportation, education, child care, and health care. We use total consumption excluding health care expenditures to avoid artificially inflating the consumption of individuals with severe health conditions.
- **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 worker reports as involuntary, namely layoffs or firings, firm closings or moves, and strikes. For moves, we limit the measure to "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.
- **Resources.** We 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, SSI, UI, and workers' compensation.

The PSID has important limitations: a relatively small sample size, self-reported DI status, and no information about applications for SSDI or SSI. To address these shortcomings, we supplement the positive analysis with SSA administrative data and with the Survey

DRAFT

of Income and Program Participation (SIPP) linked to SSA administrative data. We use SSA administrative data for measures of two types of adverse non-health events: financial events (bankruptcies, evictions, and foreclosures) and mass layoffs. First, we use the Master Beneficiary Record linked to nationwide financial records from Deshpande, Gross and Su (2019) to measure financial events such as bankruptcies, evictions, and foreclosures for DI recipients and applicants. Second, we use the Continuous Work History Sample, a 10% sample of all SSNs in the United States, to identify firms where mass layoffs occur, and link SSDI 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.[9]

We also use SIPP-SSA linked data to verify our main cross-section results and to observe SSDI and SSI applications. We use the 1996, 2001, 2004, 2008, and 2014 panels from the SIPP; each panel of the SIPP 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 allows us to observe applications in the SSA administrative data, which is important for understanding mechanisms behind classification errors. The main disadvantage of the SIPP is that each panel lasts only a few years, so it is not useful for following the same household for an extended period.

4 Positive Analysis: Non-Health Risk and Selection into DI

Our goal is to understand the nature of selection into DI, especially selection into DI on the basis of non-health shocks conditional on health. As discussed in Section 2, such selection determines the extent to which DI and its classification errors insure health and non-health risks and is a key determinant of DI's value. To that end, we characterize DI's classification error groups—non-severe recipients and severe non-recipients—on non-health dimensions. We focus on non-health characteristics likely to be predictive of the key determinants of the value of receiving DI: the marginal utility of income and counterfactual earnings.[10]

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

[10] The main text focuses on raw comparisons for simplicity and transparency. In the appendix, we report analogous results with different sets of controls for ability. Within-ability comparisons are relevant for the insurance and welfare values of DI, 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 A1 compares education level, our main measure of ability group, across DI-by-severity groups. Appendix Figure A2 shows all outcomes conditional on education and early-life (ages

DRAFT

We categorize individuals as having a severe condition (S) or not having a severe condition (NS) and start with a comparison between these groups. We further categorize individuals as DI recipients (DI) or non-recipients (NDI). Inclusion errors are indicated as NS-DI and exclusion errors as S-NDI. Our measure of disability severity in the PSID is based on self-reported work-limiting health problems. We classify individuals as "severe" if they report that their health condition limits "a lot" the amount of work they can do or if they "can do nothing." We classify them as "non-severe" if it limits "somewhat," "just a little," or "not at all," or if they report no health condition (see Appendix Table A5).[11] In the SSA administrative data, our measure of disability severity is the SSA adjudicator's judgment of the likelihood of medical improvement for DI recipients (not available for applicants). We classify DI recipients as "severe" if they receive a "medical improvement not expected" designation and as "non-severe" if they receive a "medical improvement expected" or "medical improvement possible" designation.

From the latest wave of the PSID (2017), the share of the working-age population in each disability-by-DI receipt group is the following:

- DI recipients with severe conditions (S-DI): 4.3%
- DI recipients without severe conditions (NS-DI): 6.1%
- Non-recipients with severe conditions (S-NDI): 3.4%
- Non-recipients without severe conditions (NS-NDI): 86%

Based on the PSID self-reports of severity, there are non-negligible classification errors with respect to health (and/or substantial measurement error): Among individuals reporting a severe condition, over one third (40%) report not receiving DI. Among individuals reporting DI receipt, more than half (53%) report not having a severe condition.[12] Given these magnitudes, the characteristics of DI's classification error groups are critical to the overall value of DI.

DRAFT

31-35) earnings. Appendix Figures B1-B6 show the main results using SSDI only.

[11] Measurement error is an obvious concern with self-reported severity, especially for the comparison of NS-DI to S-DI. One concern is that individuals who experience severe non-health shocks might be more likely to incorrectly report having a severe condition. In this case, our measures of non-health shocks for S-DI will be overstated, which works in the opposite direction of our results. Another concern is that DI recipients might be more likely to incorrectly report having a severe condition. In this case, some NS-DI will be mistakenly categorized as S-DI, and if this error is random, then it will work in the opposite direction of our results. If selection is non-random and non-severe recipients with the worst (best) non-health realizations are most likely to incorrectly report having a severe condition, then this will work in the opposite (same) direction of our results.

[12] Among NS-DI, 53% report having no work-limiting condition, 28% report having a health condition that limits "somewhat" the amount of work they can do, 17% "a little," and 2% "not at all."

Fact 1: On average, severe individuals are worse off than non-severe individuals on all observable dimensions—both health and non-health—but there is substantial variation.

We start by studying the characteristics of severe and non-severe individuals. On both health and non-health measures, severe individuals are consistently much worse off than non-severe individuals. For example, Figure 2 shows that average consumption for individuals with severe conditions is just one-half that of individuals without severe conditions. This pattern is consistent across all of the measures that we observe: Having a severe health condition is strongly associated with low living standards, low earnings, and non-health shocks, even after controlling for proxies for ability group.

Yet these strong average tendencies mask substantial variation within each severity group. For example, the 90-10 intervals of consumption reveal that some individuals without severe conditions have low consumption, while some individuals with severe conditions have high consumption. More generally, across different measures, we find that although health is strongly associated with better circumstances, it is far from a perfect indicator.

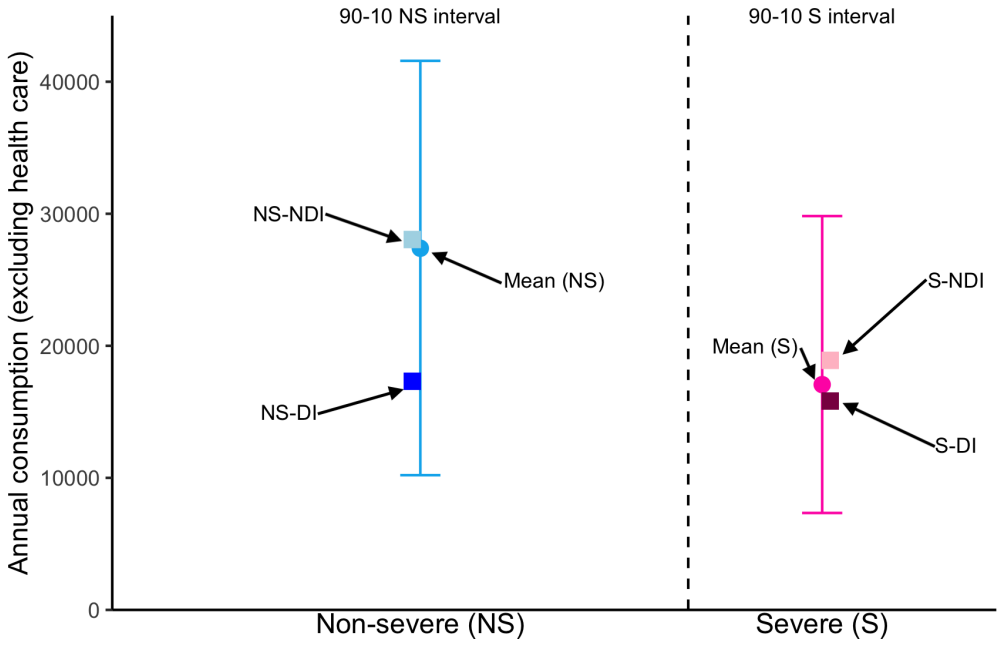
Implications for DI's value. These facts suggest that *representative* classification errors would reduce DI's insurance against not only health risk but also non-health risk, and so its overall insurance value. Severe individuals are on average worse off than non-severe individuals on all observable dimensions. Yet the substantial heterogeneity in non-health risk within severity groups means that DI's *actual* classification errors could potentially increase or decrease DI's overall insurance value. It is therefore important to go beyond health to understand the non-health characteristics of DI's actual classification error groups.

DRAFT

Fact 2: Non-severe recipients are much worse off than non-severe non-recipients, and by many non-health measures are even worse off than severe recipients.

To determine whether DI's actual inclusion errors are representative, we compare NS-DI to NS-NDI. We find that non-severe recipients are much worse off than the average non-severe individual. Figure 2 shows that NS-DI have an average consumption level that is just over one-half that of NS-NDI. This pattern is consistent across all of the measures that we observe: NS-DI are substantially worse off than other non-severe individuals, even after controlling for proxies for ability group (Appendix Figure A2).

Figure 2: PSID: Consumption range



Notes: Figure presents the 90-10 percentile interval (and average) of consumption for severe and non-severe households in the 2017 PSID: households with a severe health condition (566 S), and households without a severe health condition (6,755 NS). It also gives average consumption for households not currently receiving DI (SSDI or SSI) without a severe health condition (6,312 NS-NDI), households not currently receiving DI with a severe health condition (250 S-NDI), households currently receiving DI without a severe health condition (443 NS-DI), and households currently receiving DI with a severe health condition (316 S-DI). Household consumption excludes health care and is divided by the square root of household size. "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 "non-severe." Square markers for the four groups are offset to facilitate easier reading of the graph, not as a representation of relative severity.

In fact, NS-DI are so strongly selected that on non-health dimensions they are comparable to, and even sometimes worse off than, S-DI. Figure 2 shows that NS-DI and S-DI have similar current consumption levels. Of course, current consumption might reflect the effects of DI itself and so is not a good measure of selection into DI. For this reason, we go back to the year before each recipient entered DI and find that NS-DI are equally likely as S-DI to be at the bottom of the consumption distribution prior to receiving DI (see Appendix Figure A3).[13]

Figure 3 shows that S-DI and NS-DI have similarly high rates of job loss, distressed moves, and divorce in the PSID in the four years prior to DI receipt. For job loss, we include only separations that respondents report being involuntary, namely layoffs or firings. About 21% of NS-DI have experienced a head job loss and 14% a spouse job loss, compared to 18% and 8.5% for S-DI.[14] Distressed moves, which include evictions and foreclosures as well as a broader set of moves attributed to financial difficulties, are common: 41% for NS-DI and 39% for S-DI in the four years before DI entry. Divorce rates are also similar between the two: 4.9% for NS-DI and 6.2% for S-DI. Overall, the likelihood of experiencing any of these events in the four years prior to DI receipt is 57% for NS-DI and 50% for S-DI.[15]

We can also compare adverse non-health events for S-DI and NS-DI in the SSA administrative data. Figure 4 shows that NS-DI are more likely than S-DI to have experienced a mass layoff prior to entering DI. The mass layoff rate is 8.8% for NS-DI and 6.6% for S-DI in the three years before entry, and this difference remains even after conditioning on working.[16] Figure 4 also shows a similar pattern for financial events from SSA administrative data linked to nationwide financial records. NS-DI are more likely than S-DI to have experienced bankruptcy (3.5% vs. 3.0%), foreclosure (3.8% vs. 3.0%), and eviction (4.3% vs. 3.7%) in the years prior to receiving DI.[17]

In addition to adverse non-health events, we also analyze how NS-DI are selected on resources available to cope with adverse events. As shown in Figure 5, we find that NS-DI are less likely than S-DI to be married (28% vs. 33%) or to have a banking relationship (68% vs. 75%), and similarly likely to own a home (45% vs. 44%). To account for the possible causal

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

[14]The likelihood of job loss is roughly similar for NS-DI and S-DI after conditioning on employment.

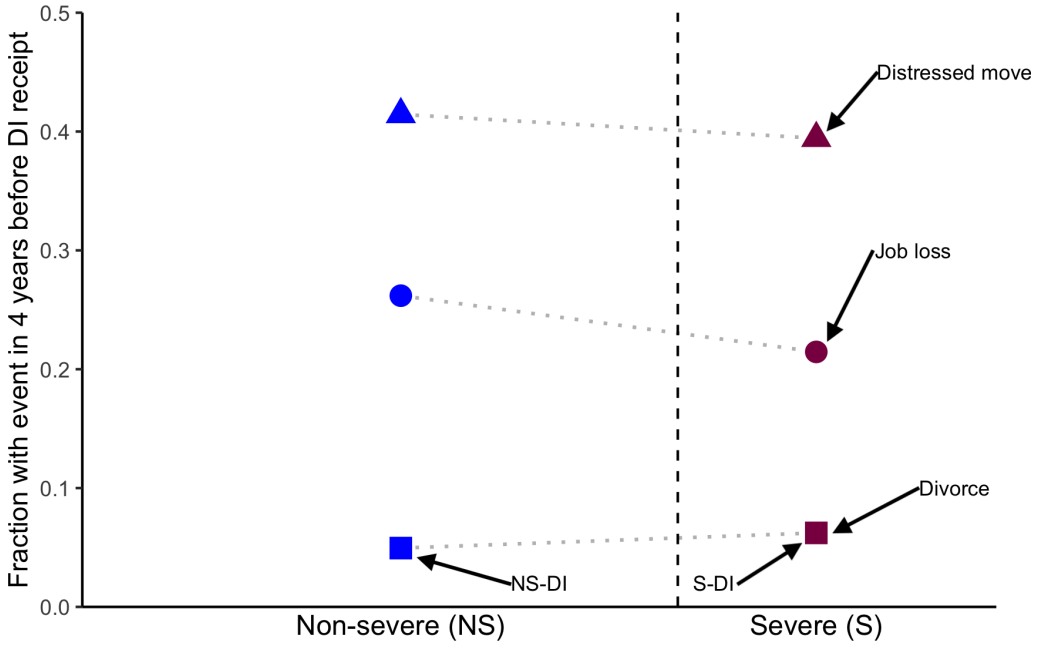
[15]These likelihoods far exceed the baseline level among non-recipients: at entry, DI recipients are 70% more likely than non-recipients of the same age to experience these adverse events (see Appendix Figure A4).

[16]Overall, mass layoff rates for DI recipients at entry are about twice as large as rates for non-recipients at the same ages (see Appendix Figure A5).

[17]In general, rejected applicants have even higher rates of these adverse financial events prior to application, consistent with a negative relationship between health risk and non-health risk in the applicant pool (Appendix Figure A6).

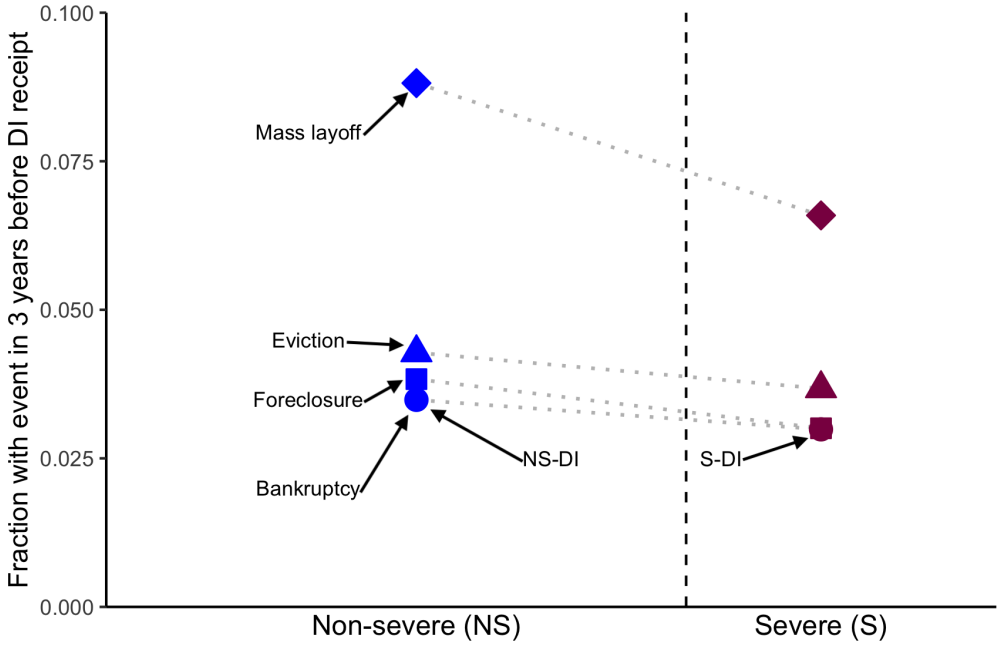
DRAFT

Figure 3: PSID: Adverse events prior to DI receipt



Notes: Figure presents rates of various adverse life events in the four years before receiving DI (SSDI or SSI) for households receiving DI in the 2017 PSID, by severe and non-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 A6). The sample for job loss is households currently receiving DI who can be observed in the four years before initial DI receipt (175 severe, 240 non-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 A7). The sample for distressed move is households currently receiving DI who can be observed in the four years before initial DI receipt (168 severe, 221 non-severe). Divorce is defined as being married in a previous survey year but not in this survey year (see Appendix Table A6). The divorce sample is households currently receiving DI who can be observed in the four years before initial DI receipt (162 severe, 210 non-severe). "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 "non-severe."

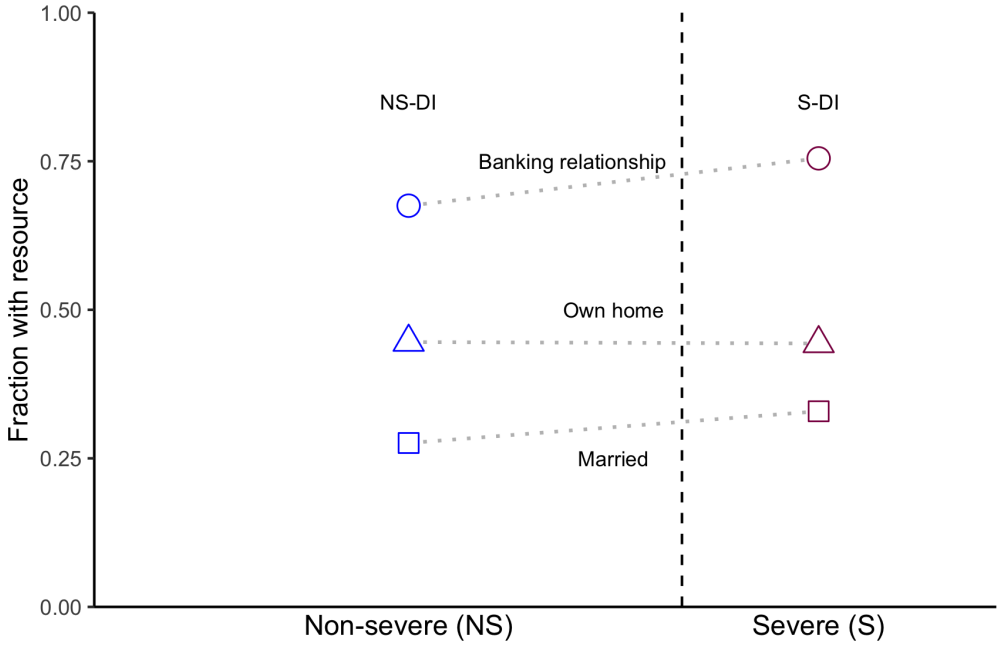
Figure 4: Administrative data: Adverse events prior to DI receipt



Notes: Figure presents rates of adverse financial events in the three years before receiving DI (SSDI or SSI) for individuals who apply for and receive DI benefits in SSA administrative data, by severe and non-severe. "Mass layoff" indicates being part of a mass layoff in the three years prior to DI application. The mass layoff sample is DI recipients in the CWHS-linked Master Beneficiary Record (MBR) who were allowed onto DI between 1985 and 2016 (520,000 less-severe, 940,000 more-severe). We determine rates of mass layoff using the Continuous Work History Sample (CWHS), which is a 10% sample of individuals in the US. In the CWHS, we identify EINs that experience a drop in employees of at least 30 percent from a base of at least 150 employees. From the MBR, we label a DI recipient as experiencing 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 DI application. We link DI applicants in the 831 files to nationwide financial records using the methods described in Deshpande, Gross and Su (2019). The bankruptcy sample is DI recipients in the 831 files between 1995 and 2009 (6.5M less-severe, 3.2M more-severe). The foreclosure sample is approved DI applicants in the 831 files between 2005 and 2014 who are homeowners (1.2M less-severe, 1.0M more-severe). The eviction sample is approved DI applicants in the 831 files between 2005 and 2014 who are not homeowners (1.4M less-severe, 770,000 more-severe). "Severe" indicates a medical diary of "medical improvement not expected," and "non-severe" indicates a medical diary of "medical improvement possible" or "medical improvement expected."

effect of DI on resources, we go back to the year before each recipient entered DI and find a similar pattern (Appendix Figure A7). Prior to DI receipt, NS-DI are also less likely to have either private or public health insurance (67% vs. 75%, from Appendix Figure A8), less likely to receive public transfers (37% vs. 40% for food stamps, from Appendix Figure A9), and less likely to receive help from relatives (23% vs. 25%, from Appendix Figure A10).

Figure 5: PSID: Outside resources

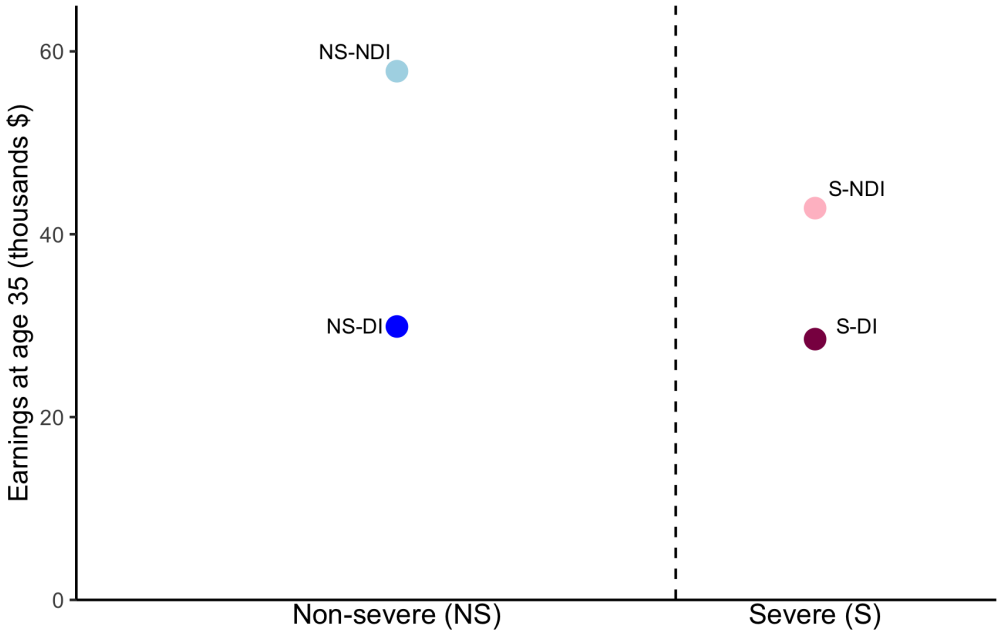


Notes: Figure presents rates of marriage, banking relationship, and homeownership for two types of households in the 2017 PSID: households currently receiving DI (SSDI or SSI) with a severe health condition (316 S-DI), and households currently receiving DI without a severe health condition (443 NS-DI). Appendix Table A8 presents the PSID questions for banking relationship and homeownership. "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 "non-severe."

The measures we have considered so far—consumption, education, adverse non-health events, and resources—are likely good predictors of the marginal utility of income, which is one of the key determinants of the value of receiving DI. We have fewer predictors of the other key determinant: counterfactual earnings. Figure 6 shows one possible such predictor: earnings at age 35, among the subset of individuals who were not on DI at age 35. Age 35 is well before most individuals enter DI but after at least some health and non-health shocks are realized, so earnings at that time could potentially be informative about (ex post) counterfactual earnings. Consistent with the other measures, the earnings at age 35, before ever entering DI, of individuals who later become NS-DI are only about one-half those of individuals who later become NS-NDI, and they are similar to those of individuals who later become S-DI.

How do NS-DI end up on DI in the first place? From the PSID, about 25% of NS-DI

Figure 6: PSID: Earnings at age 35



Notes: Figure presents average earnings at head age 35 for four types of households in the 2017 PSID: households not currently receiving DI (SSDI or SSI) without a severe health condition (1,923 NS-NDI), households not currently receiving DI with a severe health condition (90 S-NDI), households currently receiving DI with a severe health condition (93 S-DI), and households currently receiving DI without a severe health condition (152 NS-DI). Households are categorized based on their severity status and DI status in the 2017 PSID, and are included in the sample only if they are observed in the PSID at head age 35 and if they do not receive DI benefits at head age 35 or before. Results are qualitatively similar when we pool multiple years of the PSID to get a larger sample. "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 "non-severe."

had a severe condition when they applied, even though they do not currently have a severe condition. Another 15% of NS-DI received UI immediately before applying for DI, suggesting that they may have turned to DI as a form of permanent insurance following a qualifying job loss. Another 40% of NS-DI experienced some observable shock (job loss, distressed move, or divorce) before applying for DI but never received UI. The remaining 20% of NS-DI have no observable shock prior to applying for DI but have low resources.[18]

Implications for DI's value. These facts suggest that DI's inclusion errors likely increase DI's overall insurance value by providing valuable insurance against non-health risk. Non-severe recipients have experienced a variety of severe non-health shocks and appear likely to have an average marginal utility that is not only much higher than that of the average non-severe individual, but even comparable to that of the average severe recipient. Moreover, we find suggestive evidence of similar counterfactual earnings for severe and non-severe recipients.

Fact 3: Severe non-recipients are better off on non-health measures than both severe and non-severe recipients.

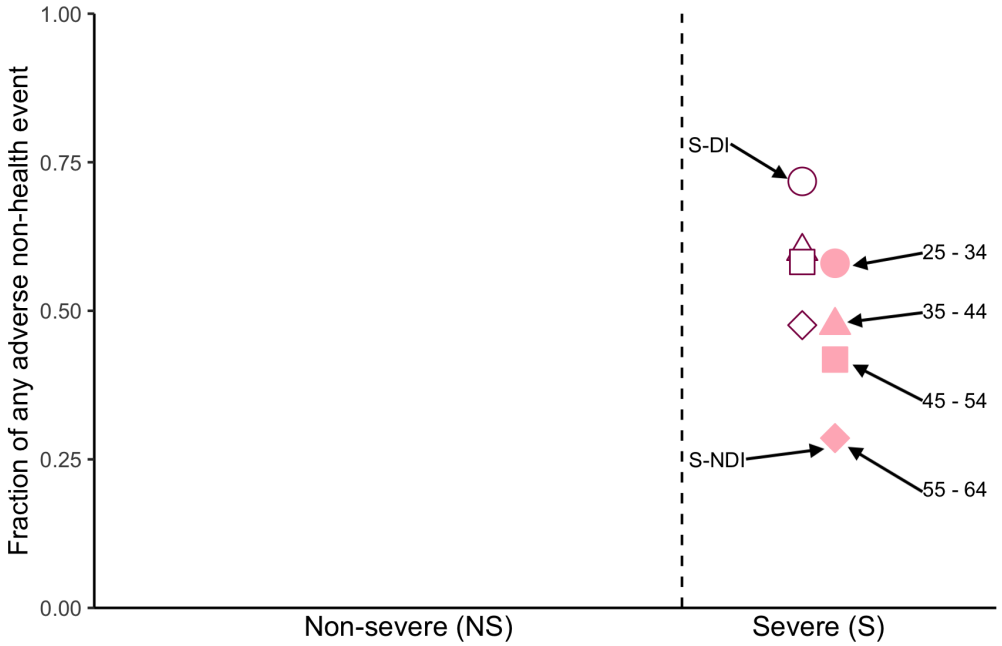
We turn next to DI's exclusion errors, severe non-recipients (S-NDI), to determine whether they are representative of all severe individuals. Comparing S-NDI to S-DI, we find that severe non-recipients are better off than the average severe individual. Figure 2 shows that S-NDI consumption is on average slightly greater than S-DI consumption, though still well below NS-NDI consumption. When we look at the distribution of consumption, we find that S-NDI are less concentrated at the bottom than S-DI or NS-DI (Appendix Figure A11). This pattern is true even when we go back to disability onset (see Appendix Figure A12).

This pattern is also apparent with respect to adverse non-health shocks and resources to cope with those shocks. Figure 7 shows the 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 "severe" side of the grid the rate of adverse non-health events for S-NDI at that age and for S-DI who enter DI in that age range at that age. We find that S-NDI are 10-30% less likely at every age range to have experienced an adverse non-health event. Appendix Figures A13–A16 show results for head and spouse job loss, distressed move, and divorce separately.

[18]From the 2014 SIPP, among NS-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.

DRAFT

Figure 7: PSID: Any adverse life event by DI status and age



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 DI (SSDI or SSI) status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI) and their age at entry, and NDI are further classified by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The figure plots the rate of experiencing an adverse life event in a specific age range (25–34, 35–44, 45–54, 55–64) among NDI and among DI who also entered DI at that age. The sample sizes for age groups 25–34, 35–44, 45–54, and 55–64 are 8,023, 7,098, 4,424, and 2,735, respectively, for NS-NDI; 896, 1199, 1054, and 848 for S-NDI; 312, 272, 271, and 354 for NS-DI; and 97, 211, 243, and 348 for S-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 A6). 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 A7). Divorce is defined as being married in a previous survey year but not in this survey year (see Appendix Table A6). "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 "non-severe." Markers for the two groups are offset to facilitate easier reading of the graph, not as a representation of relative severity.

Turning to outside resources, we find that S-NDI are more advantaged than S-DI in the year of disability onset (Appendix Figure A17). S-NDI are more likely than S-DI to be married (49% vs. 35%), to have a banking relationship (66% vs. 58%), and to own a home (47% vs. 45%).

Finally, we use earnings at age 35 among the subset of individuals who were not on DI at age 35 as a predictor of counterfactual earnings. From Figure 6, the age-35 earnings of individuals who later become S-NDI are almost 50% greater than of those who later become S-DI, though still substantially lower than of those who become NS-NDI.

Why do S-NDI not receive DI? One important question is whether these individuals never apply, or whether they apply and get rejected. From the SIPP-SSA linked data, we find that about 45% of S-NDI can be explained by either measurement error in the survey data or mistakes in the decision process: individuals with severe conditions who applied but were rejected by SSA on medical grounds. The remaining 55% of S-NDI do not apply. This appears to be largely insurance-promoting selection out of DI: relative to S-NDI who apply, S-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%). Clearly, however, there are still a nontrivial fraction of S-NDI who do not apply despite being disadvantaged.[19]

Implications for DI's value. These facts suggest that DI's exclusion errors increase its insurance against non-health risk but likely decrease its overall insurance value. Severe non-recipients are better off on non-health measures than severe recipients, meaning that exclusion errors insure non-health risk. But since they are still much worse off than non-recipients as a whole, exclusion errors reduce DI's overall insurance value.

[19] Among S-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).

DRAFT

5 Normative Analysis: Estimating the Value of DI and Its Classification Errors

5.1 Theory: Value of receiving DI in different states

Our goal is to estimate the value of receiving DI benefits in different states of the world: the extent to which the ex ante value of different groups of actual or hypothetical DI benefits exceeds the costs of paying for them through higher net taxes, including costs from the distortions created by DI receipt. This value reflects the classic tradeoff between the insurance value of transfers to high marginal utility states and the efficiency cost of behavioral responses.

We start by constructing the ex post value and cost of receiving DI in a particular state of the world. The state of the world reflects the realization of all risks, health and non-health (e.g., no severe disability and few outside resources). We use these ex post values and costs to construct the ex ante values and costs of a given set of DI benefits to a given "risk type"—individuals that face a particular distribution of possible ex post states of the world.

DRAFT

Ex ante value of receiving DI in states Ω_b: $EAWTP_\theta(\Omega_b)$. We start with the *ex post* marginal value of a DI-like expansion of the ex-post budget constraint in state of the world ω:

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

where $z_\omega^{DI=0}$ is counterfactual earnings in state ω, $\bar{z}$ is the statutory earnings limit for DI recipients, and b is the DI benefit. The indicator function indicates whether earnings in state ω would exceed the statutory earnings limit if, counterfactually, the individual did not receive DI and so was not subject to the earnings test. This is a conservative approximation that tends to understate the value of a DI-like expansion of the budget constraint.[20]

Ex ante "risk type" θ faces a given set of possible ex post states of the world, Ω_θ. We wish

[20]This is conservative in that it tends to overstate the cost to DI recipients of the earnings test and so understate the value of DI. It assumes that the ex post value of receiving DI is zero in states in which the individual would have earned more than the earnings limit in the absence of DI. A less conservative bound would essentially allow free disposal of earnings. Imagine someone who would earn $1 more than the earnings limit without DI telling their boss, "I will continue to work exactly as before, just pay me $1 less." They would lose $1 of earnings but gain the entire DI benefit, a large net gain. Such a free-disposal assumption corresponds to $WTP_\omega = \max\{0, b - (z_\omega^{DI=0} - \bar{z})\}$, which weakly exceeds the WTP_ω implied by equation (1). We use equation (1) both to be conservative about the value of DI and to be consistent with Hendren and Sprung-Keyser (2019) for comparability (though unlike Hendren and Sprung-Keyser (2019), we focus on DI's ex ante, not ex post, value).

to assess the ex ante value of θ hypothetically receiving DI in states $\Omega_b \subseteq \Omega_\theta$.[21] We measure the value of DI benefits to the individual in terms of income in states $\Omega_\tau \subseteq \Omega_\theta$. The *ex ante* value to type θ, in terms of income in states Ω_τ, of receiving DI in states Ω_b is the $EAWTP_\theta(\Omega_b)$ that solves

$$\underbrace{p(\Omega_b|\theta)E(\lambda_\omega \times WTP_\omega|\theta,\Omega_b)}_{\text{MB of receiving DI in states } \Omega_b} = \underbrace{p(\Omega_\tau|\theta)E(\lambda_\omega|\theta,\Omega_\tau)EAWTP_\theta(\Omega_b)}_{\text{MC of giving up } EAWTP_\theta(\Omega_b) \text{ in states } \Omega_\tau}, \tag{2}$$

where $p(\cdot)$ is the probability operator, λ_ω is marginal utility of income in state ω, and here, with some abuse of notation, Ω_X refers to the event that the realized state ω is in the set of states Ω_X ($\omega \in \Omega_X$). The left-hand-side is the ex ante marginal benefit of receiving a DI-like expansion of the ex post budget constraint 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 terms yields

$$EAWTP_\theta(\Omega_b) = \underbrace{\frac{p(\Omega_b|\theta)}{p(\Omega_\tau|\theta)}}_{\pi} E\left[\underbrace{\left(\frac{\lambda_\omega}{E(\lambda_\omega|\theta,\Omega_\tau)}\right)}_{(1+M_\omega)} WTP_\omega \middle| \theta, \Omega_b\right]. \tag{3}$$

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 DI, WTP_ω, is one plus the "markup" on transfers from Ω_τ to ω:

$$M_\omega \equiv \frac{\lambda_\omega - E(\lambda|\theta,\Omega_\tau)}{E(\lambda|\theta,\Omega_\tau)}. \tag{4}$$

The markup is the amount by which marginal utility in ω exceeds mean marginal utility in the states used to value DI (Ω_τ), as a share of the latter. One dollar worth of expected income in ω is worth \$(1+M_ω) worth 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 DI in states Ω_b: $EAG_\theta(\Omega_b)$. The cost to the government of providing DI in a particular state includes not only the cost of the DI benefit itself but also "fiscal externalities" on other aspects of the government budget from any induced behavioral responses, such as reductions in the labor supply of—and so net taxes paid by—the recipient's household. The *ex post* net cost to the government of

[21] The main sets of states Ω_b we will assess are the four sets of states defined by severity-by-DI receipt that we have emphasized throughout: severe recipient (S-DI), non-severe recipient (NS-DI), severe non-recipient (S-NDI), and non-severe non-recipient (NS-NDI). Of course, the NS-DI and NS-NDI states are ones in which the individual does not receive DI in practice.

DRAFT

providing DI benefits b in state ω is

$$G_\omega = b + fe^z_\omega, \tag{5}$$

where fe^z_ω is the earnings component of the fiscal externality: the loss in net tax revenue from the earnings response to DI in state ω. Aggregating over ex post states, the *ex ante* expected cost to the government of providing DI to type θ in states Ω_b, per state of the world in which type θ pays associated taxes, is

$$EAG_\theta(\Omega_b) = \pi E(G_\omega|\theta, \Omega_b).^{22}$$

Ex ante surplus from type θ receiving DI in states Ω_b: $s_\theta(\Omega_b)$. Ex ante surplus from type θ receiving DI in states Ω_b is

$$s_\theta(\Omega_b) = g_\theta \times EAWTP_\theta(\Omega_b) - EAG_\theta(\Omega_b),$$

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)).[23] This compares the cost to the government, accounting for behavioral responses, of providing $EAWTP_\theta(\Omega_b)$ of surplus to θ in two different ways: DI benefits in Ω_b states and income tax cuts in Ω_τ states. 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 DI benefits in Ω_b states. The difference, $s_\theta(\Omega_b)$, therefore measures the surplus from type θ receiving DI 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 DI to type θ in states Ω_b and (ii) raising taxes on θ exactly enough to leave θ indifferent between this hypothetical combined DI-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 DI benefits in Ω_b states than by cutting θ's income taxes in Ω_τ states. The surplus from DI benefits in Ω_b states relative to income tax cuts in Ω_τ states is greater when: the value of DI benefits in Ω_b states is greater (greater $EAWTP$), the efficiency cost of DI benefits in Ω_b states is smaller (smaller EAG), and the efficiency gain from cutting income taxes in Ω_τ states is smaller

[22]The fiscal externality, FE—the increase in the cost to the government from behavioral responses as a share of the mechanical effect ME—is $FE(\Omega_b) = \frac{EAG_\theta(\Omega_b)}{ME_\theta(\Omega_b)} - 1 = \frac{[1-p(z^{DI=0}_\omega \leq \bar{z}|\theta,\Omega_b)]b + E(fe^z_\omega|\theta,\Omega_b)}{p(z^{DI=0}_\omega \leq \bar{z}|\theta,\Omega_b)b} - 1$.

[23]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 (Hendren, 2020).

DRAFT

(greater g_θ).[24,25]

Total social surplus to all types: $s(\{\Omega_b(\theta)\}_{\theta\in\Theta})$. Total social surplus from providing DI benefits in each type θ's states $\{\Omega_b(\theta)\}_{\theta\in\Theta}$, in net government revenue per taxpayer, is the mean surplus across types,

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

Intuitively, this approach measures the value of a policy as the difference in the costs to the government of providing the same value to the affected types via net tax cuts versus via the policy in question. A policy generates social surplus to the extent that it provides value to the affected types more cheaply than the income tax and transfer system could.[26]

Assumptions and limitations. For tractability, this analysis aims to provide first-order approximations to the value of DI *benefits* rather than exact estimates of the value of DI itself. The latter would include additional costs of administering, applying for, and receiving DI, including any hassle and stigma costs. Even setting aside the differences between DI benefits and DI as a whole, this analysis does not speak directly to the welfare effects of policy changes. Our analysis asks whether a value-maximizing planner who had perfect information about the realized state of the world would choose to provide DI in a given state. Policy changes, in contrast, cannot be implemented state-by-state due to information and incentive compatibility constraints; they can only use tags and screens to roughly target sets of states. One implication of these constraints is that most policy changes would affect applications to DI, and any such application response is a potentially important determinant of the costs

DRAFT

[24]The smaller the efficiency gain from cutting income taxes, the more costly it is to the government to help the individual via income tax cuts and so the better is DI in comparison. The efficiency cost (marginal excess burden) of income taxes—and so the efficiency gain from cutting them—is negatively related to the inverse-optimum weight g_θ: The greater the efficiency cost of taxes, the less revenue the government would raise from a tax hike relative to the burden imposed on the individual, and so the less revenue the government would forgo from a tax cut that conferred a given benefit on the individual (smaller g_θ).

[25]The following are equivalent: These DI benefits generate positive surplus; these DI benefits are more valuable to θ ex ante than cost-equivalent tax cuts; the ex ante MVPF of these benefits ($EAWTP_\theta(\Omega_b)/EAG_\theta(\Omega_b)$) exceeds that of cutting θ's taxes in Ω_τ states ($1/g_\theta$); ex ante, θ would benefit from a government budget-neutral reform that provided these DI benefits while at the same time raising taxes in Ω_τ states enough to finance them.

[26]By expressing surplus to different types in the same units—net government revenue—the inverse-optimum weights approach enables one to calculate the efficiency consequences of a policy that affects multiple types without specifying a social welfare function or exactly how it is financed (provided it is financed through the income tax system) and accounting for the efficiency costs of the income tax and transfer system. Provided that the policy is financed through changes in the income tax and transfer system, the actual way in which it is financed is immaterial to this calculation. Of course, the way in which it is financed is relevant for the distributional consequences of the policy and so its social welfare effect, an issue we return to in Section 5.4.

of a policy change. In Appendix Section B, we consider three frequently discussed policy reforms taking into account the limited evidence on application responses: reducing inclusion errors, decreasing benefit levels, and decreasing the allowance rate.

Like all welfare analyses, our analysis relies on key assumptions for mapping observable characteristics into unobservable welfare: optimization, no externalities other than fiscal externalities, and the functional form of marginal utility. That people optimize allows us to use the envelope theorem to quantify, to first order, both the ex post value of a DI-like change in constraints based on counterfactual earnings and the ex ante value of any such ex post changes based on the marginal utility of income in different states. That there are no externalities other than fiscal externalities means that willingness to pay for DI comes entirely from recipients.

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 DI. The level of risk aversion determines the magnitude of the gap in marginal utilities—and so the estimated markup—for any given gap in consumption levels. 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 versa). Using the wrong level of risk aversion will obviously bias the calculations. Any state dependence in utility could also affect the markup. In Appendix Table A1, we report results based on a wide range of assumptions about risk aversion and state dependence.

DRAFT

5.2 Empirical implementation and results

Implementation. Our aim is to quantify, for different risk types θ, the ex ante surplus from receiving DI, $s_\theta(\Omega_b)$, in each of four sets of states of the world: $\Omega_b = \Omega_{h\text{-}di} \in$ {severe recipient (S-DI), non-severe recipient (NS-DI), severe non-recipient (S-NDI), and non-severe non-recipient (NS-NDI)}. We measure the value of DI benefits to the individual in terms of income in states in which they do not receive DI, $\Omega_\tau = \Omega_{NDI}$. 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 a given 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 A1).

For each θ-by-$\Omega_{h\text{-}di}$ group, we estimate the components of ex ante willingness to pay for DI benefits in those states from equation (3): the no moral hazard actuarially fair value, π, the markup on transfers to each state from NDI states, M_ω, and the ex post marginal value of DI in each state, WTP_ω. We also estimate the mean cost to the government of providing DI in those states, $E(b + fe^z_\omega)$ from equation (5). The key building blocks of these calculations are counterfactual earnings with and without DI ($z_\omega^{DI=1}$ and $z_\omega^{DI=0}$) and marginal utility (λ_ω).

To construct counterfactual no-DI earnings ($z_\omega^{DI=0}$) for individuals that actually receive DI, we use French and Song's (2014) quasi-experimental estimates of the effects of DI on the earnings of individuals with different types of health conditions. In particular, we assume that counterfactual no-DI earnings for individuals who receive DI are observed (with-DI) earnings plus French and Song's (2014) estimated mean reduction in earnings due to DI by detailed health condition. We assume that counterfactual *with-DI* earnings ($z^{DI=1}$) for individuals who do not receive DI are the smaller of predicted counterfactual earnings based on French and Song (2014) and the DI earnings limit. These assumptions tend to work against our main results, understating the value of DI benefits for individuals who actually receive DI and overstating the value of DI benefits for individuals who do not receive DI. The key conclusions are robust to making assumptions that are even more conservative (see Appendix Table A1).

To construct marginal utility (λ_ω), we assume that marginal utility is a power function of consumption, $\lambda = c^{-\gamma}$ (constant relative risk aversion utility), where γ is the coefficient of relative risk aversion. Our measure of household consumption comes from the PSID. Since 1998, the PSID has included detailed consumption information on food, housing and utilities, transportation, education, child care, and health. We use total household consumption excluding health care, since including health care consumption could artificially inflate the consumption level of individuals with 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 leads us to overstate the counterfactual no-DI consumption of DI recipients, and so to understate their no-DI marginal utility and the ex ante value of receiving DI in that state. Our baseline specification uses a coefficient of relative risk aversion of two and state-independent marginal utility. We test robustness to alternative plausible values of risk aversion and state-dependent utility and find that our key conclusions

DRAFT

are highly robust (see Appendix Table A1).[27]

For the other parameters in the model, we approximate their actual values in the U.S. or adopt estimates from other sources. All values are annual. The statutory DI earnings limit, $\bar{z}$, is the annualized value of the Social Security Administration's earnings limit for DI recipients ($1,260 per month in 2020, or $15,120 per year). The DI benefit, b, is $13,000, roughly the average cash benefit of SSDI and SSI recipients.[28] The earnings component of the fiscal externality, fe^z_ω, is based on Hendren and Sprung-Keyser's (2019) estimate of the average marginal net tax and transfer rate on the reduction in earnings from receiving DI of 20%. 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.

Empirical estimates of the value of DI benefits in different states. Figure 8 shows the empirical analogue of the theoretical Figure 1. For each of the four sets of DI-by-severity states of the world, we plot the fraction severe on the x-axis and the total social surplus from receiving DI in those states, s, on the y-axis. The area of the marker is proportional to the probability of those states. Table 1 reports related estimates.

The results reveal two key findings. First, DI generates substantial ex ante surplus. The states of the world in which an individual would receive DI (S-DI and NS-DI) have a large surplus from DI ($8,700 on average per recipient state per year), meaning that the average DI benefit is highly valuable ex ante. We estimate that DI produces an annual per capita ex ante surplus of $920 worth of government revenue. This means that were the government to abolish DI, the tax cuts necessary to leave individuals unharmed relative to the status quo with DI would cost the government $920 per capita per year more than what it saved from abolishing DI. In other words, DI benefits as a whole are more valuable to individuals ex ante than tax cuts with the same cost to the government.

Second, DI's classification errors are not costly—they are highly valuable. On average, health is a strong indicator of the value of receiving DI: Receiving DI in the average severe state generates substantial surplus ($5,000 on average), whereas receiving DI in the average non-severe state is extremely costly (−$16,800 on average). But DI's classification errors

[27]Note that marginal utility of consumption can be state-independent even if utility is state-dependent, e.g., if the disutility of labor is state-dependent in the way it is in Diamond and Sheshinski (1995), $U = u(c) - v(l;\theta)$, where the state θ may include health status. 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.

[28]This does not include any value of the health insurance component of the benefit. We exclude the health insurance component because French and Song (2014) exclude the value of health insurance from the denominator of their labor supply elasticity calculations.

DRAFT

Table 1: Value of DI benefits in different states

	$E(M_\omega)$	Counterfactual earnings ($)	Surplus per capita ($) $(p(\Omega_\tau)s)$	Surplus per recipient ($) $(p(\Omega_\tau)s/p(\Omega_b))$
DI	0.91	5,191	924	8,718
S-DI	0.95	3,557	488	9,858
NS-DI	0.88	6,623	436	7,719
NDI	0.00	57,476	-15,930	-17,819
S-NDI	0.64	22,591	-74	-2,203
NS-NDI	-0.02	58,838	-15,856	-18,429
Full pop	0.10	51,932	-15,005	-15,005
S	0.82	11,252	414	4,982
NS	0.03	55,620	-15,420	-16,818

Notes: Table presents mean markup (from equation (4)), counterfactual earnings, social surplus per capita, and social surplus per recipient associated with receiving DI benefits in each of several sets of states of the world: DI recipients (DI), including severe recipients (S-DI) and non-severe recipients (NS-DI); non-recipients (NDI), including severe non-recipients (S-NDI) and non-severe non-recipients (NS-NDI); and the full population (Full pop), including severe individuals (S) and non-severe individuals (NS). Surplus s (from equation (6)) is in units of government revenue per Ω_τ state, so $p(\Omega_\tau)s$ is in units of government revenue per capita and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient. Monetary amounts are in dollars per year.

DRAFT

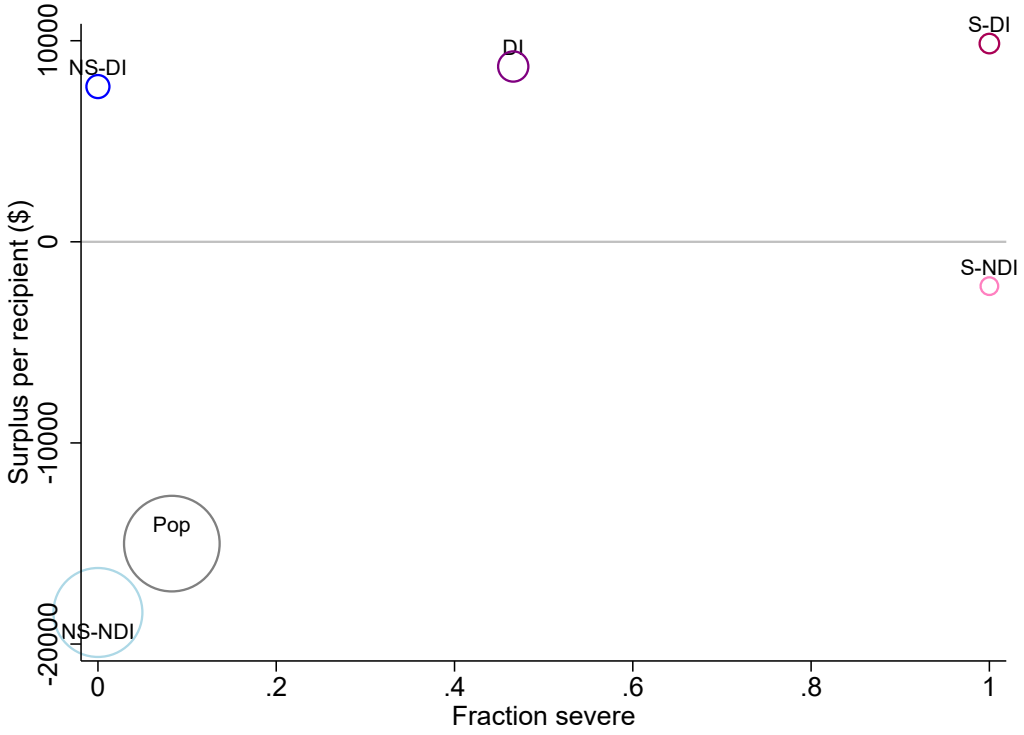
actually increase DI's overall value. From Figure 8, the ex post states of the world in which an individual would receive DI despite not having a severe condition (NS-DI, inclusion errors) have a large surplus from DI ($7,700 on average), while the ex post states of the world in which an individual would not receive DI despite having a severe condition (S-NDI, exclusion errors) have a negative surplus from DI (−$2,200 on average).

On net, DI's classification errors produce an annual per capita ex ante surplus of $510 worth of government revenue. This is mostly attributable to DI's inclusion errors: Inclusion errors produce a per capita ex ante surplus of $440, or 47% of DI's overall surplus of $920. In fact, DI's inclusion errors (NS-DI) appear to be nearly as valuable as DI's "correct" awards (S-DI). These broad patterns hold not only overall, summing across all of the different risk types in the population, but also for each risk type individually (see Appendix Table A2). They also hold under a wide range of assumptions about the key ingredients of the calculations (see Appendix Tables A1 and A3). These results indicate that classification errors are highly favorably selected, consistent with the findings from the positive analysis that classification errors are favorably selected on a wide variety of non-health measures.

Figure 9 decomposes that selection on surplus into its two key determinants: marginal utility and counterfactual earnings. DI recipients as a whole, and non-severe recipients (NS-DI) in particular, are highly selected in terms of both marginal utility and counterfactual earnings,

and so in terms of both insurance benefit and distortion cost. The markup on transfers from non-recipients to NS-DI is 0.88, about the same as the markup to S-DI 0.95. (The markup on transfers from non-recipients to non-recipients is 0 by construction.) Mean counterfactual annual earnings among NS-DI is $6,600, versus $3,600 among S-DI and $57,500 among non-recipients. Importantly, the large value of NS-DI benefits is driven not only by low consumption but also by low counterfactual earnings in those states. Selection into DI on both dimensions is so strong that awards to both S-DI and NS-DI would generate positive surplus even if markups were just 25 percent of their estimated values.

Figure 8: Surplus per recipient relative to fraction severe

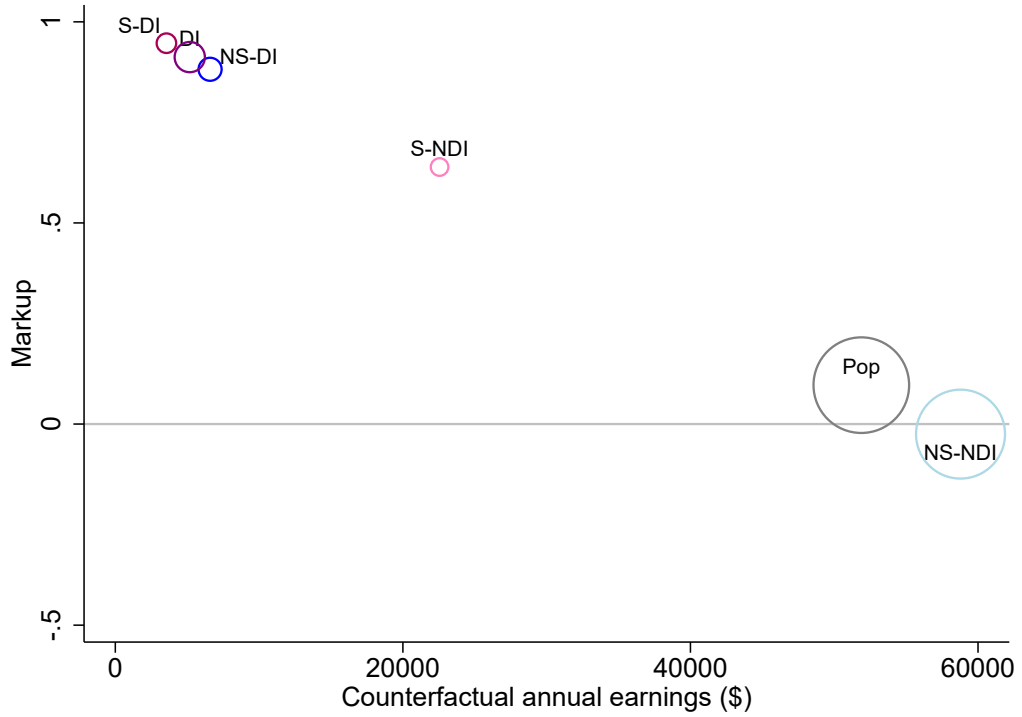


Notes: Figure plots surplus per recipient in dollars of government revenue per year, $p(\Omega_\tau)s/p(\Omega_b)$, by fraction severe for full population, DI recipients (DI), severe DI recipients (S-DI), non-severe recipients (NS-DI), severe non-recipients (S-NDI), and non-severe non-recipients (NS-NDI). Marker areas are proportional to population shares.

5.3 Key drivers of DI's value: Non-health risk and selective applications

Non-health risk as a driver: About one-half of DI's value comes from insuring risks other than that of having a severe health condition. We investigate the connections between DI's classification errors and its insurance against health and non-health risk with a decomposition analysis. For a given set of actual or hypothetical DI benefits in states $\Omega_b \subseteq \Omega_\theta$, decompose the benefit received in a particular state ω into the sum of (i)

Figure 9: Markups on transfers to, and counterfactual no-DI earnings of, various groups



Notes: Figure plots the mean markup on transfers from non-recipients (NDI) to different Ω_b states, $E(M_\omega)$, against counterfactual no-DI annual earnings in those Ω_b states, where Ω_b is the full population, severe DI recipients (S-DI), non-severe recipients (NS-DI), severe non-recipients (S-NDI), and non-severe non-recipients (NS-NDI). The markup on transfers to state ω, M_ω, is defined in equation (4). Counterfactual no-DI earnings are actual observed earnings for NDI groups and are inferred from French and Song's (2014) estimates of the effects of DI on earnings for DI groups, as described in the text. Marker areas are proportional to population shares.

the mean benefit received in states in the same health category as ω, $h_\omega \in \{\text{NS, S}\}$, 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):

$$b_\omega = \underbrace{E(b|h = h_\omega)}_{\text{Health-contingent benefit}} + \underbrace{[b_\omega - E(b|h = h_\omega)]}_{\text{Within-health transfer}}. \tag{7}$$

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

$$M_\Omega(\Omega_b) = \underbrace{Cov_h\left[\frac{p(\Omega_b|h)}{p(\Omega_b)}, E\left(\widehat{\lambda}|h\right)\right]}_{\text{Insurance against health risk}} + \underbrace{E_h\left\{\frac{p(\Omega_b|h)}{p(\Omega_b)}[1 - p(\Omega_b|h)]\left[E\left(\widehat{\lambda}|\Omega_b, h\right) - E\left(\widehat{\lambda}|\sim\Omega_b, h\right)\right]\right\}}_{\text{Insurance against non-health risk}}, \tag{8}$$

where $\widehat{\lambda} \equiv \frac{\lambda}{E(\lambda)}$ is normalized 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 A.[29]

The overall markup, $M_\Omega(\Omega_b)$, is the sum of two components. The "insurance against health risk" component comes from the health-contingent benefits in equation (7). This component is from insuring the risk of having a severe health condition ("health risk"). The more concentrated are benefits in states in which the individual has a severe health condition (greater $p(\Omega_b|S)$ relative to $p(\Omega_b|NS)$), the greater the insurance of health risk. This component would be maximized by a hypothetical program that perfectly targeted states in which the individual has a severe health condition—i.e., that had $\Omega_b = S$. It is necessarily reduced by classification errors with respect to health, since by definition such errors reduce the targeting of severe states.

The "insurance against non-health risk" component comes from the within-health transfers in equation (7). This component is from insuring "non-health risk," i.e., risk *within* health categories.[30] The greater the marginal utility gap between recipient and non-recipient states *within* severe and non-severe states, (greater $E\left(\widehat{\lambda}|\Omega_b, h\right) - E\left(\widehat{\lambda}|\sim\Omega_b, h\right)$), the more valu-

[29] As in the previous section, we calculate markups on transfers within each education group and then take the mean across education groups. Unlike that markup, which measures the value of benefits in terms of income in NDI states, this markup measures value in terms of income in all states of the world: $M_\Omega(\Omega_b;\theta) \equiv \frac{E(\lambda|\theta,\Omega_b) - E(\lambda|\theta)}{E(\lambda|\theta)}$. We do this because it simplifies equation (8) and facilitates comparison across different groups of benefits. It also leads to somewhat smaller overall markups.

[30] We term this component "non-health risk" partly as a shorthand for "risk other than that of having a severe health condition" 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 severe health condition, part of it is included in the "health risk" component).

DRAFT

able the insurance of non-health risk.[31] The "insurance against non-health risk" component is zero in the absence of classification errors. Classification errors make this component positive (negative) if (appropriately-weighted) selection into DI on marginal utility conditional on health is positive (negative).

Table 2: Decomposition of markup on transfers to various groups

	Surplus per recip ($) $(p(\Omega_\tau)s/p(\Omega_b))$	Markup, $E(M_\Omega)$	Contrib from insuring:		Share from insuring:	
			Health	Non-health	Health	Non-health
DI	8,718	0.75	0.28	0.48	37%	63%
S-DI	9,858	0.78	0.66	0.12	85%	15%
NS-DI	7,719	0.73	-0.06	0.79	-8%	108%
S-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%

Notes: Table presents social surplus per recipient, overall markup, and markup components and shares from insurance of health and non-health risk for different types of actual and hypothetical benefits, defined by the row. Contributions of insurance of health and non-health risk to overall markup are defined in equation (8). 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. S-DI are DI's "correct" awards. NS-DI are DI's inclusion errors. S-NDI are DI's exclusion errors. "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 with the same adjudication standards as DI but for which everyone applies, rather than the actual selective applications to DI. 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. The contribution to the overall markup from insurance of health risk is increasing in the extent to which receipt of the benefit in question (as defined by the table row) is concentrated in severe rather than non-severe states and is positive if and only if receipt is greater in severe than in non-severe states. The contribution to the overall markup from insurance of non-health risk is increasing in the extent to which receipt of the benefit in question (as defined by the table row) is concentrated in higher- rather than lower-marginal utility states conditional on health and is positive if and only if an appropriately-weighted average of the marginal utility gap between recipient and non-recipient states is positive.

DRAFT

Table 2 shows the results of this decomposition. The insurance that DI provides against the risk of having a severe health condition is highly valuable: It contributes 28 percentage points to DI's total markup. But the insurance that DI provides against other risks is significantly more valuable, contributing 48 percentage points—63% of DI's total markup. Even when we use additional health and disability measures to define finer health categories in equation (8),

[31] In addition to the marginal utility gap, the value of insurance against non-health risk is also increasing in the proximity of the recipiency rates in severe and non-severe states to 0.5, $p(\Omega_b|h) \approx 0.5$. The closer recipiency rates are to 0.5, the greater the total within-health transfers induced by the benefits. Intuitively, the closer the recipiency rate to 0 or 1, the smaller the effect of the benefits on the distribution of resources.

we still find that a majority (55%) of DI's insurance value comes from insuring risks *within* states with the same realizations of measured health. These results, which are consistent with the earlier result that roughly half of DI's ex ante surplus comes from its classification errors, suggest that a majority of DI's overall insurance value comes from insuring risks beyond that of having a severe health condition.

DI's classification errors drive this insurance of non-health risk, albeit at the expense of worsening insurance against the risk of having a severe health condition. Compared to a hypothetical, infeasible program that would perfectly target states with severe health conditions ("Health-DI"), DI's across-health component is less than one-half as large (0.28 for DI versus 0.66 for Health-DI). Yet DI more than makes up for its lesser insurance of the risk of having a severe health condition by providing highly valuable insurance against other risks. DI's insurance of other risks contributes 0.48 to its markup (whereas by definition Health-DI does not insure risks within health levels), vaulting its overall markup and surplus per recipient above those of Health-DI (0.75 versus 0.66 and $8,700 versus $5,000, respectively).

Although both types of DI's classification errors insure non-health risk, the vast majority of such insurance comes from DI's inclusion errors (NS-DI). The "NS-DI" row of the table decomposes the markup on transfers from all states to NS-DI states, using the same approach used to decompose the DI markup (substituting "NS-DI" for "DI" in equation (8)). NS-DI benefits worsen health risk, by transferring from all states to lower-marginal utility NS states, but provide highly valuable insurance against non-health risk, by transferring from NS-NDI states to much higher-marginal utility NS-DI states. The insurance against non-health risk provided by DI's inclusion errors accounts for over half of DI's overall markup and 85% of DI's "insurance of non-health risk" markup component.[32]

Selective applications as a driver: DI's insurance of non-health risk and overall value are driven primarily by strong selection into application rather than selective awards given applications. Selection into DI receipt reflects the combined effects of selective applications and selective awards given applications. We quantify the importance of selective applications by comparing DI to a hypothetical DI with no systematic selection into application ("Random-App-DI").[33] Whereas the distribution of actual DI benefits

[32]For S-DI ("correct" awards), the "health risk" component is large because transfers from all states to S states is in the direction of much greater marginal utility on average, and the "non-health risk" component is modestly positive because transfers from S-NDI to S-DI states are in the direction of modestly greater marginal utility on average. For S-NDI (hypothetical benefits to DI's exclusion errors), the "health risk" component is large because transfers from all states to S states is in the direction of much greater marginal utility on average, and the "non-health risk" component is modestly negative because transfers from S-DI states to S-NDI states is in the direction of modestly lower marginal utility on average.

[33]To simulate the distribution of benefits under Random-App-DI, we use Low and Pistaferri's (2015) estimates, from an indirect inference approach based on a structural model, of DI award success probabilities

depends on the combined effects of selection into application and selective awards among applicants, the distribution of Random-App-DI's benefits depends only on selective awards among applicants. Comparing the two is therefore informative about the role of selection into application in driving DI's value.

The results, shown in the last row of Table 2, suggest that selection into application plays a crucial role in driving DI's overall value. DI is much more valuable than Random-App-DI. Whereas DI generates $8,700 per recipient worth of annual surplus, Random-App-DI would generate −$6,100 per recipient worth of annual surplus, i.e., it would have *negative* value: Individuals would prefer cost-equivalent tax cuts to Random-App-DI. The much greater value of DI than Random-App-DI is driven almost entirely by DI's greater insurance of risks beyond that of having a severe health condition. DI's markup exceeds that on Random-App-DI by 53 percentage points, 47 percentage points (89%) of which are from its greater markup on within-health transfers. Selection into application is responsible for the vast majority of DI's insurance against non-health risk, which itself is highly important to DI's overall value.

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 not important. Rather, it means that the net effect of DI's award decisions and other features, including the many restrictions it imposes on applicants and recipients, is strongly favorable selection into application that greatly increases DI's value.

5.4 Policy implications

Together, the findings from the positive and normative analyses paint a picture of a DI program that targets well on marginal utility of income and on counterfactual earnings, including not only its awards to individuals with severe health conditions but also its classification errors. While random classification errors would be quite costly, our results suggest that DI's highly-selected classification errors *increase* its ex ante value, making DI more valuable than a hypothetical program that targeted perfectly on health would be.

These results come with important caveats. They depend on the usual factors underpinning most welfare analyses: They rely on existing estimates of behavioral responses, and they depend on hard-to-test assumptions about marginal utility, private optimization, and that

as a function of an applicant's work limitation status and age. We assume that each applicant applies for DI up to two times (if they are rejected the first time), to reflect that rejected applicants frequently re-apply. The results are similar if we assume that applicants apply just once. We also assume that each applicant accepts DI if awarded, even those who would have earned far more than the DI earnings limit had they not received DI. While in reality the earnings test likely deters many high-earning individuals from applying, these calculations aim to shed light on the value of DI were there not systematic selection into application.

DRAFT

there are no externalities associated with DI beyond the fiscal externalities associated with DI benefits and income taxes. Some of these assumptions seem especially important in the case of DI, including that we analyze a DI-like expansion of the budget constraint rather than DI itself (given the difficulty of quantifying stigma and hassle costs, for example) and that the assumption of private optimization requires, among other things, that individuals not overestimate their private return to being on DI, say by undervaluing work. Although our tests indicate that the key conclusions are robust to relatively large changes in these assumptions, they obviously depend on the assumptions we make. Acknowledging these important caveats, our results have implications for some of the biggest questions in the debate over DI:

1. **Is the provision of benefits to individuals without severe conditions diluting the value of the program?** We find that DI recipients without severe conditions are a highly selected group: They are much worse off than the general population of individuals without severe conditions, and by many non-health measures are even worse off than DI recipients with severe conditions. Figure 9 shows that non-severe recipients are similar to severe recipients in terms of both of the key determinants of the value of receiving DI: marginal utility and counterfactual earnings.[34] These results suggest that the benefits received by non-severe recipients are not diluting the value of DI, at least not in terms of insurance and welfare, but rather providing highly valuable insurance against a wide variety of major risks. To the extent that severe recipients are considered deserving of DI benefits on the basis of insurance and welfare gains, non-severe recipients appear nearly as deserving.

   To be sure, our results do not necessarily imply that DI should reduce its efforts to target individuals who have severe health conditions or increase efforts to target based on non-health risk. Evaluating changes to specific tags and screens requires evidence on the likely application responses to these changes, which is unfortunately sparse. It could be that DI's 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 is so crucial to its value. Similarly, it could be that the same features of DI that drive up its private cost and fiscal externality cost also drive up its insurance value, by screening out applications in lower-marginal utility states, such as those with substantial earnings capacity—an example of the targeting-distortion tradeoff emphasized by Nichols and

[34]Quasi-experimental evidence on the effects of DI awards on the earnings of individuals with different sets of health conditions suggests that not only severe recipients but also non-severe recipients would have earned very little had they not received DI (Maestas, Mullen and Strand, 2013; French and Song, 2014; Von Wachter, Song and Manchester, 2011). On average, DI recipients on the margin of allowance have counterfactual earnings on the order of $5,000 to $9,000 per year (Maestas, Mullen and Strand, 2013; French and Song, 2014).

DRAFT

Zeckhauser (1982). The strong insurance- and welfare-promoting selection into DI is consistent with DI's screens, especially its earnings test, providing a valuable screening role by leading individuals to choose not to apply in states with non-trivial earnings capacity. On the other hand, these or other features of DI might impose costs beyond any improvement, if any, in targeting they produce (as, for example, the SSA field office closings in Deshpande and Li, 2019). Quantifying the costs and benefits of specific DI tags and screens is a critical question for future research.

2. **Would expanding or contracting DI increase social welfare?** We quantify the insurance and welfare effects of key reforms to DI, with details in Appendix Section B and results in Appendix Table A4. The results suggest that expansions to DI would generate substantial surplus and that contractions—even a costless elimination of inclusion errors—would decrease it. As we discuss in that section, there is substantial uncertainty about key inputs into this analysis, including hassle and stigma costs of applying for or receiving DI and application responses to policy changes. As a result, these conclusions are tentative pending better evidence on these parameters.

   An additional consideration for the effects of DI on social welfare is possible redistribution value from DI. Throughout the paper, we have focused on DI's 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 DI would redistribute from higher- to lower-ability groups, since lower-ability groups both are more likely to receive DI and pay lower absolute amounts of the payroll taxes that fund DI. Such redistribution would increase the overall social welfare value of DI under social welfare functions that value redistribution more than the income tax and transfer implicitly does.[35] To get a rough sense of the likely magnitudes, we do back-of-the-envelope calculations of the value to someone behind the veil (or, equivalently, to a Utilitarian planner) of DI's redistribution effects—the differences between each type's expected benefits and expected DI payroll taxes—holding other policies fixed. We find that accounting for redistribution increases the value of DI to someone behind the veil from \$1,130 to \$1,670 per capita, a 32% increase.[36]

[35] The income tax and transfer system can be viewed as implicitly valuing redistribution from type θ_t to type θ_b at the welfare cost to θ_t of the tax on θ_t necessary to benefit θ_b by \$1. This cost is $g_{\theta_b}/g_{\theta_t}$, the ratio of the two types' inverse-optimum weights, since it costs the government g_{θ_b} to benefit θ_b by \$1 and raising \$1 from θ_t hurts θ_t by $1/g_{\theta_t}$. Redistribution from θ_t to θ_b through the income tax thus increases social welfare if and only if the ratio of marginal social welfare weights exceeds the ratio of inverse-optimum weights.

[36] The value of "pure," no-redistribution DI to someone behind the veil exceeds the social surplus from DI (\$1,130 versus the \$920 reported in Table 1, respectively) because someone behind the veil values surplus to different types differently from the implicit valuations implied by the inverse optimum weights. Assuming other policies would be the same were there no DI, these calculations likely understate DI's redistribution

DRAFT

Of course, a distinct question, and one that this paper is not equipped to answer, is whether we *should* in fact expand DI. Such a normative judgment depends on considerations beyond DI's effects on insurance and welfare, including non-welfarist considerations about DI's appropriate role or about the extent to which different individuals or states are deserving of benefits. For example, whereas the efficiency cost of DI depends on what recipients' earnings would have been had they not received DI, non-welfarist judgments might also depend on *why* recipients' earnings are low, say whether due to bad luck or bad decisions. Although we do not explicitly incorporate non-welfarist considerations in our analysis, our results suggest that such considerations would have to be fairly strong in order to overturn our main conclusions. To overturn the conclusion that DI benefits to non-severe recipients are valuable would require that each dollar of benefits to non-severe recipients be viewed as being worth less than 65 cents. To overturn the conclusion that DI as a whole is valuable would require that each dollar of benefits to non-severe recipients be viewed as being worth less than 25 cents.

3. **Is DI the *best* way to insure non-health risk?** We find that insurance against non-health risk accounts for about one-half of DI's high value. This is driven not only by strong selection into DI on the basis of non-health shocks, but also by the extent to which non-health shocks are otherwise uninsured. In other words, this finding in part reflects gaps in the existing social safety net, namely its protection against non-health risks. Would not non-health risks be better addressed more directly by filling these gaps with policies specifically designed to insure non-health risk rather than by DI, a program designed to provide assistance to individuals with severe health conditions? Quite possibly. For individuals who have not experienced a permanent drop in earnings capacity, meeting DI's statutory earnings limit is privately and socially costly. An alternative program without such strong work disincentives (perhaps expanded UI or wage insurance) would likely be a more efficient way to insure them. On the other hand, the best evidence suggests that the large majority of DI recipients would earn little even if not for DI,[37] and other, non-DI ways of supporting such individuals have efficiency costs as well. Whether alternative policies could improve upon DI's high

DRAFT

value to someone behind the veil by assuming away any insurance or redistribution value from the fact that DI payroll taxes are increasing in earnings within risk types, capturing only that redistribution across risk types.

[37]The evidence suggests that 20–30% of DI recipients on the margin of allowance would earn at least full-time, full-year minimum-wage amounts (about $15,000 annually) if they did not receive DI (Maestas, Mullen and Strand, 2013; French and Song, 2014). Although it would be useful to know how many would earn substantially more than that, we are not aware of any evidence on higher earnings levels. Assuming that inframarginal recipients have lower earnings capacity than marginal recipients, these estimates put a meaningful upper bound on the fraction of DI recipients with nontrivial earnings capacity. That DI benefits to individuals with low counterfactual earnings appear to be so valuable likely reflects a growing gap in the safety net, with the shift from traditional cash assistance to in-work benefits in recent decades (Hoynes and Schanzenbach, 2018), of support for individuals and households with low potential earnings.

value is an important question for future research.

6 Conclusion

In the public debate, DI is mainly evaluated relative to its statutory purpose of providing assistance to individuals with severe health conditions. Relative to this perspective, providing benefits to individuals without severe disabilities is a major concern, one that has dominated the public debate over the program. In this paper, we go beyond this health-based perspective to evaluate DI from a broader perspective that considers insurance against all risks, non-health as well as health. From this perspective, we find that not only is DI highly valuable, but its classification errors, including awards to individuals without severe disabilities, actually increase its value. This is because DI recipients without severe conditions are a highly selected group: Compared to non-recipients, non-severe recipients are more likely to have experienced adverse non-health events, are less likely to have the resources to insure those events, and are much more likely to have very low earnings even in the absence of DI. In fact, recipients without severe conditions are so highly selected that they look similar to recipients with severe conditions on these dimensions.

As a result of this strong selection, DI benefits to both severe and non-severe recipients produce insurance value substantially greater than distortion cost. Our results suggest that about one-half of DI's value comes from insuring non-health risk. For this reason, reforms to increase the emphasis on DI's statutory purpose could significantly reduce DI's ex ante value.

The importance of non-health risk in driving the value of DI is likely just one example of a broader phenomenon of mismatch between the statutory intent and actual effects of social programs. No program exists in a vacuum: Its effects reflect the diversity and severity 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 uninsured risks. We find that DI insures risks well outside of its statutory intent, and that its "accidental" role actually makes it more valuable than if it fulfilled only its intended role—even if it somehow did so with no additional administrative costs. Other programs might be like DI in their costs and benefits being driven by factors outside of their statutory aims. The importance of effects outside a program's statutory intent is an empirical question, one that future research could investigate using similar methods.

DRAFT

References

Autor, David, and Mark Duggan. 2006. "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding." National Bureau of Economic Research Working Paper 12436.

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.

Benıtez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, John Rust, and Sofia Sheidvasser. 1999. "An empirical analysis of the social security disability application, appeal, and award process." *Labour Economics*, 6(2): 147–178.

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 Jr Stephens. 2018. "Disability benefit take-up and local labor market conditions." *Review of Economics and Statistics*, 100(3): 416–423.

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.

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. 2019. "Disability and Distress: The Effect of Disability Programs on Financial Outcomes." National Bureau of Economic Research.

Diamond, Peter, and Eytan Sheshinski. 1995. "Economic aspects of optimal disability benefits." *Journal of Public Economics*, 57(1): 1–23.

DRAFT

Finkelstein, Amy, Erzo FP 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(suppl_1): 221–258.

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

Hendren, Nathaniel. 2020. "Measuring economic efficiency using inverse-optimum weights." *Journal of Public Economics*, 187: 104198.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2019. "A Unified Welfare Analysis of Government Policies." National Bureau of Economic Research, Inc NBER Working Papers 26144.

Hoynes, Hilary W., and Diane Whitmore Schanzenbach. 2018. "Safety net investments in children." *BPEA Conference Draft*.

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.

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.

Morton, William R. 2015. "Concurrent Receipt of Social Security Disability Insurance (SSDI) and Unemployment Insurance (UI): Background and Legislative Proposals."

Nichols, Albert L, and Richard J Zeckhauser. 1982. "Targeting transfers through restrictions on recipients." *The American Economic Review*, 72(2): 372–377.

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.

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

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.

DRAFT

Appendix (for online publication)

A Decomposing the DI markup

Section 5.3 decomposes the markup associated with increasing different groups of actual and hypothetical DI benefits into components that reflect the underlying across- and within-health transfers. Here, we derive the key equation in Section 5.3, equation (8), 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 DI 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, $h_\omega \in \{\text{NS, S}\}$, 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 here 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, $h_\omega \in \{\text{NS, S}\}$, $E(b|h = h_\omega) = p(\Omega_b|h)b$. The mechanical effect of—the ex ante expected marginal loosening of the budget constraint from—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-

DRAFT

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, $p(\Omega_b)$, 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 DI recipient and $-p(\Omega_b|h)$ if not.[38]

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

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

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.

[38]The sum of the mechanical effects of an increase in the overall benefit level b operating through the health-contingent benefit and within-group 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.

DRAFT

Combined value of health-contingent benefits and within-health transfers. Summing the marginal values of the health-contingent benefits and the within-group transfers, the full ex-ante marginal value of increasing the overall benefit level b of benefits in states Ω_b is

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

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

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 (8) in the main text.

B Policy analysis

The normative analysis in Section 5 estimates the ex ante value and cost of receiving DI benefits in different states of the world. This analysis does not speak directly to the welfare implications of reforms to DI. In this appendix section, we consider three commonly proposed policy reforms: reducing inclusion errors, 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 DI to the individual in a particular state comprises the costs of the DI benefits and of 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 DI. Application responses tend to increase the costs of expansions of DI, 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 DI applications to changes in DI policy and in economic conditions. Section B.4 provides details about these cost calculations.

Appendix Table A4 reports the results. The following sections provide details of the analysis for each of the three policies in turn.

B.1 Reducing inclusion errors

This hypothetical policy reform reduces the number of inclusion errors—benefits to non-severe recipients. This could potentially be achieved by increasing investments in health assessments during the adjudication process or in reassessments of existing recipients. Many

DRAFT

of the reforms that have been proposed to reduce SSDI spending (see, e.g., McCrery and Pomeroy, 2016) aim to reduce inclusion errors in different ways, such as:

- Stricter medical eligibility criteria: eliminate mental and back pain allowances
- 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 inclusion errors 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 inclusion errors being valuable, we ignore any costs associated with the investments necessary to reduce such errors, including any unintended effects on individuals with 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 non-severe recipient states as a whole in terms of marginal utility and counterfactual earnings. We estimate the foregone ex ante value of receiving DI in the non-severe states that no longer receive DI as a result of the reform exactly as in Section 5: See equation (2), plugging in $\Omega_b = \Omega_{NS\text{-}DI}$. The foregone ex ante cost to the government associated with providing DI in the non-severe states that no longer receive DI as a result of the reform comprises not only the foregone costs of those benefits and the induced behavioral responses in those states as in Section 5 but also the foregone costs of applications and awards in other states that no longer apply and receive DI as a result of the reform. See Section B.4 for details.

DRAFT

B.2 Decreasing the allowance rate

This reform decreases 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 DI, while the award decisions on the 77 percent of inframarginal applicants are unchanged. We consider a decrease in the allowance rate of 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 non-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 DI in the states in which the individual is an applicant on the margin of program entry exactly as in Section 5: See equation (2), 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 DI recipiency rate $p(\Omega_{ampe}) = 0.1 \times p(\Omega_{DI})$. The ex ante cost to the government associated with providing DI in the affected states comprises not only the foregone costs of those benefits and the induced behavioral responses in those states as in Section 5, but also the cost savings of applications and awards in other states that are discouraged from applying for DI as a result of benefits in the affected states. See Section B.4 for details.

B.3 Decreasing benefit levels

This reform marginally decreases the DI benefit level. 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 DI benefits in different states described in Section 5, here focused on all states in which the individual receives DI benefits. The key differences are that here we consider application responses and that the ex post value of a change in the benefit level differs from that of introducing DI.

We estimate the ex ante value of receiving DI in the states in which the individual is an "inframarginal recipient"—i.e., receiving DI 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-DI earnings were below the DI limit. Formally, the ex post value in state of the world ω of a Δb decrease in the DI benefit is

$$WTP_\omega = -DI_\omega \times \Delta b, \tag{9}$$

where DI_ω equals one if the individual receives DI benefits in state ω under status quo policies (i.e., before the benefit decrease). Recall from equation (1) that the ex post marginal value in state ω of a DI-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 DI in state ω, $\bar{z}$ is the earnings limit for DI recipients, and b is the DI benefit. The difference arises because any costs to inframarginal recipients of DI's earnings test or other restrictions, though diminishing the full value to them of receiving DI, do not change the marginal value of receiving lower benefits. We plug this ex post value into equation (2), plugging in $\Omega_b = \Omega_{DI}$, the set of states in which the individual is an inframarginal recipient.

The ex ante cost to the government associated with changing benefit levels comprises not only the cost savings from the benefit decrease and induced behavioral responses in the states in which the individual is an "inframarginal recipient," but also the cost savings from fewer applications and awards in states in which the individual's decision of whether to apply for DI is affected by the change in benefits. See Section B.4 for details.

DRAFT

B.4 The fiscal externality of changes in DI

For each policy counterfactual, we calculate the associated costs to the government based on the available evidence from the literature on behavioral responses to DI 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 DI. 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 DI 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 DI 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) DI 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 DI (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 DI are driven by younger people, our focus on changes in DI awards and benefits to non-elderly beneficiaries tends to increase the fiscal externality as a fraction of the mechanical 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 (2018) have used local economic booms and busts to estimate an elasticity of DI 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 DI 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 DI benefits less decreased means-tested transfers received and increased tax payments made—of \$0.34. Increasing net transfers to inframarginal DI recipients by \$1 costs the government \$1.34. This is somewhat smaller than Bound et al.'s

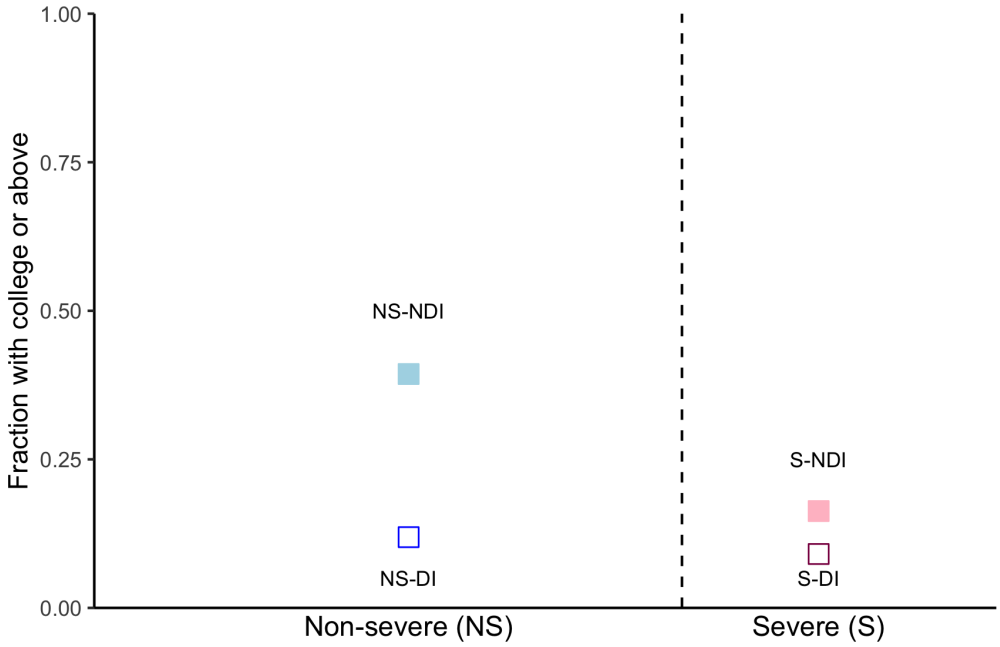
DRAFT

(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 DI reform counterfactuals and each of our ex ante risk types, we scale the mechanical effect of the reform on that risk type—the change in expected net transfers to inframarginal recipients—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.

DRAFT

Appendix Figures and Tables (for online publication)

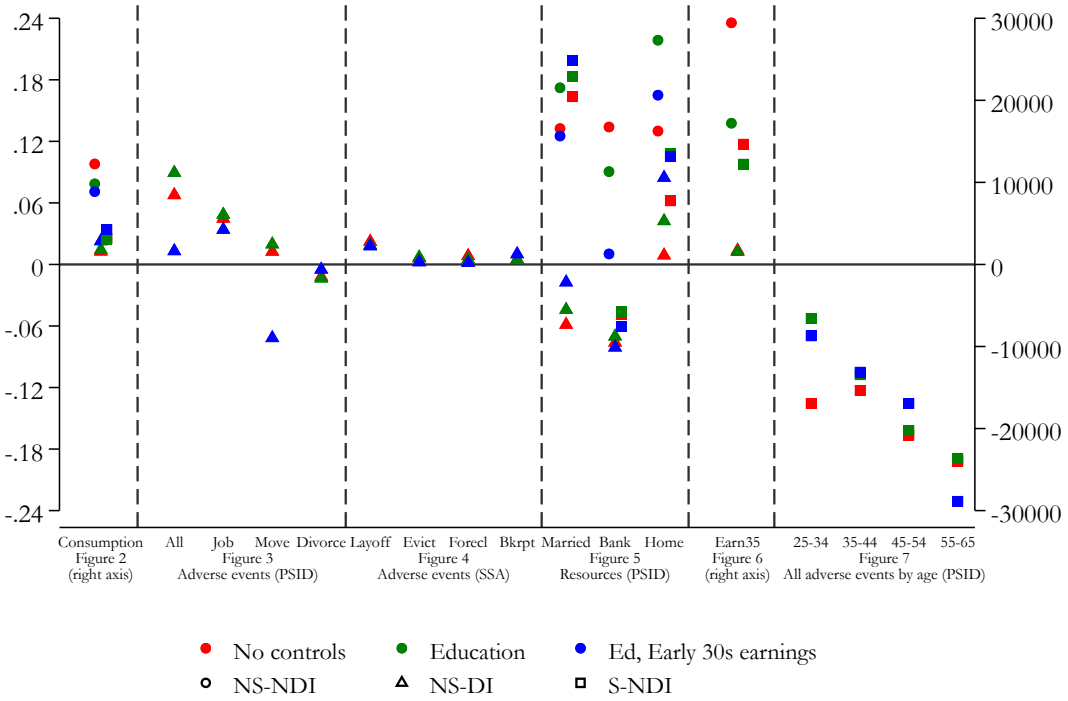
Figure A1: PSID: College or above



Notes: Figure presents rates of college degree or above obtained for four types of households in the 2017 PSID: households not currently receiving DI (SSDI or SSI) without a severe health condition (6,312 NS-NDI), households not currently receiving DI with a severe health condition (250 S-NDI), households currently receiving DI with a severe health condition (316 S-DI), and households currently receiving DI without a severe health condition (443 NS-DI). "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 "non-severe."

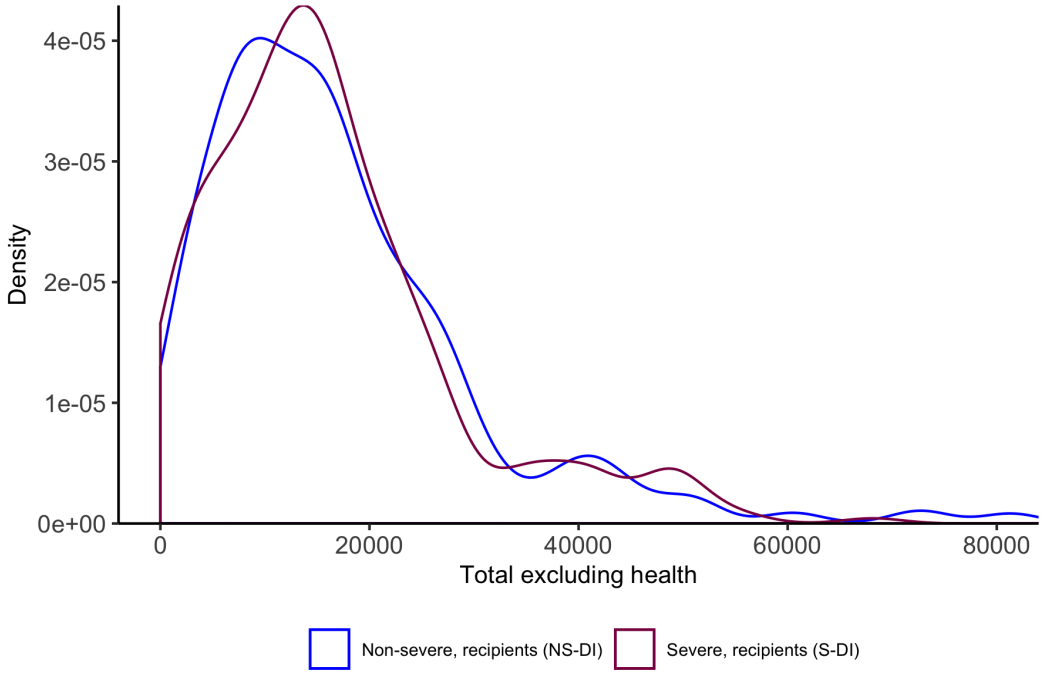
DRAFT

Figure A2: Robustness to Ability-Group Controls



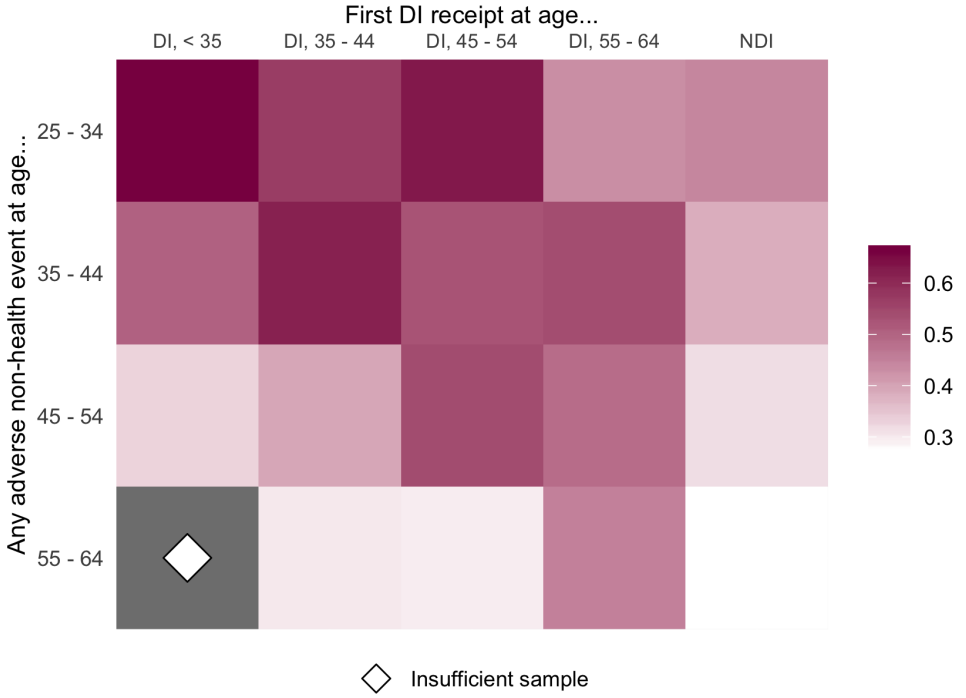
Notes: Figure presents all outcomes from Figures 2 to 7 with controls for ability group. The figure plots the coefficients on indicators for each DI-by-severity group (NS-NDI, NS-DI, and S-NDI; S-DI is the excluded group) from a regression of the outcome on those indicators and ability controls. The three groups are shown using different shapes. We use three specifications for ability controls: no controls (which reproduces the coefficients from Figures 2 to 7), 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 3 indicates experiencing at least one of the adverse events (job loss, involuntary move, or divorce).

Figure A3: PSID: Consumption distribution at application



Notes: Figure presents the distribution of consumption (excluding health care expenditures) one year prior to the year of initial DI (SSDI or SSI) receipt for two types of households in the 2017 PSID: households currently receiving DI without a severe health condition (213 S-DI), and households currently receiving DI without a severe health condition (286 NS-DI). Household consumption excludes health care and is divided by the square root of the household size. "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 "non-severe."

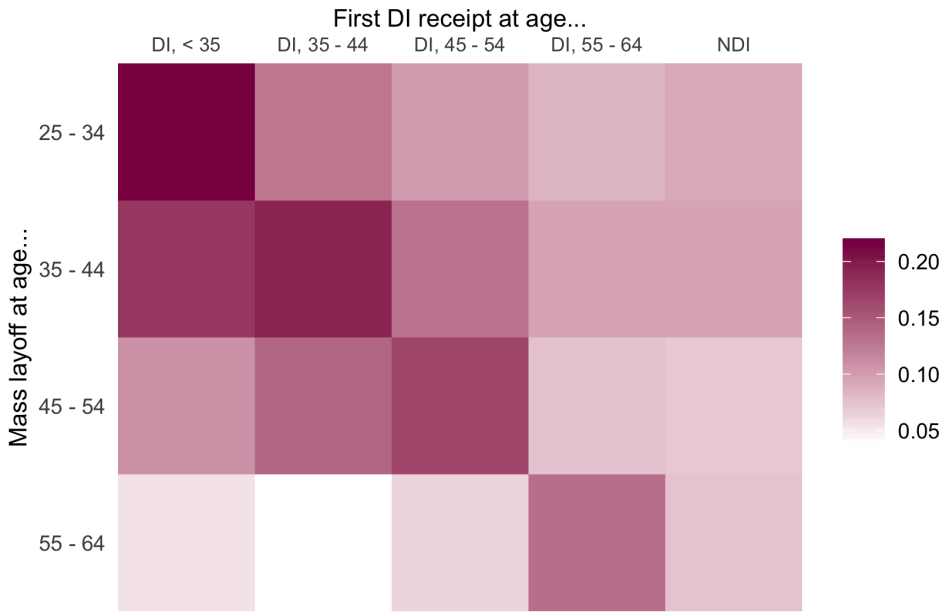
Figure A4: PSID: Any adverse life event by DI status



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 DI (SSDI or SSI) status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). The DI group is further categorized by their age at DI 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.

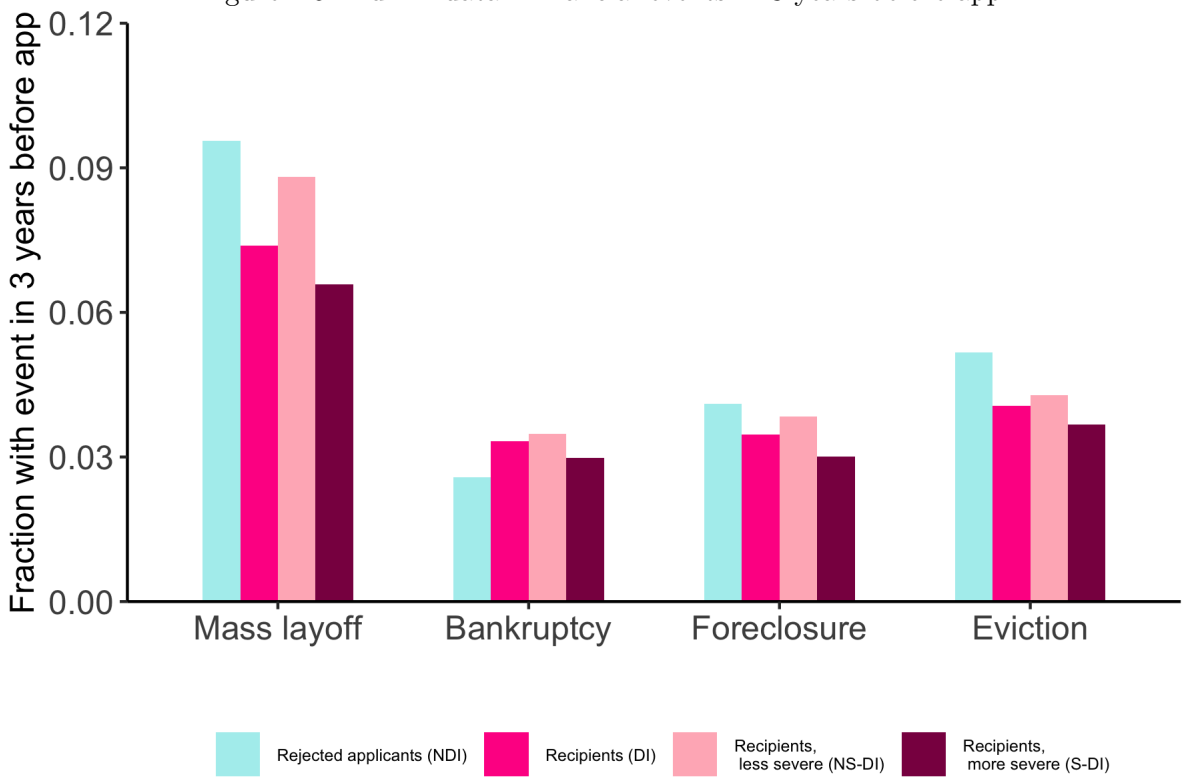
DRAFT

Figure A5: CWHS: Mass layoffs by DI status



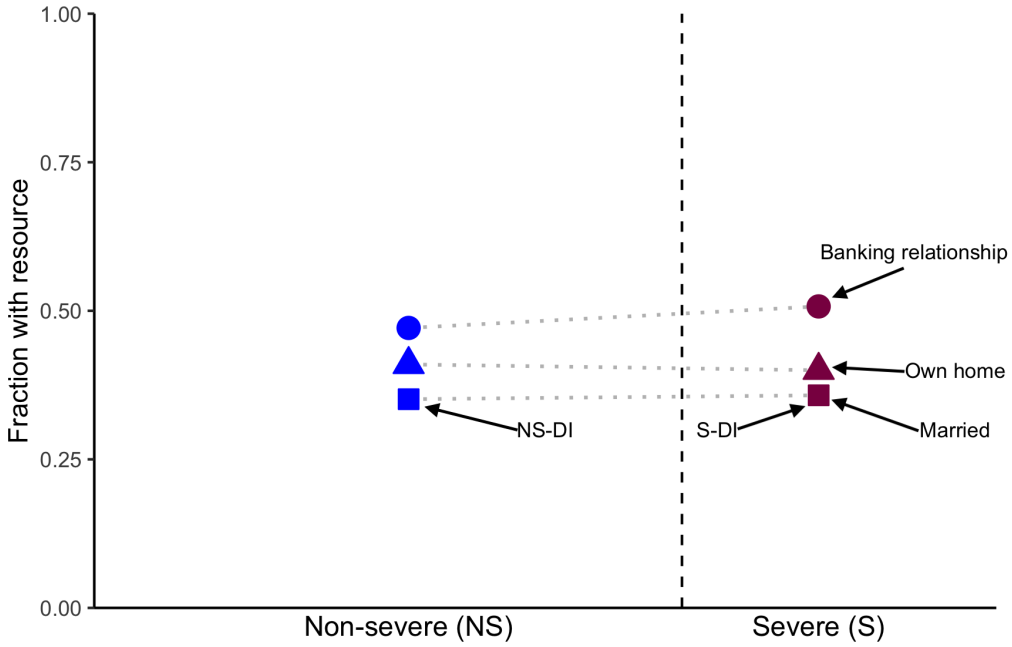
Notes: Figure represents rates of mass layoff at various age for individuals in the CWHS, by DI (SSDI or SSI) status, and age at DI entry. Individuals are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). The DI group is further categorized by their age at DI 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 A6: Admin data: Financial events in 3 years before app



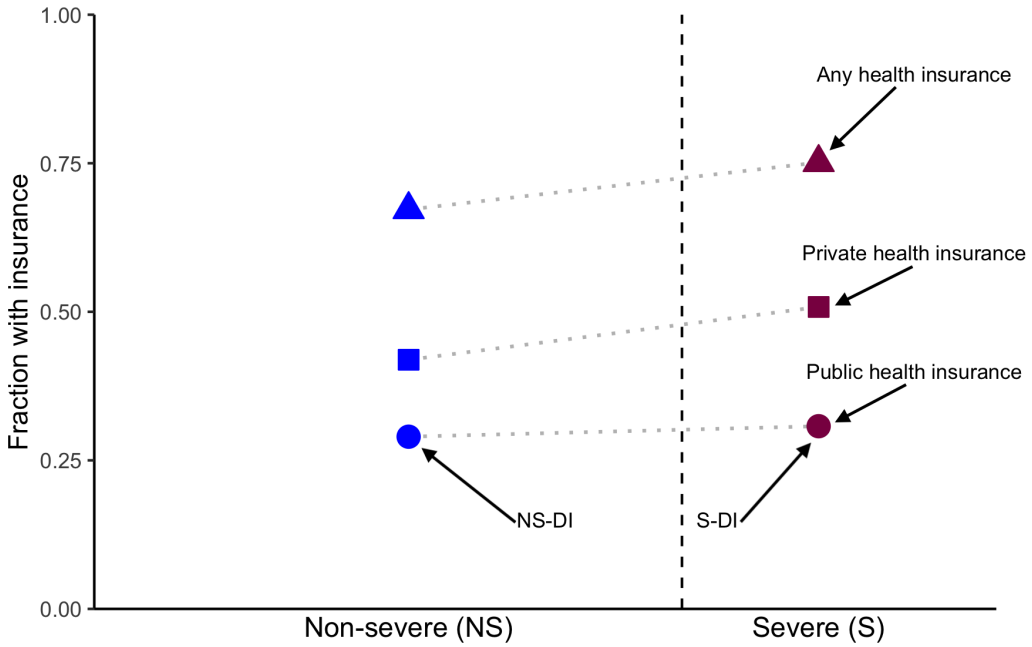
Notes: Figure presents rates of adverse financial events prior to DI (SSDI or SSI) receipt for rejected applicants and approved applicants (by more-severe and less-severe) in SSA administrative data. "Mass layoff" indicates being part of a mass layoff in the five years prior to DI application. The mass layoff sample is DI recipients in the CWHS-linked Master Beneficiary Record (MBR) who were allowed onto DI between 1985 and 2016 (520,000 less-severe, 940,000 more-severe), and who applied but rejected (980,000 rejected applicants). We determine rates of mass layoff using the Continuous Work History Sample (CWHS), which is a 10% sample of individuals in the US. In the CWHS, we identify EINs that experience a drop in employees of at least 30 percent from a base of at least 150 employees. From the MBR, we label an DI recipient as experiencing 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 DI application. We link DI applicants in the 831 files to nationwide financial records using the methods described in Deshpande, Gross and Su (2019). The bankruptcy sample is DI recipients in the 831 files between 1995 and 2009 (6.5M less-severe, 3.2M more-severe), and those applicants who were rejected (8.2M rejected applicants). The foreclosure sample is approved DI applicants in the 831 files between 2005 and 2014 who are homeowners (1.2M less-severe, 1.0M more-severe), and those applicants who were rejected (1.3M rejected applicants). The eviction sample is approved DI applicants in the 831 files between 2005 and 2014 who are not homeowners (1.4M less-severe, 770,000 more-severe), and those applicants who were rejected (2.4M rejected applicants). "More severe" indicates a medical diary of "medical improvement not expected," and "less severe" indicates a medical diary of "medical improvement possible" or "medical improvement expected."

Figure A7: PSID: Outside Resources at application



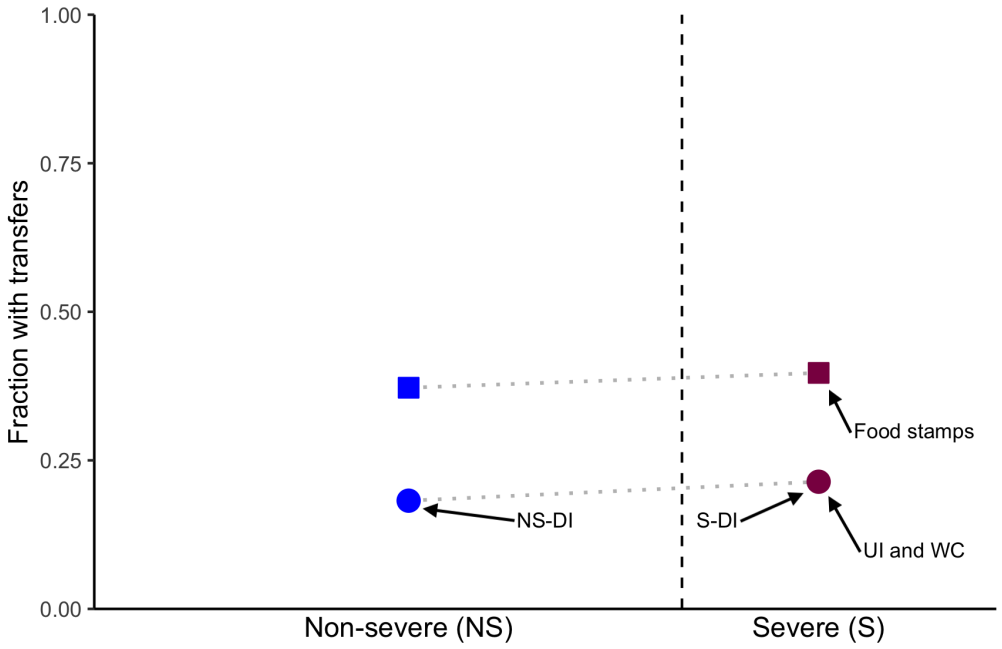
Notes: Figure presents rates of marriage, banking relationship, and homeownership one year prior to the year of initial DI (SSDI or SSI) receipt for two types of households in the 2017 PSID: households currently receiving DI with a severe health condition (243 S-DI), and households currently receiving DI without a severe health condition (320 NS-DI). "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 "non-severe." Appendix Table A8 presents the PSID questions for banking relationship and homeownership.

Figure A8: PSID: Health Insurance Coverage (Head) at application



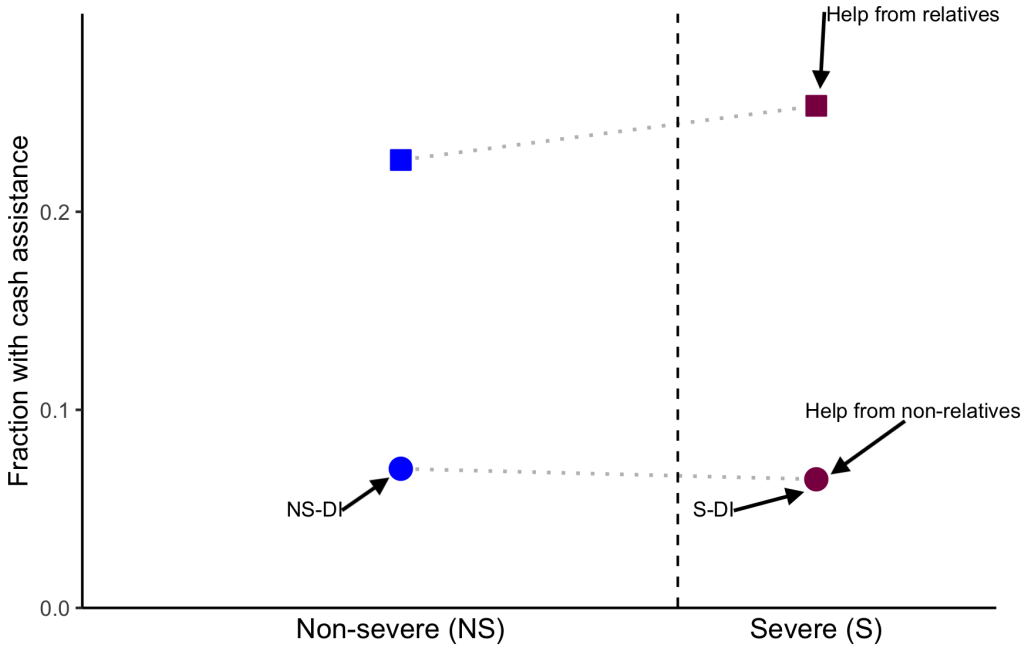
Notes: Figure presents rates of any health insurance, private health insurance, and public health insurance one year prior to the year of initial DI receipt for two types of households in the 2017 PSID: households currently receiving DI (SSDI or SSI) with a severe health condition (243 S-DI), and households currently receiving DI without a severe health condition (320 NS-DI). "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 "non-severe." Appendix Table A9 presents the PSID questions for private and public insurance.

Figure A9: PSID: Transfers at application



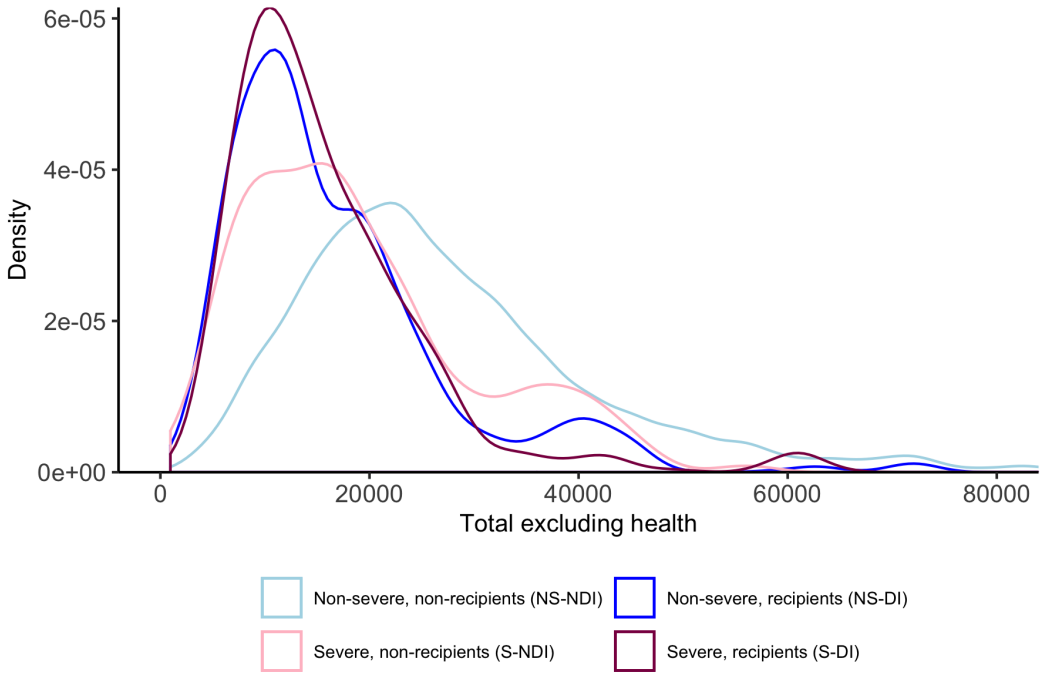
Notes: Figure presents rates of SNAP (food stamps) and unemployment insurance (UI) and workers' compensation (WC) one year prior to the year of initial DI (SSDI or SSI) receipt for two types of households in the 2017 PSID: households currently receiving DI with a severe health condition (243 S-DI), and households currently receiving DI without a severe health condition (320 NS-DI). "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 "non-severe." Appendix Table A8 presents the PSID questions for SNAP, UI, and WC.

Figure A10: PSID: Cash assistance at application



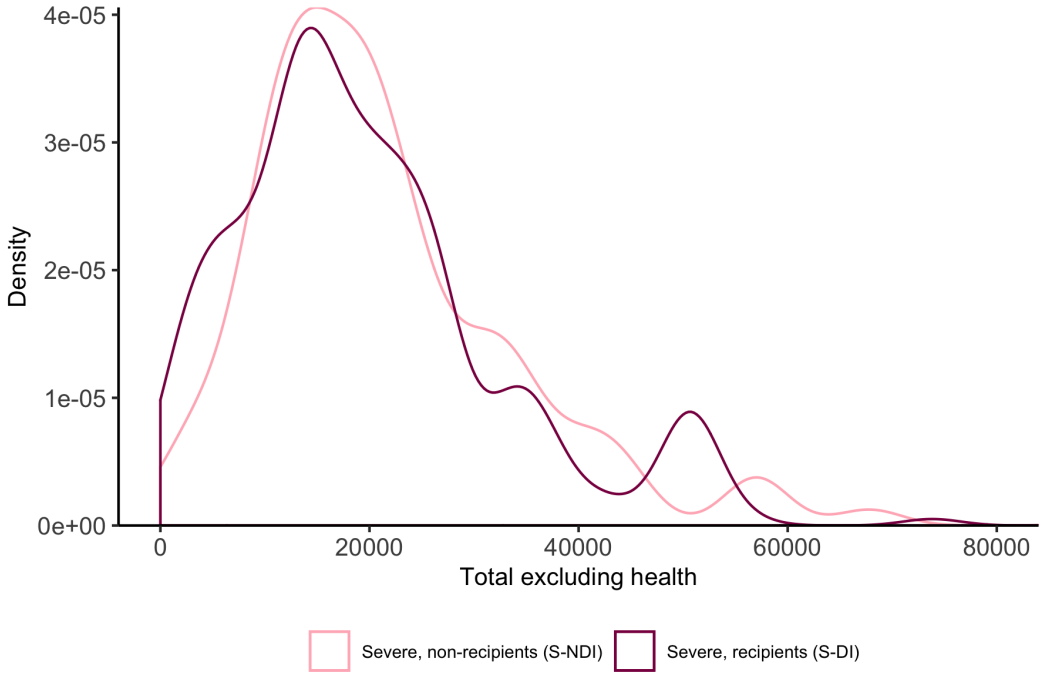
Notes: Figure presents rates of cash assistance from relatives and non-relatives one year prior to the year of initial DI receipt for two types of households in the 2017 PSID: households currently receiving DI (SSDI or SSI) with a severe health condition (243 S-DI), and households currently receiving DI without a severe health condition (320 NS-DI). "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 "non-severe." Appendix Table A8 presents the PSID questions for cash assistance from relatives and non-relatives.

Figure A11: PSID: Consumption distribution



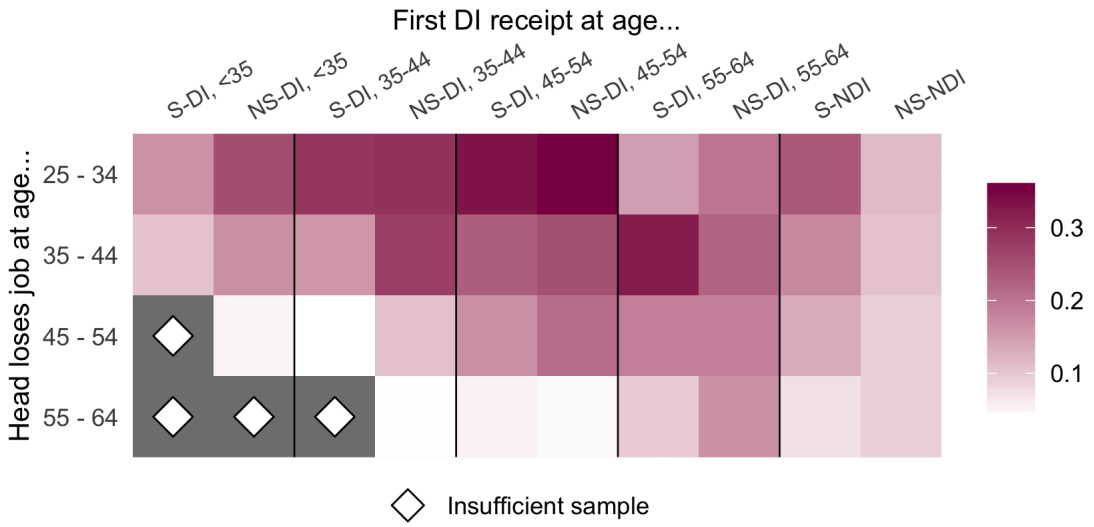
Notes: Figure presents the distribution of consumption (excluding health care expenditures) for four types of households in the 2017 PSID: households not currently receiving DI (SSDI or SSI) without a severe health condition (6,312 NS-NDI), households not currently receiving DI with a severe health condition (250 S-NDI), households currently receiving DI with a severe health condition (316 S-DI), and households currently receiving DI without a severe health condition (443 NS-DI). Household consumption excludes health care and is divided by the square root of the household size. "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 "non-severe."

Figure A12: PSID: Consumption distribution at severe onset



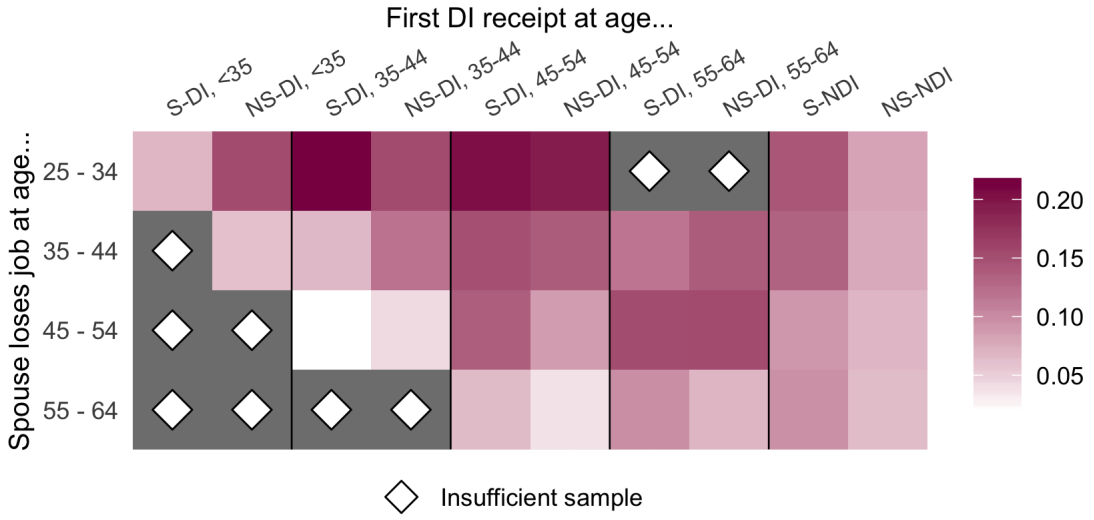
Notes: Figure presents the distribution of consumption (excluding health care expenditures) one year prior to the year of severe onset for two types of households in the 2017 PSID: households not currently receiving DI (SSDI or SSI) with a severe health condition (186 S-NDI), and households currently receiving DI with a severe health condition (197 S-DI). Household consumption excludes health care and is divided by the square root of the household size. "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 "non-severe."

Figure A13: PSID: Head job loss by DI and health status



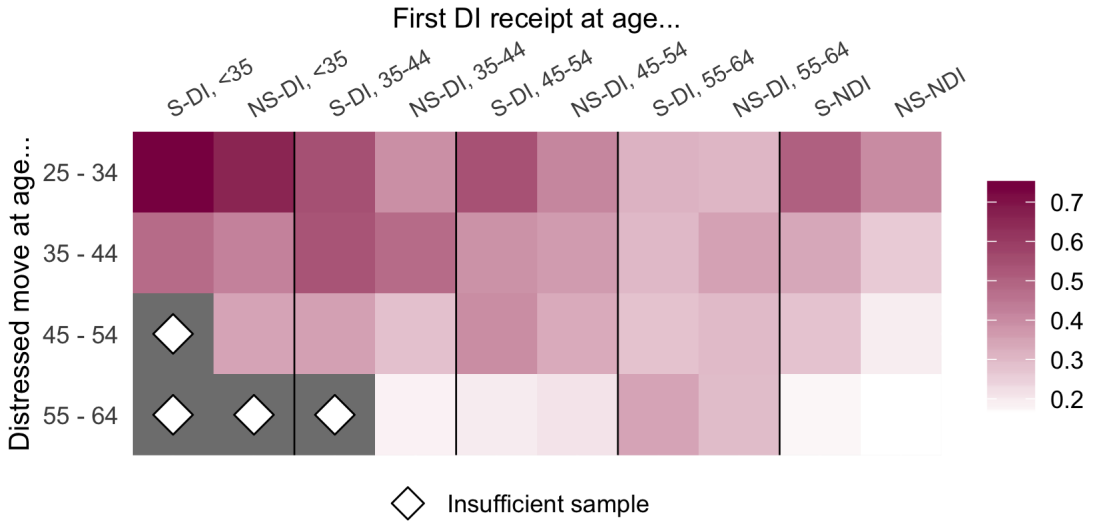
Notes: Figure presents rates of head job loss at various ages for households in the PSID, by DI (SSDI or SSI) status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI), and NDI by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The S-DI and NS-DI groups are further categorized by their age at DI 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 (see Appendix Table A6). "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 "non-severe."

Figure A14: PSID: Spouse job loss by DI and health status



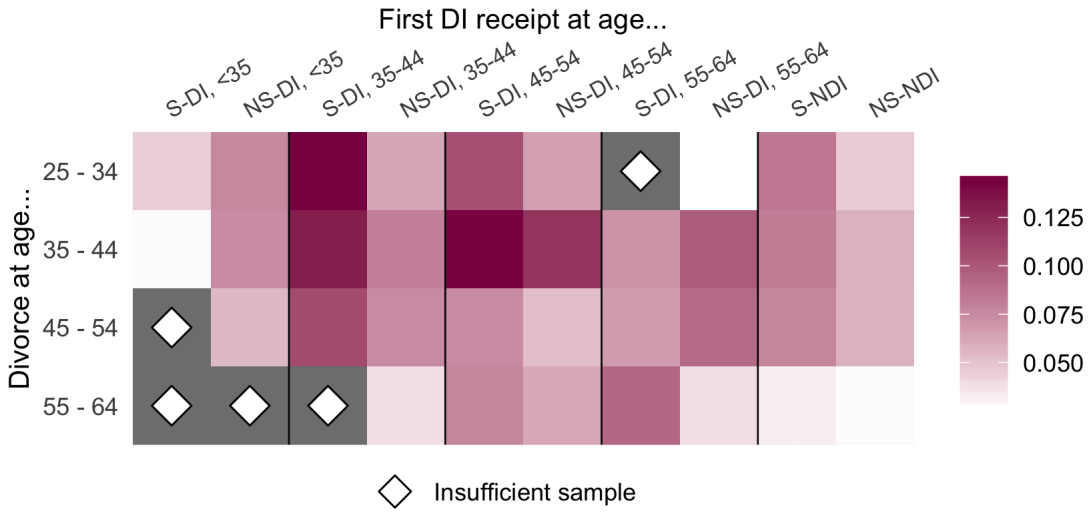
Notes: Figure presents rates of experiencing spouse job loss at various ages for households in the PSID, by DI (SSDI or SSI) status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI), and NDI by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The S-DI and NS-DI groups are further categorized by their age at DI 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 (see Appendix Table A6). "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 "non-severe."

Figure A15: PSID: Distressed move by DI and health status



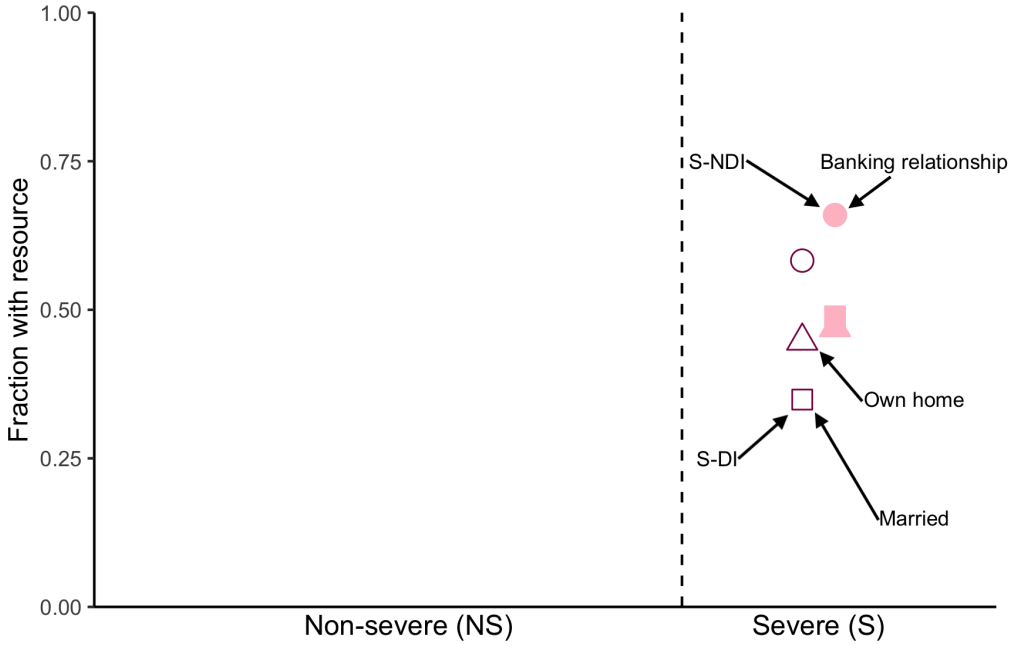
Notes: Figure presents rates of experiencing a distressed move at various ages for households in the PSID, by DI (SSDI or SSI) status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI), and NDI by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The S-DI and NS-DI groups are further categorized by their age at DI 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 (see Appendix Table A7). "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 "non-severe."

Figure A16: PSID: Divorce by DI and health status



Notes: Figure presents rates of experiencing divorce at various ages for households in the PSID, by DI status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI), and NDI by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The S-DI and NS-DI groups are further categorized by their age at DI 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. Divorce is defined as being married in a previous survey year but not in this survey year (see Appendix Table A6). "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 "non-severe."

Figure A17: PSID: Outside Resources at severe onset



Notes: Figure presents rates of marriage, banking relationship, and homeownership one year prior to the year of severe onset for two types of households in the 2017 PSID: households not currently receiving DI (SSDI or SSI) with a severe health condition (207 S-NDI), and households currently receiving DI with a severe health condition (240 S-DI). "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 "non-severe." Appendix Table A8 presents the PSID questions for banking relationship and homeownership. Markers for the two groups are offset to facilitate easier reading of the graph, not as a representation of relative severity.

DRAFT

Table A1: Value of DI benefits in different states: Robustness to key assumptions

	$E(M_\omega)$	Counterfactual earnings ($)	Surplus per capita ($) $(p(\Omega_\tau)s)$	Surplus per recipient ($) $(p(\Omega_\tau)s/p(\Omega_b))$	$MVPF$
Baseline					
DI	0.91	5,191	924	8,718	1.76
S-DI	0.95	3,557	488	9,858	1.85
NS-DI	0.88	6,623	436	7,719	1.68
S-NDI	0.64	22,591	-74	-2,203	0.93
Lower risk aversion ($\gamma = 1$, i.e., log utility)					
DI	0.41	5,191	310	2,927	1.30
S-DI	0.44	3,557	186	3,763	1.36
NS-DI	0.39	6,623	124	2,195	1.24
S-NDI	0.28	22,591	-208	-6,180	0.66
Higher risk aversion ($\gamma = 3$)					
DI	1.41	5,191	1,530	14,425	2.22
S-DI	1.43	3,557	772	15,587	2.32
NS-DI	1.39	6,623	758	13,407	2.15
S-NDI	1.02	22,591	76	2,253	1.24
Higher earnings responses to DI					
DI	0.91	8,137	821	7,740	1.66
S-DI	0.95	6,380	451	9,101	1.76
NS-DI	0.88	9,676	370	6,548	1.57
S-NDI	0.64	22,591	-76	-2,270	0.93
Earnings responses based on severity					
DI	0.91	6,366	877	8,273	1.72
S-DI	0.95	4,434	473	9,547	1.82
NS-DI	0.88	8,060	404	7,157	1.63
S-NDI	0.64	22,591	-74	-2,215	0.93
State-dependent utility: Lower when severe					
DI	0.72	5,191	676	6,372	1.58
S-DI	0.49	3,557	215	4,349	1.42
NS-DI	0.92	6,623	460	8,144	1.72
S-NDI	0.25	22,591	-184	-5,471	0.71
State-dependent utility: Higher when severe					
DI	1.10	5,191	1,164	10,979	1.94
S-DI	1.39	3,557	751	15,162	2.27
NS-DI	0.85	6,623	413	7,314	1.65
S-NDI	1.01	22,591	32	940	1.14

DRAFT

Notes: Table presents mean markup, counterfactual earnings, social surplus per capita, social surplus per recipient, and MVPF associated with receiving DI benefits in each of several sets of states of the world. The MVPF, the marginal value of public funds associated with DI receipt, is the ex ante value to the individual per dollar of net cost to the government, $MVPF(\Omega_b) = \frac{EAWTP(\Omega_b)}{EAG(\Omega_b)}$. "Baseline" is the baseline specification. "Lower risk aversion" uses $\gamma = 1$ (i.e., log utility). "Higher risk aversion" uses $\gamma = 3$. "Higher earnings responses to DI" assumes that earnings responses to DI are twice French and Song's (2014) estimates. "Earnings responses based on severity" uses French and Song's (2014) estimates of earnings responses to DI based on severity rather than diagnosis (as in the baseline specification). "State-dependent utility: Lower when severe" assumes that marginal utility is 25% lower in states with severe health conditions at a given level of consumption. This is roughly based on Finkelstein, Luttmer and Notowidigdo (2013), who estimate that a one-standard deviation increase in the number of chronic diseases is associated with a 10%-25% decline in the marginal utility of consumption. This affects the markup and surplus from transfers not only to severe states but to non-severe states as well, by changing marginal utility in some of the states in which the individual pays for benefits (severe non-recipient states). "State-dependent utility: Higher when severe" assumes that marginal utility is 25% higher in states with severe health conditions at a given level of consumption. Lillard and Weiss (1997) estimate that marginal utility is higher when health is worse. 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 capita per year and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient per year.

Table A2: Value of DI benefits in different states: Heterogeneity by education level

	$E(M_\omega)$	Counterfactual earnings ($)	Surplus per capita ($) $(p(\Omega_\tau)s)$	Surplus per recipient ($) $(p(\Omega_\tau)s/p(\Omega_b))$	$MVPF$
Full pop					
DI	0.91	5,191	924	8,718	1.76
S-DI	0.95	3,557	488	9,858	1.85
NS-DI	0.88	6,623	436	7,719	1.68
S-NDI	0.64	22,591	-74	-2,203	0.93
HS dropouts					
DI	0.46	3,775	1,255	5,567	1.39
S-DI	0.50	3,751	647	6,067	1.43
NS-DI	0.43	3,796	608	5,118	1.36
S-NDI	0.44	11,011	108	1,541	1.10
HS graduates					
DI	0.66	5,135	931	5,977	1.52
S-DI	0.62	3,416	432	6,187	1.53
NS-DI	0.70	6,530	499	5,806	1.50
S-NDI	0.44	23,088	-174	-3,855	0.80
Some college					
DI	1.56	5,094	1,663	15,787	2.36
S-DI	1.60	3,525	942	17,289	2.48
NS-DI	1.53	6,776	721	14,176	2.22
S-NDI	0.74	25,059	-90	-2,906	0.90
College or above					
DI	1.13	8,415	271	7,764	1.95
S-DI	1.27	3,682	142	10,153	2.17
NS-DI	1.03	11,596	129	6,158	1.80
S-NDI	1.15	33,704	-49	-3,025	1.03

Notes: Mean markup, counterfactual earnings, social surplus per capita, social surplus per recipient, and MVPF associated with receiving DI benefits in each of several sets of states of the world. 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 capita per year and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient per year. The MVPF, the marginal value of public funds associated with DI receipt, is the ex ante value to the individual per dollar of net cost to the government, $MVPF(\Omega_b) = \frac{EAWTP(\Omega_b)}{EAG(\Omega_b)}$

DRAFT

Table A3: Value of DI benefits in different states: Robustness to different strategies for isolating risk from heterogeneity

	$E(M_\omega)$	Counterfactual earnings ($)	Surplus per capita ($) $(p(\Omega_\tau)s)$	Surplus per recipient ($) $(p(\Omega_\tau)s/p(\Omega_b))$	$MVPF$
Risk types: education (baseline)					
DI	0.91	5,191	924	8,718	1.76
S-DI	0.95	3,557	488	9,858	1.85
NS-DI	0.88	6,623	436	7,719	1.68
S-NDI	0.64	22,591	-74	-2,203	0.93
Risk type: earnings					
DI	0.65	5,587	684	6,091	1.51
S-DI	0.85	3,047	440	9,258	1.77
NS-DI	0.51	7,445	245	3,774	1.33
S-NDI	0.58	27,812	-173	-4,873	0.75
Estimate markups with controls					
DI	0.33	5,178	117	1,106	1.20
S-DI	0.34	3,555	91	1,846	1.26
NS-DI	0.32	6,593	26	460	1.14
S-NDI	0.33	22,551	-210	-6,279	0.68

Notes: Table presents mean markup, counterfactual earnings, social surplus per capita, social surplus per recipient, and MVPF associated with receiving DI benefits in each of several sets of states of the world. 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 capita per year and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient per year. The MVPF, the marginal value of public funds associated with DI receipt, is the ex ante value to the individual per dollar of net cost to the government, $MVPF(\Omega_b) = \frac{EAWTP(\Omega_b)}{EAG(\Omega_b)}$. 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 DI 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 DI-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.

DRAFT

Table A4: Policy analysis

	$E(M_\omega)$	Surplus per capita ($) $(p(\Omega_\tau)s)$	Surplus per recipient ($) $(p(\Omega_\tau)s/p(\Omega_b))$	$MVPF$
Eliminate inclusion errors	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 capita, social surplus per recipient, and MVPF associated with different policies. Each of the policies contracts DI, so a negative surplus means that contracting DI in that way reduces social surplus. The MVPF, the marginal value of public funds associated with each policy, is the ex ante value to the individual per dollar of net cost to the government of *expanding* DI along that dimension, $MVPF(\Omega_b) = \frac{EAWTP(\Omega_b)}{EAG(\Omega_b)}$. "Eliminate inclusion errors" is a hypothetical, infeasible policy that eliminates inclusion errors (benefits to non-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 DI recipiency decreases by 10%, from about 11% to about 10%. We assume that this decrease in allowances affects "DI 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 non-severe recipients. We assume that each of these contractions of DI produce cost savings to the government from reducing applications to DI 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 capita per year and $p(\Omega_\tau)s/p(\Omega_b)$ is in units of government revenue per recipient per year. See Section B for details.

DRAFT

Table A5: Severity questions

Severity	
(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
Severe	if "Yes" in (1) and "A lot" in (2), or "Can do nothing" in (1)
Non-severe	if others

DRAFT

NOT FOR DISTRIBUTION/HAS NOT BEEN REVIEWED BY SSA

Table A6: 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.

DRAFT

NOT FOR DISTRIBUTION/HAS NOT BEEN REVIEWED BY SSA

Table A7: 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; physial 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.

DRAFT

NOT FOR DISTRIBUTION/HAS NOT BEEN REVIEWED BY SSA

Table A8: 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 jontly;mobile home ownders 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 mch would they amount to right now?
Help from relatives	Did you receive any help from relatives? [THIS MUCT BE FROM NON-FU MEMBERS.]
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 receie any income from unemployment compensation?
Workers compensation	Did you receive any income from workers compensation?

DRAFT

NOT FOR DISTRIBUTION/HAS NOT BEEN REVIEWED BY SSA

Table A9: Insurance question

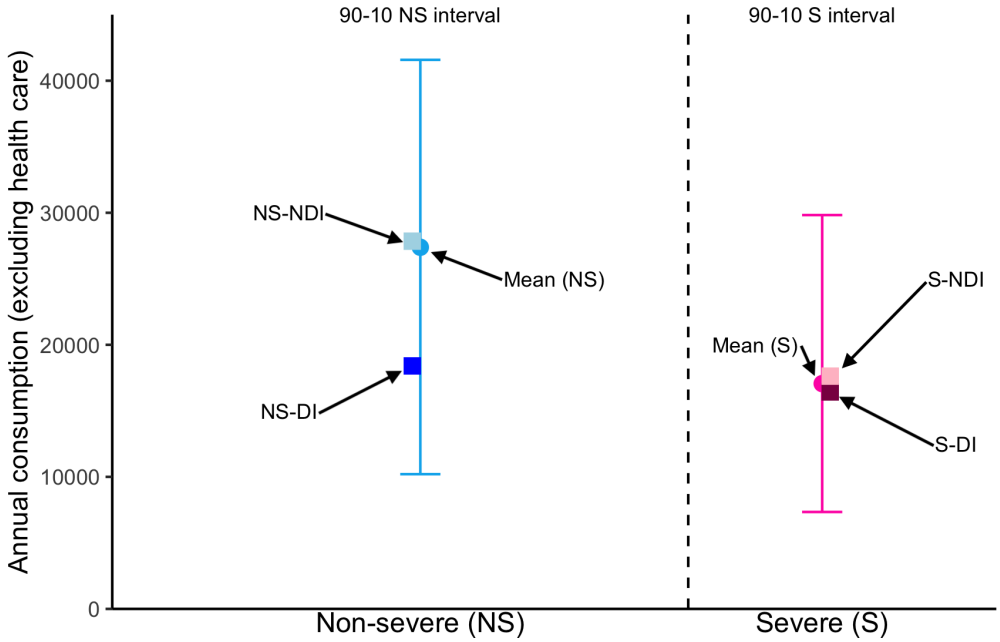
Insurance		What type of health insurance plan or health insurance coverage did you have in 2015 and 2016? If you had more than one type of plan or coverage please tell me about each of them.
Private	Public	
✓		1. Employer provided health insurance
✓		2. Private health insurance purchased directly
	✓	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. Oher state-sponsored plan (not Medicaid)
✓		10. Other government program
✓		11. Other

DRAFT

NOT FOR DISTRIBUTION/HAS NOT BEEN REVIEWED BY SSA

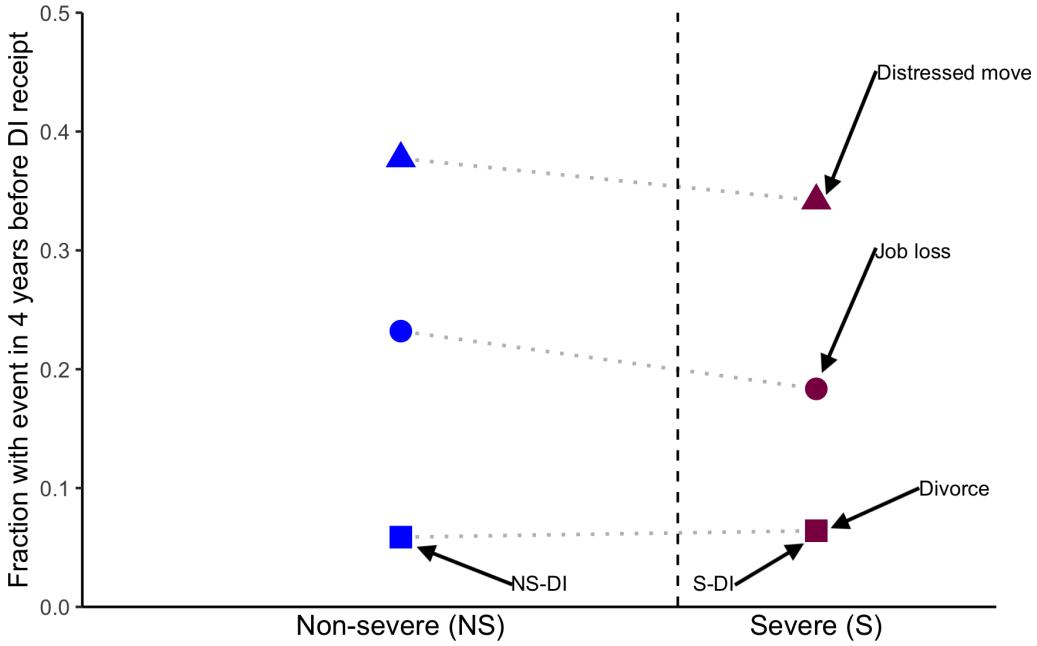
Appendix Figures and Tables for SSDI only (for online publication)

Figure B1: PSID: Consumption range (SSDI only)



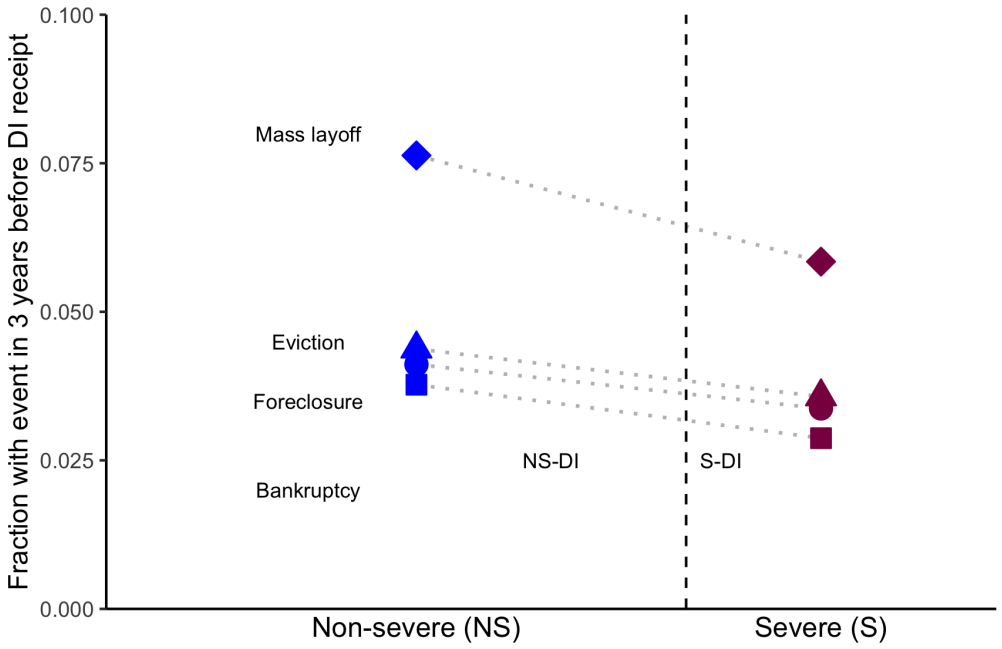
Notes: Figure presents the 90-10 percentile interval (average) of consumption for severe and non-severe households in the 2017 PSID: households with a severe health condition (566 S), and households without a severe health condition (6,754 NS). It also gives average consumption for households not currently receiving DI (SSDI only) without a severe health condition (6,442 NS-NDI), households not currently receiving DI with a severe health condition (316 S-NDI), households currently receiving DI without a severe health condition (312 NS-DI), and households currently receiving DI with a severe health condition (250 S-DI). Household consumption excludes health care and is divided by the square root of household size. "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 "non-severe." Square markers for the four groups are offset to facilitate easier reading of the graph, not as a representation of relative severity.

Figure B2: PSID: Adverse events prior to DI receipt by health status (SSDI only)



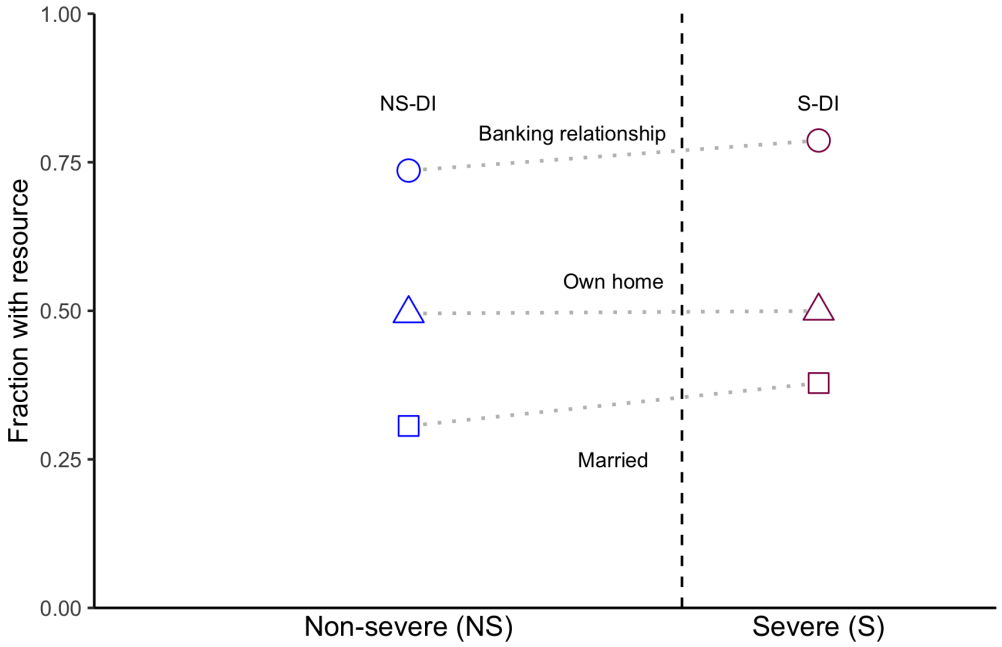
Notes: Figure presents rates of various adverse life events in the four years before receiving DI (SSDI only) for households receiving DI in the 2017 PSID, by severe and non-severe. Job loss includes "involuntary" reasons for separation: strike/lockout, laid off/fired, or company going out of business or leaving town (see Appendix Table A6). The sample for job loss is households currently receiving DI who can be observed in the four years before initial DI receipt (161 severe, 201 non-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 A7). The sample for distressed move is households currently receiving DI who can be observed in the four years before initial DI receipt (186 severe, 156 non-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 DI who can be observed in the four years before initial DI receipt (175 severe, 154 non-severe).

Figure B3: Admin: Adverse events prior to DI receipt by health status (SSDI only)



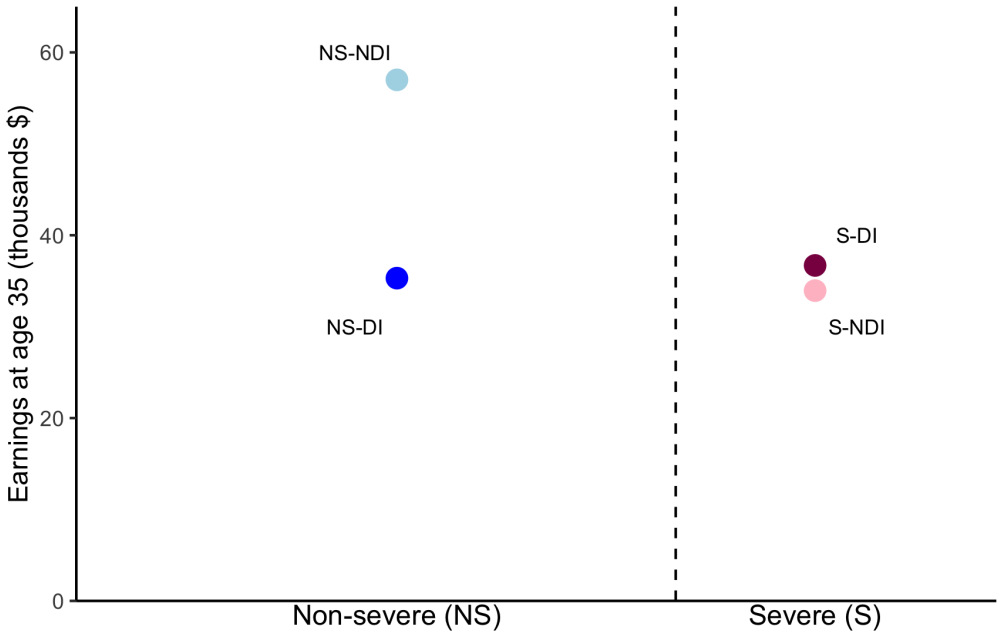
Notes: Figure presents rates of adverse financial events in the three years before receiving DI (SSDI only) receipt for individuals who apply for and receive DI benefits in SSA administrative data, by severe and non-severe. "Mass layoff" indicates being part of a mass layoff in the three years prior to DI application. The mass layoff sample is DI recipients in the CWHS-linked Master Beneficiary Record (MBR) who were allowed onto DI between 1985 and 2016 (420,000 less-severe, 800,000 more-severe). We determine rates of mass layoff using the Continuous Work History Sample (CWHS), which is a 10% sample of individuals in the US. In the CWHS, we identify EINs that experience a drop in employees of at least 30 percent from a base of at least 150 employees. From the MBR, we label a DI recipient as experiencing 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 DI application. We link DI applicants in the 831 files to nationwide financial records using the methods described in Deshpande, Gross and Su (2019). The bankruptcy sample is DI recipients in the 831 files between 1995 and 2009 (4.5M less-severe, 2.3M more-severe). The foreclosure sample is approved DI applicants in the 831 files between 2005 and 2014 who are homeowners (1.1M less-severe, 920,000 more-severe). The eviction sample is approved DI applicants in the 831 files between 2005 and 2014 who are not homeowners (870,000 less-severe, 520,000 more-severe). "Severe" indicates a medical diary of "medical improvement not expected," and "non-severe" indicates a medical diary of "medical improvement possible" or "medical improvement expected."

Figure B4: PSID: Outside Resources (SSDI only)



Notes: Figure presents rates of marriage, banking relationship, and homeownership for two types of households in the 2017 PSID: households currently receiving DI (SSDI only) with a severe health condition (250 S-DI), and households currently receiving DI without a severe health condition (312 NS-DI). Appendix Table A8 presents the PSID questions for banking relationship and homeownership. "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 "non-severe."

Figure B5: PSID: Earnings at age 35 (SSDI only)



Notes: Figure presents average earnings at head age 35 for four types of households in the 2017 PSID: households not currently receiving DI (SSDI only) without a severe health condition (1,967 NS-NDI), households not currently receiving DI with a severe health condition (108 S-NDI), households currently receiving DI with a severe health condition (93 S-DI), and households currently receiving DI without a severe health condition (75 NS-DI). Households are categorized based on their severity status and DI status in the 2017 PSID, and are included in the sample only if they are observed in the PSID at head age 35 and if they do not receive DI benefits at head age 35 or before. Results are qualitatively similar when we pool multiple years of the PSID to get a larger sample. "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 "non-severe."

Figure B6: PSID: Any adverse life event by DI status and age (SSDI only)

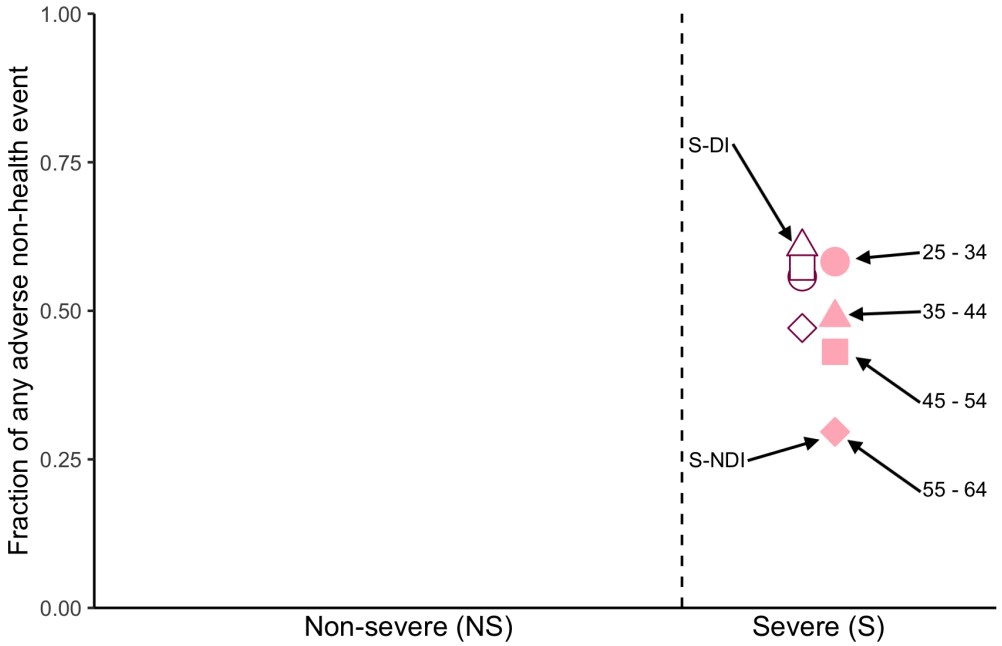
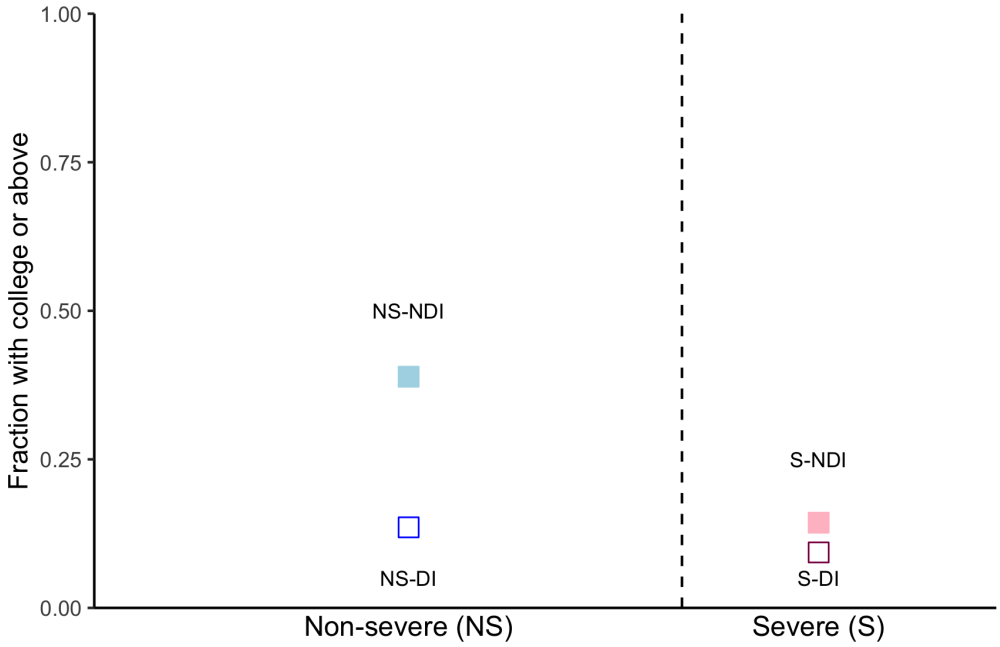


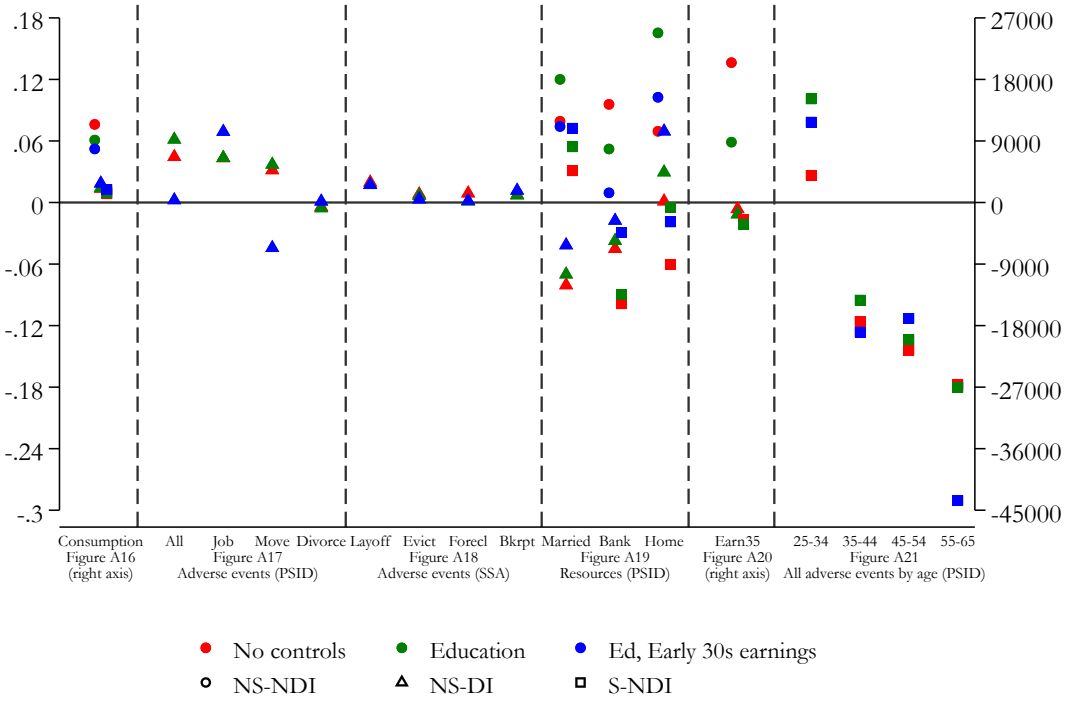
Figure presents rates of experiencing an "adverse life event"—head or spouse job loss, involuntary move, or divorce—at various ages for households in the PSID, by DI (SSDI only) status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI) and their age at entry, and NDI are further classified by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The figure plots the rate of experiencing an adverse life event in a specific age range (25–34, 35–44, 45–54, 55–64) among NDI and among DI who also entered DI at that age. The sample sizes for age groups 25–34, 35–44, 45–54, and 55–64 are 8,290, 7,379, 4,618, and 2,924, respectively, for NS-NDI; 1039, 1407, 1254, and 1039 for S-NDI; 159, 143, 192, and 309 for NS-DI; and 75, 172, 223, and 339 for S-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 A6). 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 A7). Divorce is defined as being married in a previous survey year but not in this survey year. "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 "non-severe." Markers for the two groups are offset to facilitate easier reading of the graph, not as a representation of relative severity.

Figure B7: PSID: College or above (SSDI only)



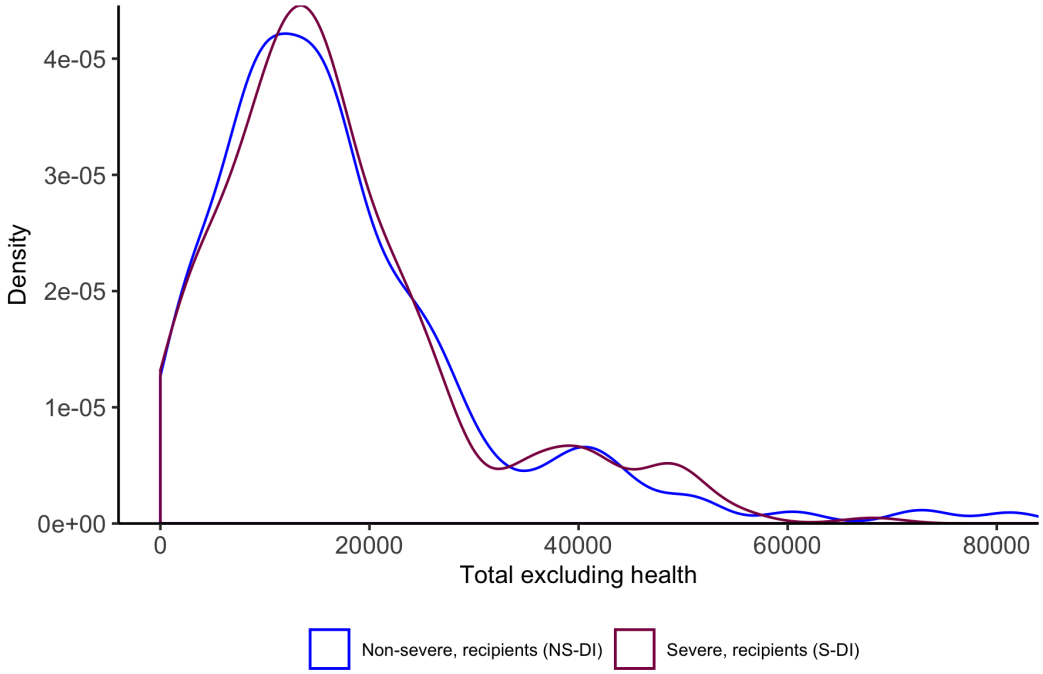
Notes: Figure presents educational attainment for four types of households in the 2017 PSID: households not currently receiving DI (SSDI only) without a severe health condition (6,442 NS-NDI), households not currently receiving DI with a severe health condition (316 S-NDI), households currently receiving DI with a severe health condition (250 S-DI), and households currently receiving DI without a severe health condition (312 NS-DI). "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 "non-severe."

Figure B8: Robustness to Ability-Group Controls (SSDI only)



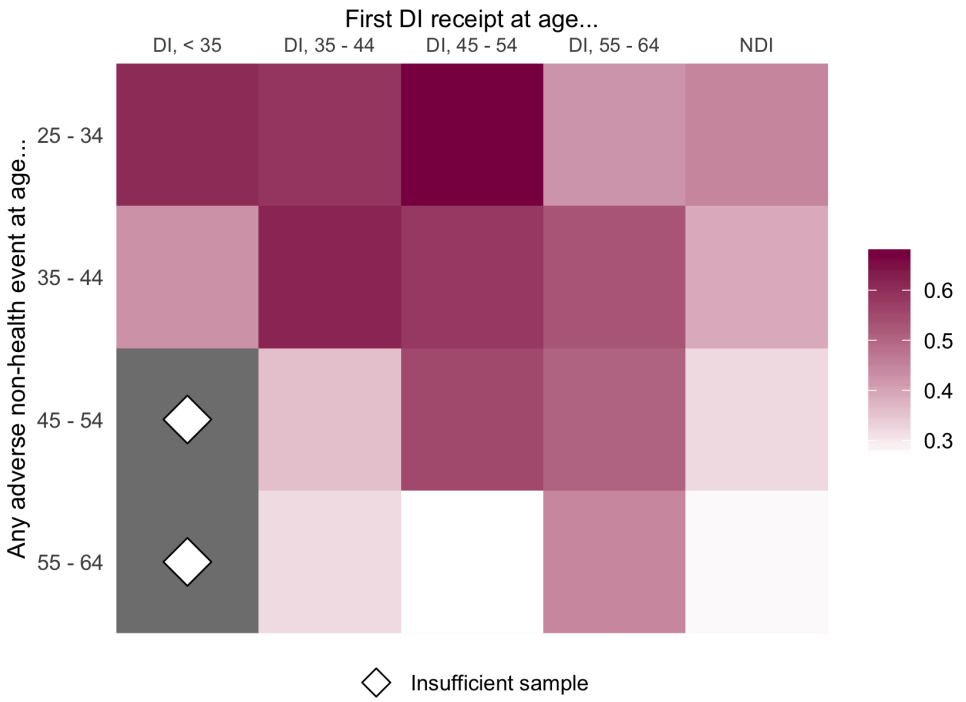
Notes: Figure presents all outcomes from Figures B1 to B6 with controls for ability group. The figure plots the coefficients on indicators for each DI-by-severity group (NS-NDI, NS-DI, and S-NDI; S-DI is the excluded group) from a regression of the outcome on those indicators and ability controls. The three groups are shown using different shapes. We use three specifications for ability controls: no controls (which reproduces the coefficients from Figures B1 to B6), 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 B2 indicates experiencing at least one of the adverse events (job loss, involuntary move, or divorce).

Figure B9: PSID: Consumption distribution at application (SSDI only)



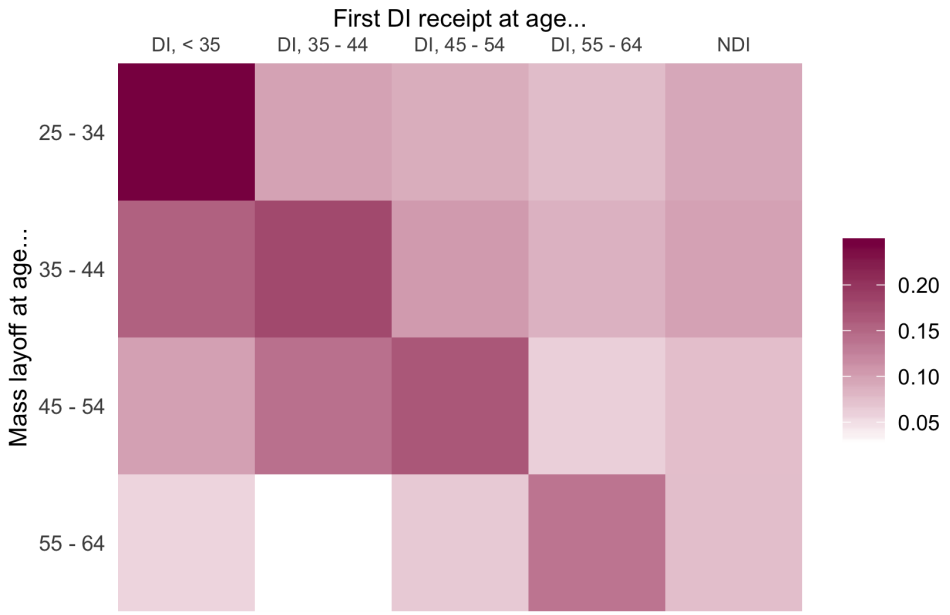
Notes: Figure presents the distribution of consumption (excluding health care expenditures) one year prior to the year of initial DI (SSDI or SSI) receipt for two types of households in the 2017 PSID: households currently receiving DI without a severe health condition (199 S-DI), and households currently receiving DI without a severe health condition (242 NS-DI). Household consumption excludes health care and is divided by the square root of the household size. "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 "non-severe."

Figure B10: PSID: Any adverse life event by DI status (SSDI only)



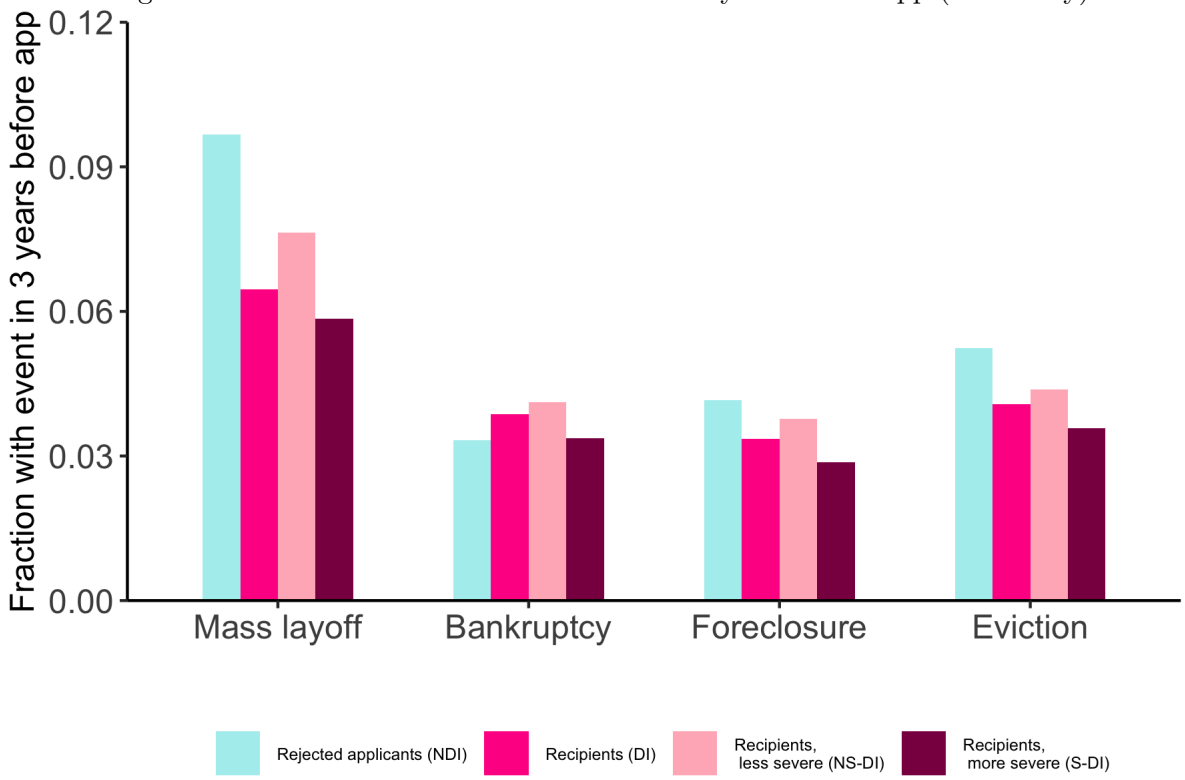
Notes: Notes: Figure presents rates of experiencing an "adverse life event"—head or spouse job loss, involuntary move, or divorce—at various ages for households in the PSID, by DI (SSDI only) status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). The DI group is further categorized by their age at DI 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 B11: CWHS: Mass layoffs by DI status



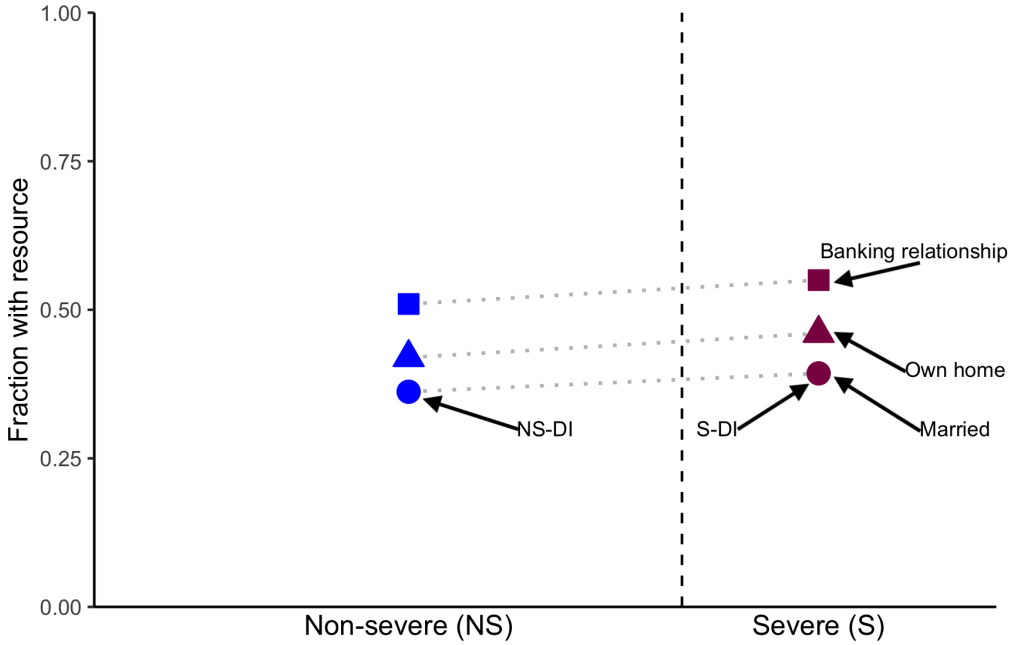
Notes: Figure represents rates of mass layoff at various age for individuals in the CWHS, by DI (SSDI only) status, and age at DI entry. Individuals are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). The DI group is further categorized by their age at DI 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 B12: Admin data: Financial events in 3 years before app (SSDI only)



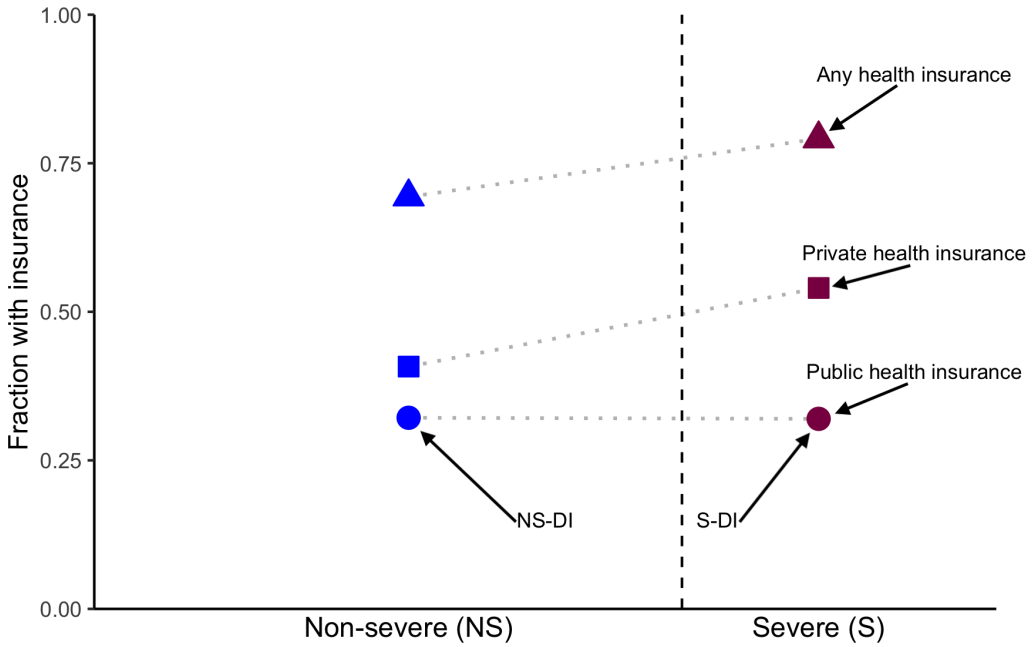
Notes: Figure presents rates of adverse financial events prior to DI (SSDI or SSI) receipt for rejected applicants and approved applicants (by more-severe and less-severe) in SSA administrative data. "Mass layoff" indicates being part of a mass layoff in the five years prior to DI application. The mass layoff sample is DI recipients in the CWHS-linked Master Beneficiary Record (MBR) who were allowed onto DI between 1985 and 2016 (420,000 less-severe, 800,000 more-severe), and who applied but rejected (810,000 rejected applicants). We determine rates of mass layoff using the Continuous Work History Sample (CWHS), which is a 10% sample of US workers. In the CWHS, we identify EINs that experience a drop in employees of at least 30 percent from a base of at least 150 employees. From the MBR, we label an DI recipient as experiencing 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 DI application. We link DI applicants in the 831 files to nationwide financial records using the methods described in Deshpande, Gross and Su (2019). The bankruptcy sample is DI recipients in the 831 files between 1995 and 2009 (4.5M less-severe, 2.3M more-severe), and those applicants who were rejected (4.4M rejected applicants). The foreclosure sample is approved DI applicants in the 831 files between 2005 and 2014 who are homeowners (1.1M less-severe, 920,000 more-severe), and those applicants who were rejected (970,000 rejected applicants). The eviction sample is approved DI applicants in the 831 files between 2005 and 2014 who are not homeowners (870,000 less-severe, 520,000 more-severe), and those applicants who were rejected (1.3M rejected applicants). "More severe" indicates a medical diary of "medical improvement not expected," and "less severe" indicates a medical diary of "medical improvement possible" or "medical improvement expected."

Figure B13: PSID: Outside Resources at application (SSDI only)



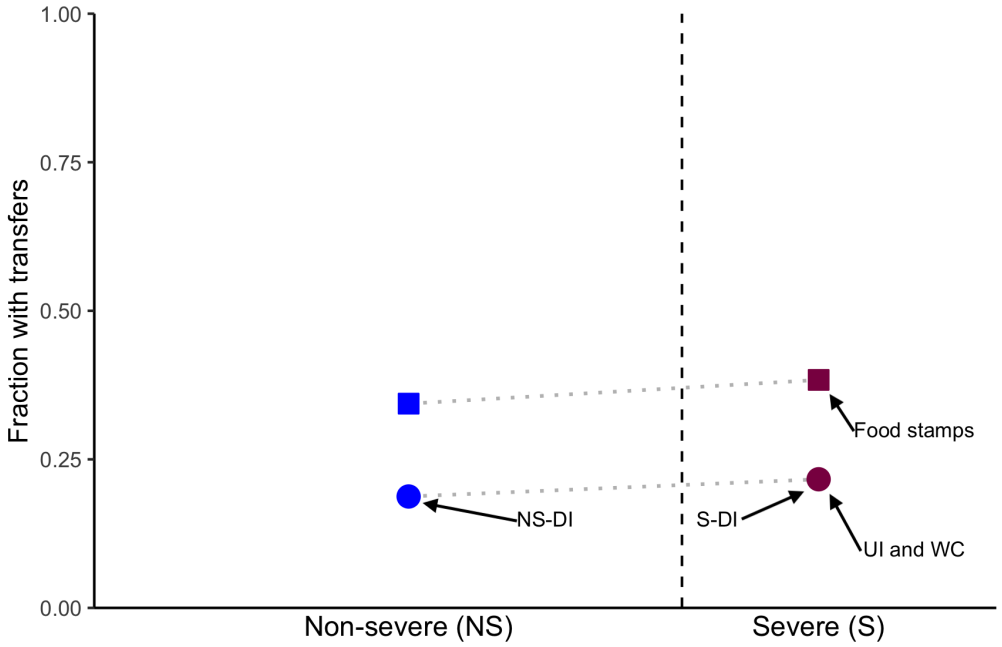
Notes: Figure presents rates of marriage, banking relationship, and homeownership one year prior to the year of initial DI receipt for two types of households in the 2017 PSID: households currently receiving DI (SSDI only) with a severe health condition (216 S-DI), and households currently receiving DI without a severe health condition (274 NS-DI). "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 "non-severe." Appendix Table A8 presents the PSID questions for banking relationship and homeownership.

Figure B14: PSID: Health Insurance Coverage (Head) at application (SSDI only)



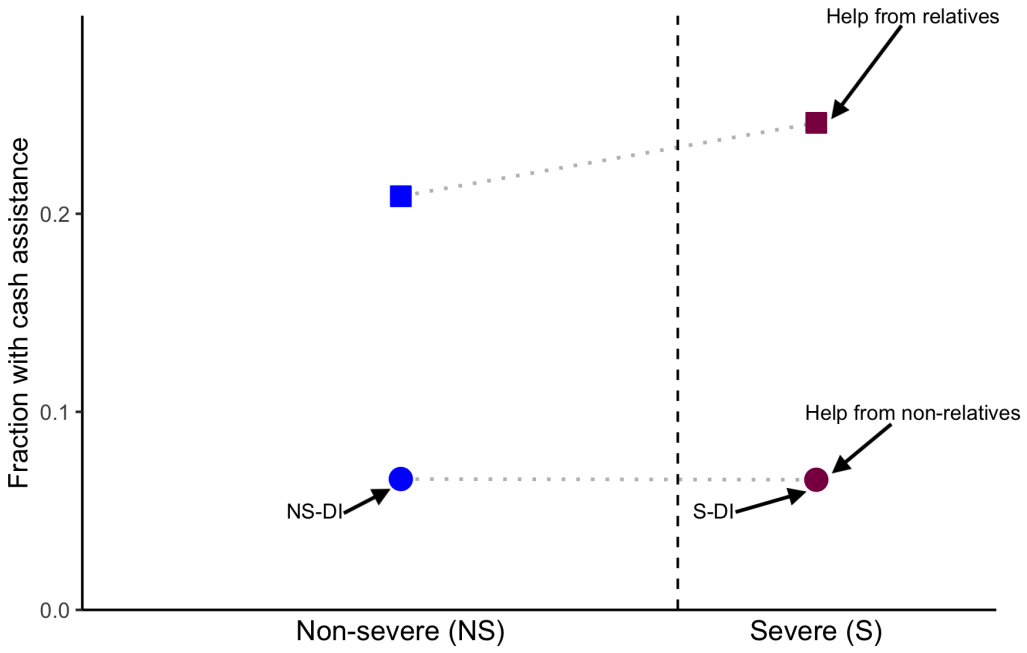
Notes: Figure presents rates of any health insurance, private health insurance, and public health insurance one year prior to the year of initial DI receipt for two types of households in the 2017 PSID: households currently receiving DI with a severe health condition (216 S-DI), and households currently receiving DI without a severe health condition (274 NS-DI). "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 "non-severe." Appendix Table A9 presents the PSID questions for private and public insurance.

Figure B15: PSID: Transfers at application (SSDI only)



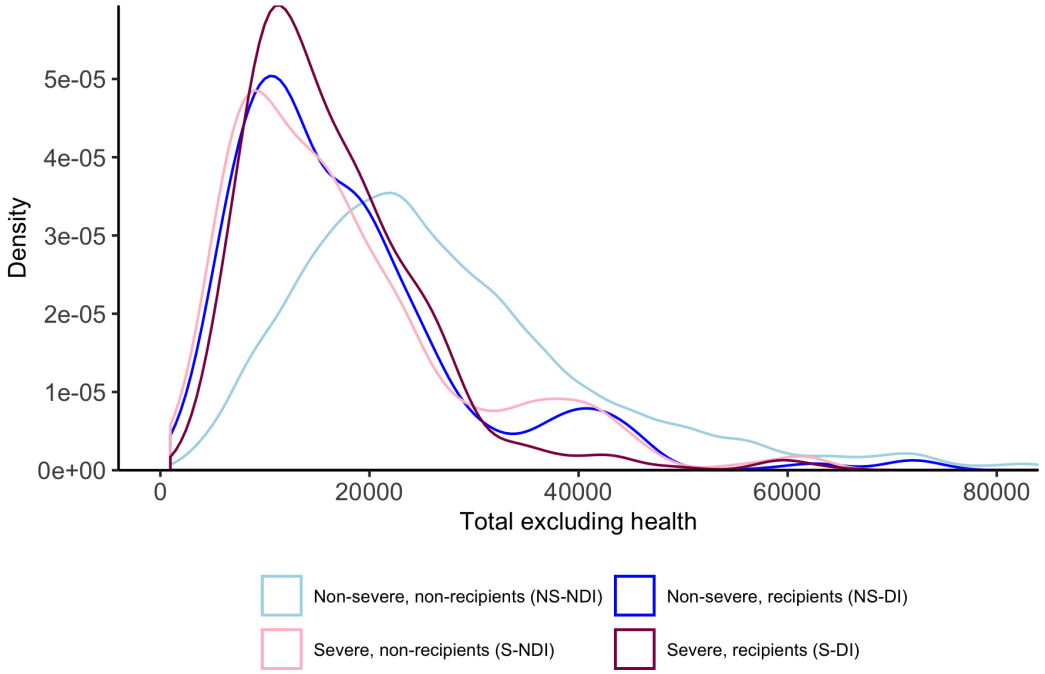
Notes: Figure presents rates of SNAP (food stamps) and unemployment insurance (UI) and workers' compensation (WC) one year prior to the year of initial DI (SSDI only) receipt for two types of households in the 2017 PSID: households currently receiving DI with a severe health condition (216 S-DI), and households currently receiving DI without a severe health condition (274 NS-DI). "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 "non-severe." Appendix Table A8 presents the PSID questions for SNAP, UI, and WC.

Figure B16: PSID: Cash assistance at application (SSDI only)



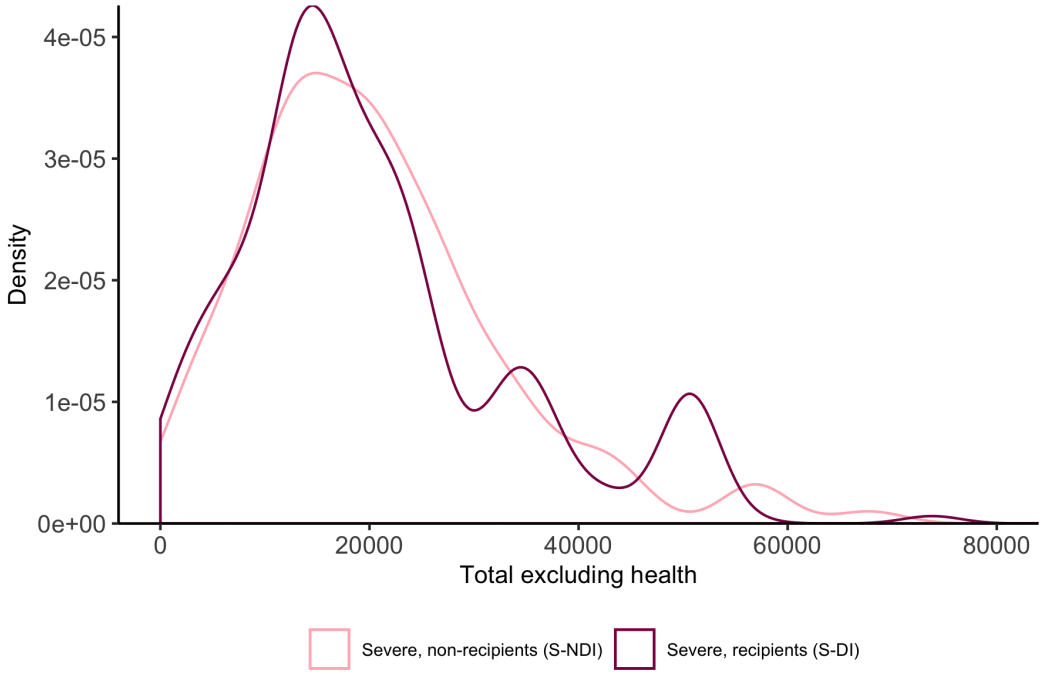
Notes: Figure presents rates of cash assistance from relatives and non-relatives one year prior to the year of initial DI receipt for two types of households in the 2017 PSID: households currently receiving DI (SSDI only) with a severe health condition (216 S-DI), and households currently receiving DI without a severe health condition (274 NS-DI). "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 "non-severe." Appendix Table A8 presents the PSID questions for cash assistance from relatives and non-relatives.

Figure B17: PSID: Consumption distribution (SSDI only)



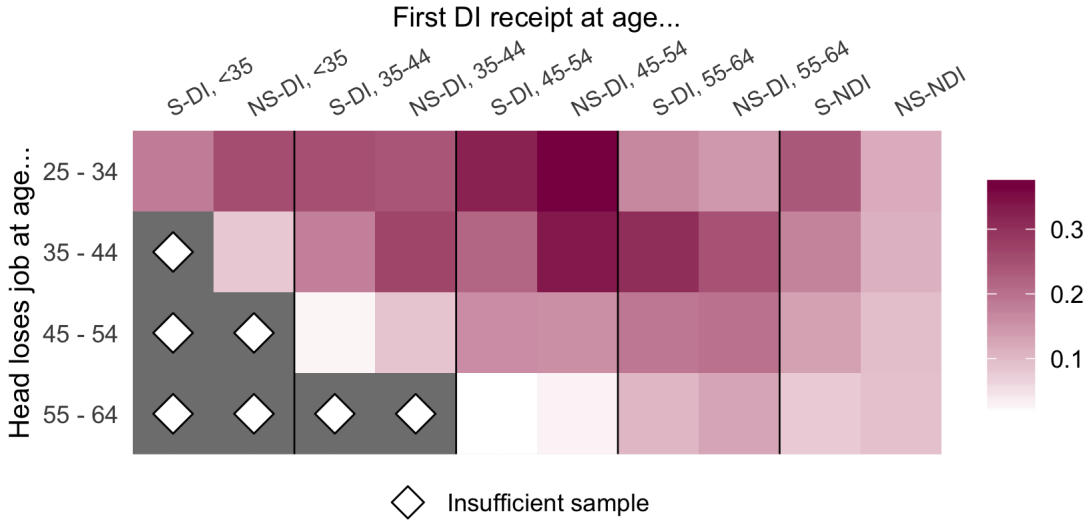
Notes: Figure presents the distribution of consumption (excluding health care expenditures) for four types of households in the 2017 PSID: households not currently receiving DI (SSDI only) without a severe health condition (6442 NS-NDI), households not currently receiving DI with a severe health condition (316 S-NDI), households currently receiving DI with a severe health condition (250 S-DI), and households currently receiving DI without a severe health condition (312 NS-DI). Household consumption is controlled and divided by the square root of the household size. "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 "non-severe."

Figure B18: PSID: Consumption distribution at severe onset (SSDI only)



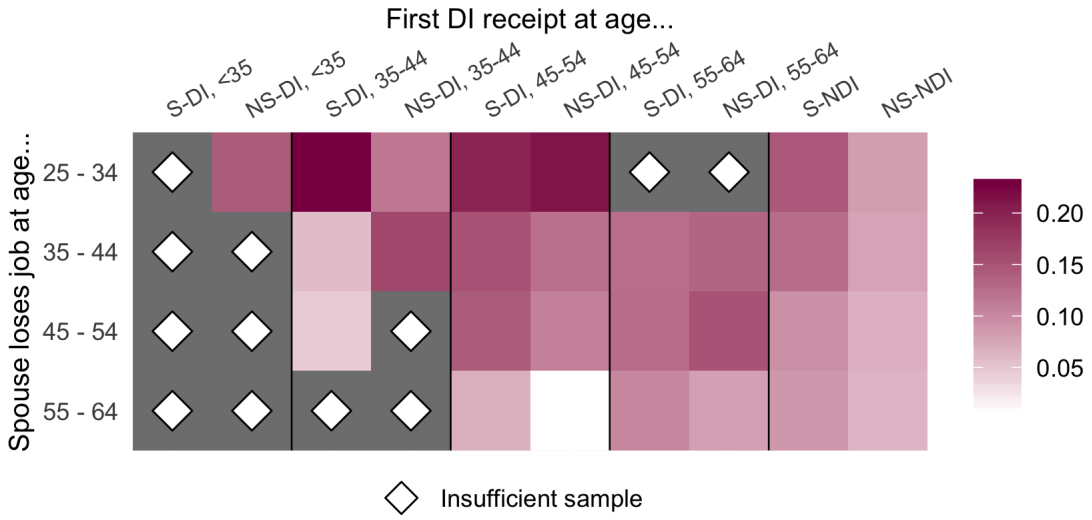
Notes: Figure presents the distribution of consumption (excluding health care expenditures) one year prior to the year of severe onset for two types of households in the 2017 PSID: households not currently receiving DI (SSDI or SSI) with a severe health condition (217 S-NDI), and households currently receiving DI with a severe health condition (166 S-DI). Household consumption excludes health care and is divided by the square root of the household size. "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 "non-severe."

Figure B19: PSID: Head job loss by DI and health status (SSDI only)



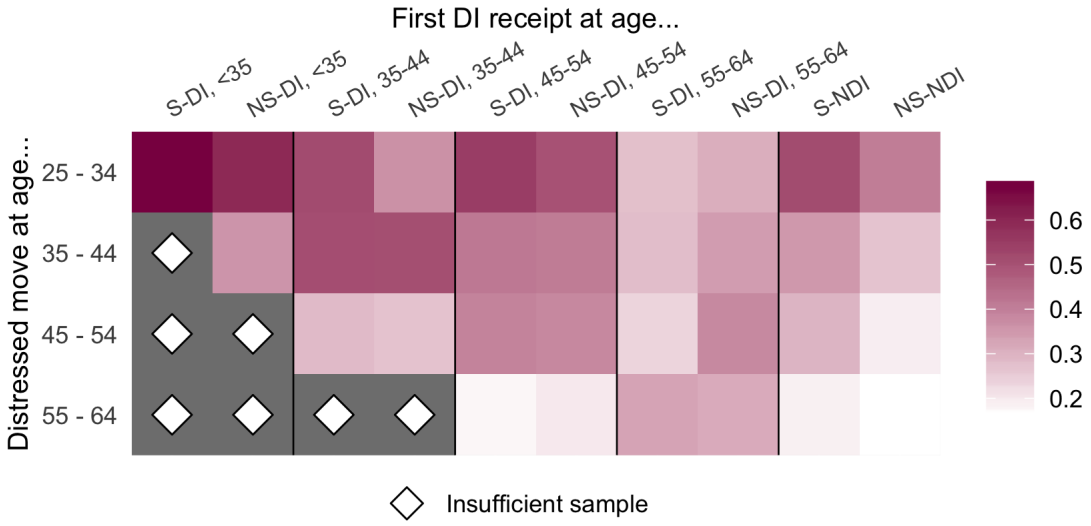
Notes: Figure presents rates of head job loss at various ages for households in the PSID, by DI (SSDI only) status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI), and NDI are further classified by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The S-DI and NS-DI groups are further categorized by their age at DI 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 (see Appendix Table A6). "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 "non-severe."

Figure B20: PSID: Spouse job loss by DI and health status (SSDI only)



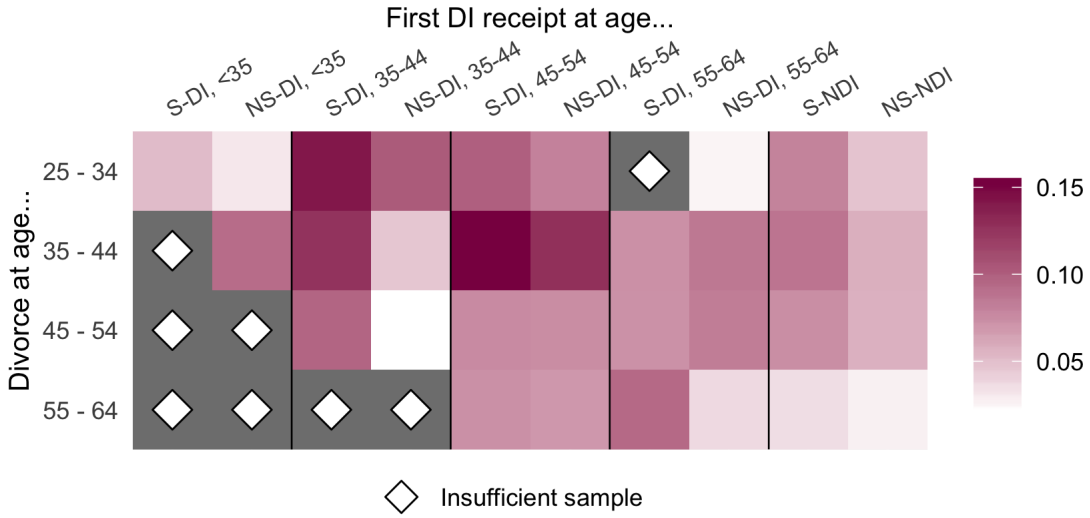
Notes: Figure presents rates of experiencing spouse job loss at various ages for households in the PSID, by DI (SSDI only) status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI), and NDI are further classified by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The S-DI and NS-DI groups are further categorized by their age at DI 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 (see Appendix Table A6). "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 "non-severe."

Figure B21: PSID: Distressed move by DI and health status (SSDI only)



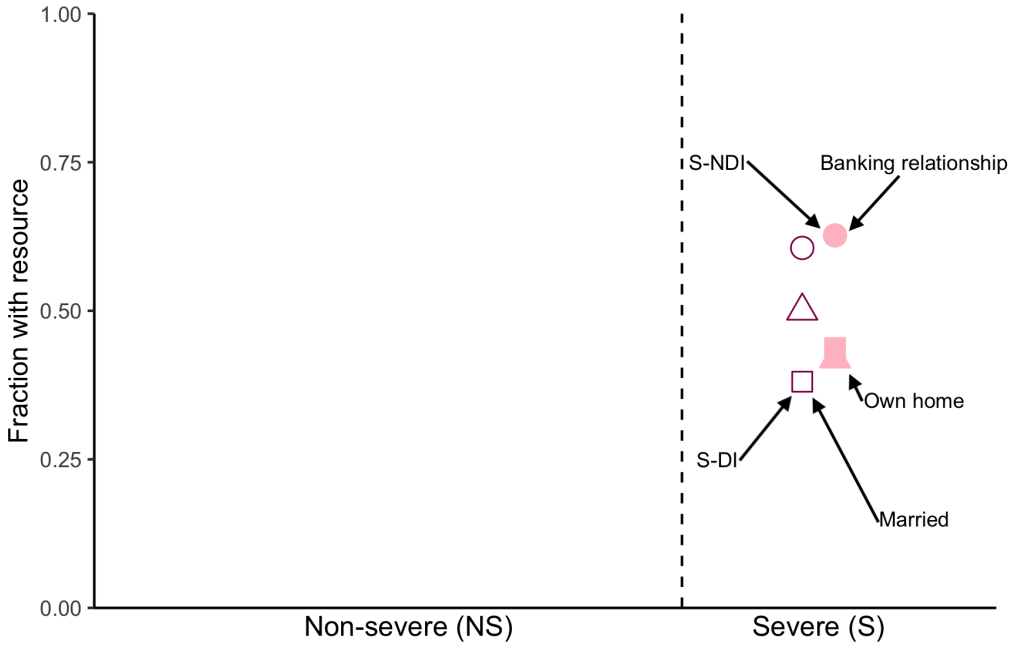
Notes: Figure presents rates of experiencing involuntary move at various ages for households in the PSID, by DI (SSDI only) status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI), and NDI are further classified by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The S-DI and NS-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 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 (see Appendix Table A7). "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 "non-severe."

Figure B22: PSID: Divorce by DI and health status (SSDI only)



Notes: Figure presents rates of experiencing divorce at various ages for households in the PSID, by DI (SSDI only) status, health status, and age at DI entry. Households are first categorized by their DI status: whether they ever received DI (DI) or never received DI (NDI). DI are further classified by their health status at DI entry (severe S-DI, or non-severe NS-DI), and NDI are further classified by whether they ever had a severe health condition (ever severe S-NDI, never severe NS-NDI). The S-DI and NS-DI groups are further categorized by their age at DI 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. Divorce is defined as being married in a previous survey year but not in this survey year (see Appendix Table A6). "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 "non-severe."

Figure B23: PSID: Outside Resources at severe onset (SSDI only)



Notes: Figure presents rates of marriage, banking relationship, and homeownership one year prior to the year of severe onset for two types of households in the 2017 PSID: households not currently receiving DI (SSDI only) with a severe health condition (207 S-NDI), and households currently receiving DI with a severe health condition (258 S-DI). "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 "non-severe." Appendix Table A8 presents the PSID questions for banking relationship and homeownership. Markers for the two groups are offset to facilitate easier reading of the graph, not as a representation of relative severity.