The Effect of Disability Insurance Receipt on Mortality

Bernard Black, Eric French, Jeremy McCauley and Jae Song*

March 1, 2018

Abstract

This paper estimates the effect of Disability Insurance and Supplemental Security Income benefit receipt on mortality, for those on the margin to receive benefits or not. Those receiving benefits receive large cash transfers, and health insurance from Medicare or Medicaid, but also face important work disincentives. Each of these factors could affect mortality. The income and health insurance benefits likely reduce mortality, but the work disincentive could increase mortality. Identifying the overall mortality effect is difficult, however, because those allowed benefits may be unobservably less healthy than those denied. We exploit the random assignment of judges to disability insurance cases to create instrumental variables that address this selection problem. We find considerable heterogeneity in the mortality response. For marginal recipients, who receive benefits if seen by lenient judges, but would be denied by stricter judges, we find no detrimental effects of being denied on mortality. Instead, we find higher mortality for these individuals within the first 10 years of benefit receipt, consistent with the view that working is beneficial for health. However, Marginal Treatment Effects estimates suggest that benefit receipt reduces mortality for inframarginal benefit recipients, who would receive benefits even if seen by a relatively strict judge. These findings suggest that for maximizing the longevity of DI applicants, the current disability thresholds are close to the right level.

*Comments welcome at bblack@northwestern.edu, eric.french.econ@gmail.com, jeremymccauley@gmail.com, and jae.song@ssa.gov. Affiliations are Northwestern University, Pritzker School of Law and Kellogg School of Management; University College London, Institute for Fiscal Studies; University College London; and Social Security Administration, respectively. We are grateful for comments from Camille Landais, Magne Mogstad, and Timothy Moore, participants from the IFS-ONS conference on administrative data, Conference on The Welfare State and the Labour Market: Micro meets Macro, UBC, ETC-Zurich, NBER Conference on Working Longer and Queen's University Belfast, and also funding from the Social Security Administration through the Michigan Retirement Research Center (MRRC grant UM17-15) and the Economic and Social Research Council (Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (RES-544-28-50001) and Grant Inequality and the insurance value of transfers across the life cycle (ES/P001831/1)). The views in this paper are not necessarily those of the above funding agencies or the Social Security Administration.

1 Introduction

This paper estimates the effect of Disability Insurance (DI) and Supplemental Security Income (SSI) benefit receipt on mortality, for those persons who would receive benefits if their case is heard by a lenient Administrative Law Judge (ALJ), but not if their case is heard by a stricter judge. We compare mortality rates of individuals who applied for and received disability insurance benefits to the mortality rates of those who applied for benefits but were denied. Those receiving benefits receive large cash transfers, and health insurance from Medicare or Medicaid. However, beneficiaries also face important work disincentives. Each of these factors could affect mortality. The income and health insurance benefits, taken together, likely reduce mortality, but the work disincentive could increase mortality. Identifying the overall mortality effect is difficult, however, because those allowed benefits may be unobservably less healthy than those denied.

Using Social Security administrative data, we exploit the essentially random assignment of DI cases to ALJs. We document large differences in allowance rates across judges, and show that these differences are unrelated to the health or earnings potential of DI applicants. We use judge specific allowance rates to construct an instrumental variable, which we call "judge leniency". We use our judge leniency instrument to predict the allowance of individual cases. We then use predicted allowance to estimate the effect of allowance on mortality. Because we have the population of DI applicants whose case was heard by an ALJ over 1995-2004, we can obtain precise estimates, even for relatively small subgroups of this population. We find heterogeneous effects. For persons aged 55-64 when assigned to an ALJ, DI benefit allowance increases the mortality rate, 10 years after assignment, by a statistically significant 2.81%, relative to a baseline 10-year mortality rate of 22.0%. This increase in mortality is surprising given that benefit allowance provides a cash benefit, which is likely to improve mortality, and health insurance, which may also improve mortality for this population. Balanced against this, allowance creates a large work disincentive. We find evidence suggesting that all three effects are important, with the net effect on mortality varying based on age and prior health.

This point estimate is a weighted average of Marginal Treatment Effects (MTEs) among those impacted by variations in judge leniency. We also estimate the distribution of MTEs within the observed range of judge leniency. Thus we can identify MTEs for healthier individuals who would be denied by almost all judges (apart from the most lenient) as well as less healthy individuals who would be allowed by almost all judges (apart from the strictest). Among the healthier individuals, DI receipt increases mortality. However, among the less healthy individuals, DI receipt tends to reduce estimated mortality, especially for those aged 55-64. Therefore, inframarginal individuals (who would be allowed benefits, whether seen by a strict or lenient judge and are the majority of DI recipients) likely benefit from DI receipt. Thus, our findings are consistent with the view that DI receipt reduces mortality on average.

Our results suggest that measured by impact on mortality, the current DI screening threshold is close to optimal. Making it less strict (and thus increasing the allowance rate) will likely increase mortality for marginal applicants, especially for those assigned to more lenient judges. Conversely, a modest tightening of the allowance rules, which brings them closer to the current decisions of stricter judges, should reduce mortality.

We find heterogeneous effects based on recipients' health conditions, although estimates are not precise. Benefit receipt lowers mortality among those with cancer, which is the highest mortality rate condition, and often requires expensive medical treatment that health insurance could help to fund. Benefit receipt also lowers mortality for recipients with respiratory and nervous system conditions, which are also high mortality conditions. Conversely, for recipients with conditions with lower medical spending, but likely strong labor supply effects, such as musculoskeletal disorders, benefit receipt predicts higher mortality.

We rely on Social Security Administration (SSA) mortality records. We compare these records to mortality records from the National Death Index (considered to be the best available source US mortality rates). SSA mortality records have historically been of suspect quality, but as we show, they have substantially improved in recent years. Mortality rates in the two databases are extremely similar for the age 55+ population: the SSA data appear to understate mortality rates for these persons by less than 1%, with larger understatement for younger age groups. Thus, our results are unlikely to be materially affected by any tendency for SSA being more likely to record deaths for DI recipients than for non-recipients. We assess the robustness of our results to potential under-reporting of mortality among those denied benefits, and find that allowing for potential under-reporting only slightly reduces our estimates of the effect of DI receipt on mortality, especially for persons aged 55-64.

Section 2 gives a literature review, section 3 describes the DI system, section 4 describes our estimation methods, and section 5 shows the data and discusses the data quality. In section 6 we present our main results, with Marginal Treatment Effect estimates being displayed in section 6.4. Section 7 discusses some channels by which DI receipt could impact mortality and how the effects vary by health condition. Section 8 shows that our estimates are robust to other specifications and methods of handling the data. Section 9 concludes.

2 Literature Review

Despite the great cost of the Disability Insurance program, relatively little research has been done on how the program affects the health and mortality of the disabled population. We might think that receiving benefits would impact health and mortality, since being allowed benefits impacts the health insurance, income, and employment of those receiving benefits. There is an active literature assessing the separate effects of health insurance, income, and work on health. In this section we review the evidence.

2.1 Income Effect from Receiving Disability Insurance

To the best of our knowledge, the only other paper to estimate the effect of DI on mortality is Gelber et al. (2017). They estimate the effect of Disability Insurance benefit income on mortality rates. They exploit the kinks in the DI benefit formulas. They measure the effect of benefit generosity on mortality, whereas we measure the effect of receiving benefits versus not receiving them. Receiving benefits not only affects income, but also affects health insurance, and affects labor supply incentives in a different way than receiving a slightly larger or smaller benefit. They find evidence that higher income benefits lead to lower mortality at the lower bend point of the DI benefit formula, but find no robust evidence of an effect at the upper bend point.

2.2 Health Insurance

Several important studies, including results from the RAND Health Insurance Experiment (Brook et al., 1983), analyses of Medicare (Finkelstein and McKnight, 2008; Card et al., 2009), and Medicaid (the Oregon Health Insurance Experiment) (Finkelstein et al., 2012) find that for the adult and elderly population, the near-term effect of health insurance on subsequent health outcomes is small. Card et al. (2009) find overall small, statistically insignificant effects of turning 65 and becoming Medicare eligible, but find that access to Medicare does modestly reduce mortality after emergency visits. Some studies do find significant effects of health insurance on mortality. For example, Sommers et al. (2014) finds that after the Massachusetts 2006 health care reform, which attained near-universal insurance coverage in the state, all-cause and health care-amenable mortality decreased when compared with similar counties in other states. Hernandez-Pizarro (2016) estimates the effect of a Spanish system of publicly-allocated long-term care benefits on mortality. She finds that access to greater benefits reduces mortality, particularly for those with only moderate needs. None of these studies focus on the disabled.

To the best of our knowledge, Weathers and Stegman (2012) is the only study that focuses on the value of health insurance for the disabled. They exploit a randomized experiment that reduced the wait time before DI recipients received Medicare benefits from 2 years to 0 years. They find no significant effect of immediate versus delayed receipt of health insurance on mortality; their point estimates imply higher mortality among those who received Medicare immediately.

These studies focus on short run effects, have limited sample sizes, or both. Thus, it is difficult to know if there is no average effect of health insurance on mortality, or if the sample size is too small or the sample period too short to detect an effect. Black et al. (2017) study longer-term effects, but also have a limited sample and use pure observational study methods, rather than a true or natural experiment. Using our data, we can estimate 10 year mortality rates for a large sample.

2.3 Income and Employment

Most papers that estimate an effect of income on mortality are estimating the joint effect on mortality of income from employment, and employment itself.

For example, Sullivan and von Wachter (2009) find that job loss significantly increases mortality, potentially reflecting loss of health insurance and loss of income. Several papers using European administrative data, such as Rege et al. (2009) and Eliason and Storrie (2006), find similar results.

Multiple papers have found that reductions in employment lead to poorer health and higher mortality. Fitzpatrick and Moore (2016) document a two percent increase in overall male mortality immediately after age 62, and suggest decreasing labor force participation as the possible key factor. To similar effect, Snyder and Evans (2006) assess the mortality effect of Social Security benefits for members of the "Social Security notch" cohort (those born in the years before 1917), who received benefits at a younger age than those born afterwards.

They find that the notch cohort had higher mortality rates and lower employment levels, and conclude that greater work effort has beneficial health impacts, which more than offset any mortality gains from greater Social Security income.

Several recent papers use European retirement reforms to estimate the impact of employment on mortality. While the evidence is mixed, the bulk of the evidence suggests that early retirement increases mortality. For example, Kuhn et al. (2017) find that an early retirement scheme in Austria led to higher mortality among males, with the higher mortality concentrated among heart diseases, diseases related to alcohol consumption, and vehicle accidents. This evidence suggests adverse changes in health behavior as a causal mechanism.

3 The Disability Insurance System

Social Security Disability Insurance (SSDI or DI) is one of America's largest social insurance programs. Furthermore, many disabled individuals with low income receive Supplemental Security Income (SSI) benefits. In 2014, 6.4% of people ages 18-64 and 16.3% of those aged 55-64 were receiving either DI or SSI benefits (U.S. Social Security Administration, 2014a).[1] Most DI and SSI beneficiaries also receive health insurance benefits through Medicare (for DI beneficiaries) or Medicaid (for SSI beneficiaries). The combined cost of these programs was $428 billion in 2008 (Livermore et al., 2011), making these programs several times more expensive than unemployment insurance. The costs have risen rapidly, generating many policy proposals to reform the system (Autor and Duggan, 2010; Burkhauser and Daly, 2011; Burkhauser et al., 2014).

3.1 Exit Rates from the DI program

Relatively few people lose disability benefits for reasons other than death.[2]

For example, of 7.1 million individuals (DI worker beneficiaries) drawing DI benefits in 2007, 0.5% had benefits terminated because they earned above the Substantial Gainful Activity (SGA) limit for an extended period of time in 2007. Another 0.3% had benefits terminated because they were deemed medically able to work after a continuing disability review, which is a periodic review conducted by SSA of the health of DI beneficiaries (U.S. Social Security Administration, 2007). Thus, the disability allowance decision is high stakes. If the individual is allowed benefits, that individual is typically given disability benefits until normal retirement age (age 65 during the 1990s and now 66), when the person becomes eligible for regular Social Security benefits.

[1] The percentage for persons aged 55-64 is based on authors' calculations using statistics from (U.S. Social Security Administration, 2014a) for the number receiving DI, (U.S. Social Security Administration, 2014b) for the number receiving SSI, and (U.S. Census Bureau, 2015) for population estimates. The total number of people in this age group receiving both DI and SSI is not reported by the SSA. We assume that the percentage of people receiving both is not dependant on age and therefore use the same percentage of 9.6% for those aged 18-64, which is reported in (U.S. Social Security Administration, 2014a).

[2] DI benefits are converted into retiree benefits once the beneficiary turns the normal retirement age. The statistics above are for DI benefits before the conversion to retiree benefits.

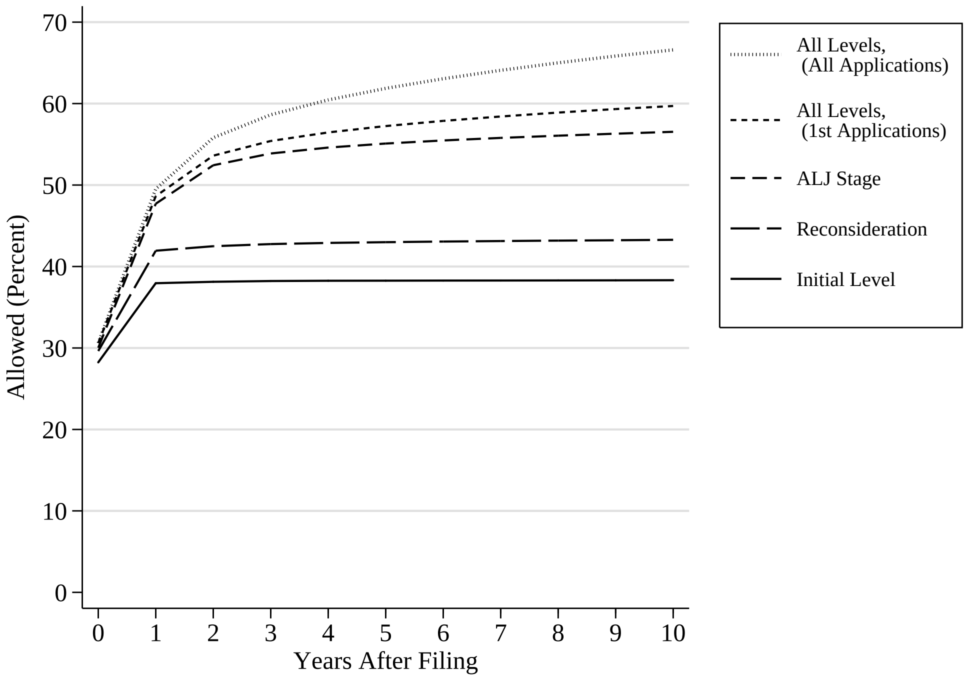


Figure 1: Allowance at different stages of the applications and appeals process.

3.2 Determining Eligibility for DI Benefits

An individual is deemed eligible for benefits if they meet certain work requirements and are deemed medically disabled. Although the exact algorithm is complex[3], one of two conditions must be met for the individual to be deemed disabled.

The first is a "listed impairment". Individuals who have one of over 100 specific listed impairments are given immediate benefits. Examples include statutory blindness (i.e., corrected vision of 20/200 or worse in the better eye) and multiple sclerosis.

The second condition is inability to work, either at their past work or other work. Eligibility under this condition turns on a combination of medical impairment and vocational factors such as education, work experience, and age. These cases can be especially difficult to evaluate. Myers (1993), a former Social Security Administration Deputy Commissioner, points out that if a worker "can do only sedentary work, then disability is presumed in the case where the person is aged 55 and older, has less than a high school education, and has worked only in unskilled jobs, but this is not so presumed in the case of a similar young worker. Clearly, borderline cases arise frequently and are difficult to adjudicate in an equitable manner!"

The disability determination is a multi-step process. Figure 1 shows the share of applicants who are allowed at different steps during our sample period. After an initial 5-month waiting period, DI applicants have their case reviewed by a Disability Determination Service review board. Figure 1 shows that 39% of applicants are allowed and 61% are denied at this stage. At this stage the most clear-cut cases are allowed, such as those with a listed

[3]See Hu et al. (2001) or Benitez-Silva et al. (1999) for details.

impairment. Cases that are harder to judge (such as musculoskeletal problems) are usually denied at this stage. About half of all applicants who are initially denied appeal at the disability determination service reconsideration stage. About 7.5% of those that appealed in 2013 were allowed benefits at this stage (U.S. Social Security Administration, 2014a). Sixty days after the disability determination service decision, a DI appeal can be requested. DI appeals are reviewed by Administrative Law Judges (ALJs) after a delay of about one year.[4] 14% of all initial claims, or 59% of all claims that are appealed, are allowed at the ALJ level.[5] If the case is denied at the ALJ level, the applicant can appeal to the SSA Appeals Council. If the applicant is denied at this level, she can then appeal after 60 days to Federal Court. However, Figure 1 shows that appeals at the higher levels are rarely successful: only about 2% of all initial claimants receive benefits at the Appeals Council or Federal Court level. Lastly, denied applicants can re-apply for benefits. The last line on Figure 1 includes those who re-apply for benefits. Another 7% of all initial claims are eventually allowed benefits through a re-application. 33% do not get benefits at any stage after 10 years.

Because we identify the causal effect of DI on mortality using variation at the ALJ level, the estimated effect applies only to marginal cases. The least healthy individuals, such as those with listed impairments, will almost always be allowed at the Disability Determination Service stage. The healthiest individuals will almost always be denied, whichever ALJ they see. Thus, our results are not generalizable to all DI applicants. However, the marginal cases are of great policy interest, because these are the individuals most likely to be affected by changes in the leniency of the appeals level of the DI system.

3.3 Assignment of DI Cases to Judges

Judicial independence means that judges have a great deal of latitude to determine eligibility (Taylor, 2007). As a result, two different judges can have very different allowance rates even though they see similar applicants.

Administrative Law Judges (ALJs) are assigned to hearing offices, and within hearing office, hear cases on a rotating basis.[6] When a judge finishes a case, that judge received the oldest pending case at his or her hearing office. Therefore, for applicants who apply at a

[4]Judges can make one of three decisions: allowed, denied, or remand. A "remand" is a request for more information from the disability determination service. Our measure of "allowed" is the final determination at the ALJ stage, and thus includes the final decision on remands.

[5]The allowance rate varies by age, and is significantly higher, at 84%, for those age 55-64, who are the principal focus of this study.

[6]Title 5, Part III, Subpart B, Chapter 31, Subchapter I, Section 3105 of the US Code states that "Administrative law judges shall be assigned to cases in rotation so far as practicable" (United States, 2007). The Social Security Administration's Hearings, Appeals and Litigation Law Manual (HALLEX) Volume I Chapter 2 Section 1-55 states that "the Hearing Office Chief Administrative Law Judge generally assigns cases to ALJs from the master docket on a rotational basis, with the earliest (i.e., oldest) Request for Hearing receiving priority." (U.S. Social Security Administration, 2009). HALLEX gives 11 exceptions to this rule. For example, the exceptions include "critical cases", such as individuals with terminal conditions and military service personnel, as well as remand cases. These cases are expedited and reviewed by Senior Attorneys. If there is a clear cut decision to be made, then the Senior Attorney will make the decision without a hearing. If the case is not clear cut, then the case is put back in the master docket and is assigned to a judge in rotation. We can identify cases that were decided without a hearing and delete them from our sample. We study the remaining cases where there was a hearing.

given office at a given point in time, the assignment of cases to ALJs is "essentially random" (Social Security Advisory Board, 2006). Judges do not pick the cases they handle. Judges are not assigned cases based on the expertise of the judge. Furthermore, an applicant cannot choose an alternate judge after being assigned a judge.

The initially assigned judge is not necessarily the judge who decides the case. Paletta (2011) documents a judge who took assigned cases from other judges and made decisions on those cases. Thus, the cases were not randomly assigned to the deciding judge.[7] We have information on the assigned judge in addition to the deciding judge. Although the deciding judge is not necessarily randomly assigned, the initially assigned judge is. We use initial assignment to a judge as our source of exogenous variation. The initially assigned judge is the deciding judge in 96% of all cases.

As we confirm below, the assigned judge is for all practical purposes randomly assigned conditional on hearing office and day. However, individuals are not randomly assigned to hearing offices. The zip code in which a person lives determines the hearing office to which they are assigned. Applicant characteristics can vary by location (e.g., black lung disease is more common near mining towns) as well as across time (e.g., the share of DI applicants listing mental illness as the main health problem has risen over time). For this reason we condition on hearing office and day in the estimations below. In doing so, we exploit only within hearing office-day variation in judge level leniency. This variation should be essentially random.

4 Estimating Equations

To estimate the effect of DI allowance on mortality, we use a two-step procedure. In the first step we generate an instrumental variable that is a measure of relative judge leniency, within a given hearing office and hearing day. This variable is correlated with the probability of allowance, but is independent of applicant health and other characteristics. In the second step we use instrumental variables procedures to estimate the effect of DI on mortality, as well as other factors that potentially affect mortality, such as employment, earnings and benefits. We focus principally on applicants age 55-64 at time of application, because SSA death records are more accurate for older applicants, which we explain in more detail in section 5, but also present results for younger applicants.

4.1 Basic Specification

Our basic estimating approach is a modified instrumental variables regression where in a first stage we estimate

$$A_i = j_i\gamma + X_i\delta_A + e_i. \tag{1}$$

[7] Furthermore, an individual can potentially reject the assigned judge. For example, if an individual misses her court case, she may be reassigned to a different judge. Also, some cases in remote areas are held via video conference where the judge and claimant are not in the same room. Claimants can demand that the judge be present at a hearing, and thus the judge must travel to the claimant. Some judges refuse to travel, and thus another judge will be reassigned to the case.

where A_i is a 0-1 indicator equal to 1 if individual i is allowed benefits by the ALJ, j_i are judge indicator variables (equal to 1 if judge j heard individual i's case), and X_i are hearing office-day indicators (equal to 1 if individual i's case is assigned on that hearing office-day pair). In some specifications we add further covariates such as gender, age, race, past income, legal representation, application type (SSDI or SSI), education, and main health condition of the individual. For the second stage we adopt the random coefficients model of Bjorklund and Moffitt (1987):

$$y_{i\tau} = A_i \phi_{i\tau} + X_i \delta_{y\tau} + u_{i\tau} \tag{2}$$

where $y_{i\tau}$ is mortality (or another outcome variable such as earnings, participation, appeals or allowance), τ years after after assignment to an ALJ. We allow for heterogeneity in the parameter $\phi_{i\tau}$ to capture heterogeneity in the effect of benefit receipt on outcomes, both across individuals and over time. We allow the variables $u_{i\tau}$ and $\phi_{i\tau}$ to be potentially correlated with A_i, and with each other.

We focus on the effect of ALJ allowance *at first hearing* on mortality and other outcomes after 5 years and 10 years. ALJ allowance after a first hearing and eventual allowance can differ because some people denied by an ALJ are allowed upon reapplication or appeal (as shown in Figure 1). We use ALJ allowance at first hearing rather than eventual allowance because those who die soon after this hearing cannot reapply or appeal: eventual allowance is thus itself a function of mortality, creating a spurious correlation between eventual allowance and mortality. This problem is circumvented by using ALJ allowance.

4.2 Estimating Equations

When estimating equation (2) we are confronted with three concerns. First, we wish to allow for heterogeneity in the parameter $\phi_{i\tau}$. Second, we have 1,404 judges in our sample, each of whom is a potential instrument. IV estimators can suffer from small sample bias when both the number of instruments and the number of observations is large (e.g., Hausman et al. (2012)). Third, we have just under 200,000 hearing office-day interactions in the covariate set X_i.

To solve these three concerns, we first construct the judge-specific allowance rate of the judge who heard individual i's case, averaged over all cases other than individual i's case. Formally this is

$$Z_i = \frac{1}{N_j - 1} \sum_{s \in \{J\}, s \neq i} A_s \tag{3}$$

where N_j is the number of cases heard by judge j_i over the sample period, and $\{J\}$ is the set of cases heard by judge j_i. This has been used as an instrument by Maestas et al. (2013) Dahl et al. (2014), and Autor et al. (2015), for example. We then de-mean this object by hearing office and day, creating $\widetilde{Z}_i$. In what follows "$\sim$" represents a de-meaned variable (e.g., $\widetilde{Z}_i = Z_i - \overline{Z_i}$ where $\overline{Z_i}$ is the mean value of Z_i on all cases that were assigned on the same day and at the same hearing office as case i).

Thus our instrument compares the fraction of cases allowed by judge j with the corresponding average probability for all other judges in the same office-day.[8] We refer to our

[8] Doyle Jr (2007) and French and Song (2014) construct a slightly different judge leniency variable–this

instrument as judge leniency. Judge leniency will be positive (negative) to the extent that a judge is more (less) likely to allow than other judges making decisions in that same office-day. Because we remove observation i, estimated judge leniency is independent of e_{it} or $u_{i\tau}$, even in a small sample.

Finally, we estimate the equations

$$\widetilde{A}_i = \lambda \widetilde{Z}_i + \epsilon_i, \tag{4}$$

$$\widetilde{y}_{i\tau} = \phi_\tau \widehat{\widetilde{A}}_i + \widetilde{u}_{i\tau} \tag{5}$$

jointly using two stage least squares. In specifications where we include additional covariates we also demean each covariate by hearing office and day.

Given the above assumptions, the estimated effect can be interpreted as a Local Average Treatment Effect (LATE). The object we identify is not technically a LATE, since a LATE assumes a binary instrument, whereas our instrument is continuous. However, some papers refer to this as a LATE. More precisely, our procedure identifies a weighted average of $\phi_{i\tau}$ for the individuals affected by the instrument (see Heckman et al. (2006) and French and Taber (2011) for more details).

We identify the LATE if three conditions are met. First, if judges are randomly assigned to cases, conditional on date and hearing office, then assignment satisfies the "independence assumption". Second, if judges differ only in leniency and rank applicants the same with respect to relative severity of their disability, then the Imbens and Angrist (1994) "monotonicity assumption" is satisfied. The monotonicity assumption implies that a case allowed by a strict judge will always be allowed by a more lenient one. Third, we assume that the instrument causes variation in allowance rates, sometimes known as the rank or existence condition. Sections 6.1 and 6.2 provide evidence on the extent to which these assumptions hold.

4.3 Marginal Treatment Effects

We are interested both in the LATE – the average effect of allowance for the marginal cases for which we can identify this effect – and also how the treatment effect varies with judge leniency, within the range of leniencies that we observe. Section 6.4 presents estimated Marginal Treatment Effects (MTEs), which measure how the mortality response varies with (de-meaned) allowance rates. We use a polynomial estimating equation to estimate the MTE. Heckman et al. (2006) experiment with different approaches to estimating the MTE, such as local polynomial smoothers. They find that the polynomial approach works about

alternative approach is described in appendix C.2. When we replace $\widetilde{Z}_i$ with their instrument we obtain similar results (see Section 8).

as well as other procedures.[9] We estimate the equations

$$\widetilde{A}_i = \sum_{k=1}^{K} \lambda_k (\widetilde{Z}_i)^k + \eta_i, \tag{6}$$

$$\widetilde{y}_{i\tau} = \sum_{k=1}^{K} \varphi_{k\tau} (\widetilde{\widehat{\widetilde{A}}_i})^k + \mu_{i\tau} \tag{7}$$

where $\widehat{\widetilde{A}}_i$ in equation 7 is the predicted value of $\widetilde{A}_i$ from equation (6), and K is the order of the polynomial.

As shown by Heckman et al. (2006) and French and Taber (2011), as well as appendix C, the estimated MTE(a) is

$$\sum_{k=1}^{K} k\varphi_{k\tau} (\widetilde{\widehat{\widetilde{A}}_i})^{k-1} = \hat{E}[\phi_{i\tau} | \text{allowed only if } \widehat{\widetilde{A}}_i \geq a, \text{not allowed if } \widehat{\widetilde{A}}_i < a,] \tag{8}$$

where a is a particular realization of the (de-meaned) allowance rate. Equation (8) shows that MTE(a) is the mean value of $\phi_{i\tau}$ for those who would be allowed if their assigned judge allowed slightly higher than a share a of cases, and would be denied if assigned to a judge allowing slightly lower than this share. As a increases, we estimate the effect of the instrument for individuals with less severe disability. Appendix section C.1 provides more details on interpretation and estimation of the MTE.

5 Data

Our initial sample is all individuals aged 25-64 who appealed either a DI or SSI initial benefit denial, and were assigned to an ALJ during 1995-2004. Using Social Security Numbers, we match together data from the SSA 831 file, the Office of Hearings and Appeals Case Control System (OHACCS), the Hearing Office Tracking System (HOTS), the Appeals Council Automated Processing System (ACAPS), the Litigation Overview Tracking System (LOTS), the Master Earnings file (MEF), and mortality data from the Numerical Identification file (NUMIDENT). These data are described in greater detail in the appendix. We study mortality outcomes up to 10 years following assignment to a judge. Thus, our mortality data run from 1995 to 2014.

We drop all observations heard by a judge who heard less than 200 cases during the sample period. We also drop cases with missing education information. Table A1 in Appendix A presents more details on sample selection criteria.

Those who die before their case was heard may possibly be recorded as "not allowed," which could inflate near-term mortality for those denied benefits. To address this problem we drop all cases where the individual died before her case was heard. In addition, to address any mismeasurement in whether a case was heard before death, we also drop 30,807 cases

[9] Our Monte Carlo simulations suggest there is very little bias when using polynomials. Furthermore, the polynomial procedure is computationally feasible with large numbers of covariates, such as a full set of hearing office-day interactions.

where the individual died in the year of assignment to an ALJ. This selection decision has only a modest effect on our estimates, which is shown in robustness checks in section 8. Our full estimation sample has 2,759,907 DI or SSI cases heard by 1,436 judges, with a mean allowance rate [*BSB: Isn't this the "eventual" allowance rate, after appeals?] at the ALJ stage of 70.8%. Our main estimation subsample of those ages 55-64 includes 610,231 cases, with a mean allowance rate at the ALJ stage of 84.1%. All dollar amounts below are in 2014 dollars, deflated by the CPI.

Cases in our sample were heard on 195,935 hearing office-day pairs. Thus, on an average $2,759,907/195,935 = 14.1$ cases were heard at each hearing office-day pair. Although we have a large number of hearing office-day fixed-effects, consistency in fixed effects estimators depends on the number of observations going to infinity, not the number of observations per fixed effect going to infinity. A non-trivial number of cases were heard when there was only a single judge at the hearing office on that day. These observations do not contribute any identifying variation.

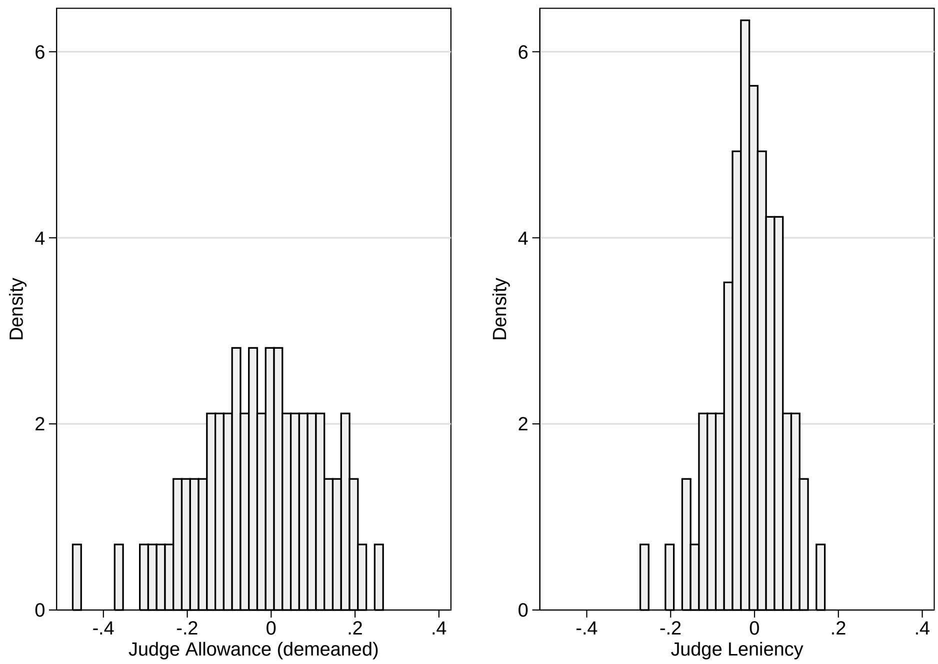


Figure 2: Allowance rate of ALJs, de-meaned (left panel), and de-meaned by hearing office and day (right panel).

Figure 2 plots the distribution of judge specific allowance rates, both unconditional (left panel) and also the judge leniency variable constructed in section 4.2 , which is conditional on hearing office-day (right panel). There is less variation in allowance rates after conditioning on hearing office and day; the standard deviation for the unconditional judge allowance rate is .149, but the standard deviation of the judge leniency variable is .096 (weighted by the number of cases handled by each judge). This means that being assigned to a judge one standard deviation more lenient than the office-day average increases the probability of allowance by 9.6 percentage points.

5.1 SSA Mortality Data

A core data issue for this study is the quality of the SSA mortality data, which comes from SSA's confidential NUMIDENT file. The SSA uses these data to process DI, SSI, and Social Security benefits. These data have been extensively used in previous research, but differ from the data used to construct the official mortality statistics for the US. The SSA obtains death records from various sources, including states, family members, funeral directors, post offices, financial institutions, and other federal agencies. It has a financial incentive to record deaths, especially if it was paying benefits to that individual.

One concern is that because the SSA has a greater financial incentive to record deaths of beneficiaries than non-beneficiaries, it will do a better job of capturing deaths of those who are allowed benefits than those who are denied benefits. Furthermore, it is easier for the SSA to measure deaths of beneficiaries, because if it sends payments to a deceased beneficiary, institutions such as banks (if benefit payments are electronically deposited) or post-offices (if benefit payments are sent by mail) will be more likely to report the death to the SSA (see GAO (2013) for information on how death information is collected). Any undercounting of deaths of those denied benefits will bias down their estimated mortality and could make it appear as if receiving benefits causes higher mortality. As we show below, however, SSA undercounting of deaths, which was formerly a major concern, is no longer an important issue, and should have at most a small effect on our estimates.

Older studies using older versions of the SSA data have shown that the SSA mortality data understates the National Death Index (NDI) data (which are considered the "gold standard" of US mortality data).[10] For example, Hill and Rosenwaike (2002) show that, in the years 1995-1997, the SSA capture approximately 80% of all deaths in the 55-64 year old population, and 95% of all deaths of those 65 and older. However, in separate work (Black et al., 2016), summarized below, we show that the SSA data have greatly improved in recent years, including retroactive updating for prior years.

We estimate the ratio of deaths in the SSA data to NDI deaths over 1995-2014, by age group. We construct these statistics to be as comparable as possible to the SSA data. Thus, we adjust the NDI data to include deaths of people in US territories and exclude foreign residents in the US, because the SSA data includes deaths of US nationals living abroad, whereas the NDI data does not. See Black et al. (2016) for details.

The top panel of Table 1 shows our estimates for the years 1995-2014, which is our sample period. Estimates broken down by each year can be found in Appendix Table A2. The estimates show that the SSA data have improved considerably relative to the estimates shown in Hill and Rosenwaike (2002). The SSA data capture 98% of all deaths over the 1995-2014 period, and is very close to complete for those over age 55. Indeed, in recent years, SSA data capture somewhat more deaths than the NDI for persons over 65. However, the

[10] The NDI is maintained by National Center for Health Statistics, and made available to researchers through the National Association for Public Health Statistics and Information Systems (NAPHSIS). These data are used to construct the Vital Statistics data for the US.

	All (20+)	20-44	45-54	55-64	65+
Estimated Ratio of Deaths in SSA to NDI data					
1995-1999	0.970	0.944	0.965	0.969	0.973
2000-2004	0.981	0.945	0.975	0.989	0.983
2005-2009	0.991	0.947	0.976	0.993	0.996
2010-2014	0.995	0.957	0.976	0.990	1.000
Average	0.948	0.948	0.973	0.985	0.985
Estimated Ratio of Non-Beneficiary Deaths that are Reported (p)					
1995-1999	--	0.929	0.948	0.955	--
2000-2004	--	0.919	0.962	0.984	--
2005-2009	--	0.918	0.961	0.989	--
2010-2014	--	0.928	0.957	0.983	--
Average	--	0.923	0.957	0.978	--

Notes: Estimated ratio of deaths in the SSA Numident data to adjusted National Death Index deaths over 1995-2014, by age group. Total (20+) column excludes children (age 0-19). The estimated ratio is calculated as D_{kt}/O_{kt} where D_{kt} represents the number of deaths reported in the SSA data for age group k occurring in year t and O_{kt} represents the official number of deaths of U.S. residents reported in the NDI for age group k during year t. Estimated ratio of non-beneficiary deaths that are reported (p) is calculated as in equation (40).

Table 1: Estimated Percentage of U.S. Deaths Included in the SSA Death Data and Underreporting Correction, by Age Group

ratio of SSA/NDI deaths is lower for those under 55. Thus, we focus our principal analyses on applicants age 55-64, where the quality of the SSA mortality data is excellent.

We provide further results on the quality of the SSA mortality data in appendix section B.2.

5.2 Correction for Underreporting in the SSA Mortality Data

While any underreporting of mortality for those denied benefits should be small, nonetheless, to account for possible underreporting, we calculate a correction factor, p, to offset any potential bias. We assume that SSA captures all deaths of allowed individuals but misses a fraction $(1-p)$ of non-beneficiaries' deaths – thus assuming that all of the SSA undercount comes from non-beneficiaries. This is a worst case bound – there are other reasons why SSA may count fewer deaths than NDI – but it gives a sense of how important underreporting among those denied could be for our results.[11] To see why this is likely a worst case bound, consider the fact that if both those allowed and those denied had the same under-reporting probabilities then the bias would only come from usual attenuation bias. In appendix section

[11] Although we made several adjustments to the data to make SSA mortality records comparable to the NDI, we cannot fully match the two. For example, illegal immigrants who lack an Social Security number should be captured in the NDI statistics if they die in the US. But SSA records deaths only for persons with Social Security numbers. Thus, the difference between NDI recorded deaths and SSA recorded deaths, likely overstates the number of missing deaths in the SSA data.

C.3 we show that p can be calculated as:

$$p = \frac{\text{\#of deaths in the SSA data} - \text{\#of deaths of beneficiaries in SSA data}}{\text{\#of deaths in the NDI data} - \text{\#of deaths of beneficiaries in SSA data}} \tag{9}$$

We calculate the average of p for each individual in our sample, using their age and year of application, over the sample period in which we observe them which we define as $\bar{p}_i$.[12] This approach allows us to reflect in our estimate of p the higher quality of the mortality data at older ages (when most deaths occur) and in more recent years.

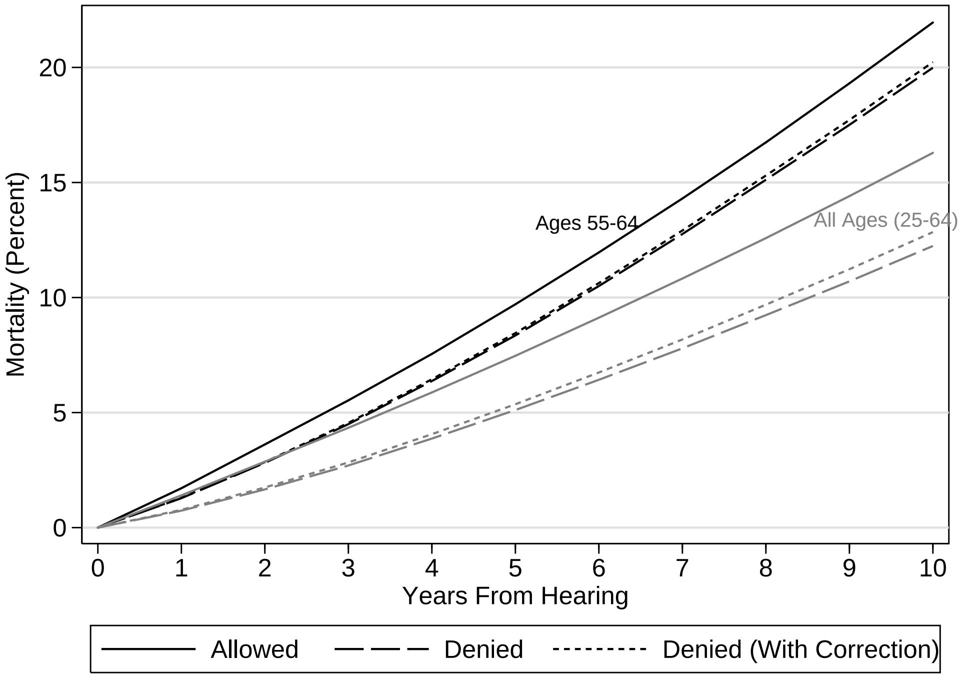
We use $\bar{p}_i$ to calculate a lower bound for the effect of receiving benefits on mortality by multiplying the observed mortality rate for persons denied benefits by $\frac{1}{\bar{p}_i}$.

5.3 Mortality Rates of Those Denied and Allowed

In this section we document some basic facts about mortality rates of those allowed versus denied. Figure 3 shows cumulative mortality rates conditional on assignment to an ALJ. For those aged 55-64 at time of application, the cumulative mortality rates in the year after assignment to an ALJ are 1.3% for those denied, versus 1.7% for those allowed, respectively. In the subsequent year the rates are 2.8% for those denied and 3.6% for those allowed. Over time, the mortality of those allowed rises faster than those denied, with a 10-year cumulative mortality rate of 22.0% for those allowed and 20.0% for those denied, a difference of 2.0%. For the full sample (aged 25-64), the 10-year cumulative mortality rate is 16.3% for those allowed and 12.2% for those denied, a difference of 4.1%. These differences should not be taken as causal, since those allowed may be less healthy. Our IV strategy seeks to address this issue. The mortality rates for those denied, with and without the correction for underreporting described in the previous section, can been seen in Figure 3. The underreporting correction has only a modest effect on our estimates: the estimated difference in 10 year cumulative mortality rates for applicants aged 55-64 between those allowed and denied falls from 2.0% to 1.8%.

Our estimated mortality rates are lower than Parsons (1991). He reports a six year mortality rate for all applicants of 12.9% for those denied and 17.5% for those allowed at ALJ stage. Our estimated six year mortality rates for all ages are 5.8% for those denied versus 8.6% for those allowed, and for the aged 55-64 are at 10.0% for those denied versus 11.1% for those allowed. Our estimates are likely lower because Parsons' cohort is from 1970 whereas ours is from 1995-2005. We also find a much smaller gap than Parsons between mortality for those allowed and denied. This could reflect more complete SSA capture of deaths of those denied benefits. More recent DI beneficiaries tend to be healthier than older ones and have primary diagnoses less related to mortality, as shown in Autor and Duggan (2006). Note that Parsons (1991) shows that mortality rates of those allowed at the initial stage is much higher than mortality rates of those who are allowed at subsequent stages of

[12] In practice we calculate p for each year for the following age groups: 25-44, 45-54 and 55-64. Using these values, we then calculate the two values $\bar{p}_5$ and $\bar{p}_{10}$ for each age and year of application combination using the mean values of the observations for the 5 or 10 years periods from the application year. We assume p is equal to 1 for those ages 65+, and therefore $\bar{p}_x$ is calculated as: $\bar{p}_{x,age,birthyear} = (\sum_{age=a}^{a+x} p_{g(a),a+birthyear})/x$ where $g(a)$ is the age group 25-44, 45-54, 55-64 or 65+, and $x \in \{5, 10\}$.



Notes: Cumulative mortality rates for applicants aged 55-64, and aged 25-64, at time of hearing with separate mortality rate curves for those allowed benefits, those denied benefits, and those denied benefits with correction for underreporting of mortality.

Figure 3: Cumulative Mortality rates, Allowed versus Denied

the adjudication process. Our sample is limited to those who are initially denied and thus get to an ALJ hearing, and our IV estimates of the effect of benefit receipt on mortality apply to those on the margin for being allowed at the ALJ stage. These estimates should not be extrapolated to applicants who receive benefits at the initial stage. Those who apply at the ALJ stage are healthier than those allowed at the initial stage (but should be less healthy than all persons denied at the initial stage, some of whom do not appeal to an ALJ). Nevertheless, we think that our sample is particularly interesting from a policy perspective, since these are the individuals whose allowance rates are likely to be affected by policy reforms that affect which persons receive benefits.

6 Results

6.1 Establishing the Validity of the Randomization

In previous sections we claimed that the assignment of cases to judges is random, conditional on hearing office and day. Random assignment implies that we should not be able to predict judge leniency using observable characteristics of the applicants who appear before that judge. Table 2 presents tests of this hypothesis for persons aged 55-64 when they apply. For similar tests on the full sample see Table A3 in the appendix.

First, we consider which variables predict allowance. Column 1 of Table 2 presents estimates from regressing an allowance indicator (de-meaned by hearing office and day) on the gender, age, race, labor force and earnings histories, legal representation, application type, education and health conditions of individuals in our estimation sample. Women, older individuals, whites, those with strong attachment to the labor market, high earners, those represented by a lawyer, and those who did not complete high school are more likely to be allowed benefits. Column 2 presents $t - statistics$ (all standard errors throughout are clustered by judge). Almost all of the covariates are highly statistically significant, due to the large sample size. The R^2 shows that the covariates explain 1.3% of the variation in allowance rates.

Our instrumental variable is judge leniency, $\tilde{Z}_i$. Column 3 presents estimates from a regression of judge leniency on the same covariates. Column 4 provides $t - statistics$.

Of the 20 covariates, only one has a coefficient that is statistically different than 0 at the 5% level, and not strongly so. For the full sample of those aged 25-64 we again only find one covariate that has a coefficient that is statistically different than 0 at the 5% level (see Table A3). All the estimated coefficients are small in comparison to the coefficients on the same variables in the allowance equation. The R^2 shows that the covariates explain 0.22% of the variation in judge specific allowance rates. These results could easily arise by chance, and are consistent with random assignment, which satisfies the independence assumption described in section 4. The next section provides some evidence on whether the rank and monotonicity conditions hold.

6.2 First Stage: The Effect of Judge Leniency on Allowance

Table 3 in the text and Table A4 in the appendix present estimates of the effect of judge leniency on allowance rates for the main estimation sample and the full sample, respectively. Column 1 shows the number of observations for different subsamples. Column 2 shows the allowance rate at the ALJ stage for that group. It shows, for example, that older individuals, high earners, and those represented by lawyers have relatively high allowance rates.[13] For health conditions, those with neoplasms (e.g., cancer), circulatory problems (e.g., heart disease), and musculoskeletal disorders (e.g., back pain) have high allowance rates, whereas those with mental disorders or retardation have lower allowance rates. Nevertheless, differences in allowance rates across subgroups are small.

Column 3 shows the estimated first stage regression coefficient $\hat{\lambda}$ from a regression of allowance on judge leniency using equation (4). The estimated value of $\hat{\lambda}$ for the main estimation sample is .68, meaning that the probability that case i is allowed at assignment rises .68 percentage points for every 1 percentage point increase in judge leniency (the de-meaned allowance rate for all other cases heard by case i's judge). Column 4 shows the standard error and column 5 the t-statistic: the estimate of $\hat{\lambda}$ is highly statistically significant for all subgroups. For the full sample in Appendix Table A4 the estimate of $\hat{\lambda}$ is .97. The difference in the two estimates arises because we measure judge leniency using the full sample. There is more dispersion in allowance rates in the full sample than for the 55-64 sample, so a

[13] The high allowance rate of cases represented by lawyers could be the result of lawyers representing only the most disabled claimants or lawyers causing the allowance probability to rise. We cannot distinguish between these two hypotheses.

Covariate	Dependent Variable: Allowed Coefficient (1)	t-stat (2)	Dependent Variable: Judge Leniency Coefficient (3)	t-stat (4)
Sex				
Female	0.0074	7.3	0.0007	1.9
Age				
55 to 59	-0.0089	-9.5	-0.0019	-2.2
Race				
Black	-0.0170	-10.2	-0.0016	-1.0
Other (non-black, non-white) or unknown	-0.0079	-4.2	-0.0013	-0.9
Labor force participation and income				
Average participation rate, years -11 to -2	0.0068	7.8	0.0006	1.0
Average earnings/billion, years -11 to -2 ($2006)	0.0004	8.9	0.0000	1.1
Represented by lawyer				
Represented by lawyer	0.0185	3.1	-0.0075	-1.8
Application type				
SSDI	-0.0134	-5.3	0.0010	0.5
Education				
High school graduate, no college	-0.0109	-10.8	-0.0012	-1.0
Some college	-0.0234	-14.9	-0.0019	-0.8
College graduate	-0.0269	-12.7	-0.0029	-1.2
Health conditions (by diagnosis group)				
Neoplasms (e.g., cancer)	0.0347	12.2	0.0031	1.2
Mental disorders	0.0019	0.9	0.0003	0.3
Mental retardation	0.0186	3.3	0.0001	0.1
Nervous system	0.0155	7.1	0.0011	1.0
Circulatory system (e.g., heart disease)	0.0325	17.5	0.0031	1.3
Musculoskeletal disorders (e.g., back pain)	0.0281	16.4	0.0031	1.6
Respiratory system	0.0194	8.8	0.0009	0.6
Injuries	0.0218	9.5	0.0016	0.9
Endocrine system (e.g., diabetes)	0.0281	12.8	0.0017	1.0
Standard deviation of dependent variable	0.2887		0.0955	
R^2	0.0127		0.0022	
Number of Applicants = 610,231	Number of Judges =1,436			

Notes: Column (1) is from a regression of de-meaned allowance on all the covariates listed. Column (3) is from a regression of judge leniency on all the covariates listed. Omitted category is male, 60-64s, white, not represented by a lawyer, applying for SSI or SSI and DI concurrently, not a high school graduate, with a health condition other than those listed above. The sample includes applicants aged 55 to 64, and we exclude applicants who died the year of application. Standard errors clustered by judge.

Table 2: Predictors of Allowance and Judge Leniency, Aged 55-64

	Obs.	Allowance Rate at ALJ Stage	Coefficient on Judge Leniency	Std. Error	T-Ratio	Relative Likelihood*
	(1)	(2)	(3)	(4)	(5)	(6)
All groups						
All groups	610,231	0.841	0.676	(0.008)	81	1.000
Sex						
Male	291,994	0.839	0.670	(0.010)	64	0.991
Female	318,237	0.843	0.682	(0.010)	71	1.009
Age						
55 to 59	390,600	0.836	0.686	(0.009)	77	1.015
60 to 64	219,631	0.850	0.657	(0.011)	60	0.972
Race						
White	415,125	0.853	0.653	(0.009)	72	0.966
Black	98,698	0.823	0.695	(0.016)	44	1.028
Other or unknown	96,408	0.806	0.747	(0.014)	55	1.104
Income						
Average earnings < \$10000	283,146	0.785	0.765	(0.012)	62	1.131
Average earnings ≥ \$10000	327,085	0.889	0.578	(0.010)	56	0.855
Represented by lawyer						
Represented by lawyer	385,118	0.854	0.652	(0.011)	59	0.964
Not represented by lawyer	225,113	0.820	0.727	(0.017)	44	1.076
Application type						
SSDI	352,991	0.856	0.647	(0.010)	66	0.956
SSI or Concurrent (both SSDI and SSI)	257,240	0.821	0.713	(0.011)	68	1.054
Education						
Less than high school	218,871	0.841	0.664	(0.011)	62	0.982
High school graduate, no college	267,634	0.847	0.668	(0.010)	69	0.988
Some college	77,685	0.830	0.706	(0.015)	46	1.044
College graduate	46,041	0.823	0.740	(0.018)	41	1.094
Health conditions (by diagnosis group)						
Neoplasms (e.g., cancer)	20,000	0.871	0.609	(0.025)	24	0.901
Mental disorders	61,508	0.795	0.817	(0.017)	47	1.209
Mental retardation	3,193	0.812	0.693	(0.056)	12	1.024
Nervous system	34,444	0.828	0.671	(0.022)	30	0.993
Circulatory system (e.g., heart disease)	103,725	0.861	0.637	(0.013)	50	0.942
Musculoskeletal disorders	231,391	0.856	0.648	(0.011)	62	0.959
Respiratory system	30,066	0.845	0.656	(0.020)	32	0.971
Injuries	27,091	0.840	0.689	(0.029)	24	1.019
Endocrine system (e.g., diabetes)	39,331	0.841	0.674	(0.018)	38	0.997
All other	59,482	0.793	0.719	(0.020)	35	1.063

Notes: Column (3) displays the first stage estimate of the coefficient λ from the regression of de-meaned allowance rates on judge leniency for those aged 55-64. Average earnings is calculated on income between 11 and 2 years before application. Standard errors clustered by judge. *Relative likelihood is the ratio of the group specific coefficient on judge leniency (presented in column 4) to the full sample coefficient.

Table 3: First Stage Estimates: Regression of Allowance Rates on Judge Leniency Variable, by Demographics, Aged 55-64

judge who is 1 percentage point more lenient on the full sample is only .68 percentage points more lenient for the 55-64 sample, who already have high allowance rates.

Column 3 shows that the estimated coefficient $\hat{\lambda}$ is larger for younger individuals, those with lower labor force participation and earnings prior to appealing, those not represented by a lawyer, and those whose primary health problem is a mental disorder. Abadie (2003) shows that the ratio of the group specific estimate of $\hat{\lambda}$ to the full sample estimate of $\hat{\lambda}$ is informative for understanding the characteristics of those allowed due to a small increase in the ALJ allowance rate. This ratio, shown in column 6, provides the relative likelihood that someone with a given characteristic is allowed given a small increase in judge leniency. Thus, an increase in the allowance threshold of all judges would increase the allowance rate of those with low participation and earnings, those not represented by a lawyer, and those with mental disorders more than for other groups, holding the applicant pool and the rest of the re-applications and appeals process constant. However, all relative likelihoods are close to 1, implying that more lenient judges are lenient across all applicants, to a similar extent.

The monotonicity assumption described in section 4 implies that the probability of allowance is non-decreasing in judge leniency for all subgroups of the population. Column 6 provides evidence supporting the monotonicity assumption. Furthermore, all estimates are highly significant, so the rank condition holds.

6.3 Second Stage: The Effect of Benefit Receipt on Mortality

Panel (a) of Table 4 presents estimates of the effect of disability recipiency on mortality 5 and 10 years after assignment to an ALJ for our main estimation sample. For example, the first two rows show that 21.95% of those allowed benefits in our sample die within 10 years, whereas 19.99% of those denied benefits die within 10 years. This difference of 1.97% is shown in the third row. These estimates suggest that those allowed benefits are more likely to die. An equivalent, way of obtaining this difference is to take the coefficient on allowance from a regression of mortality on allowance. This approach produces a standard error, reported in the fourth row. The difference in mortality between those allowed and denied is statistically significant. However, these are simple OLS estimates, without covariates, which do not address selection effects and do not provide causal estimates.

Panel (d) of Table 4 presents the same estimates as Panel (a), but for the full sample (ages 25-64). Panels (b) and (c) display the estimates for the populations aged 45-54 and 25-44, respectively. Perhaps surprisingly, the full sample (Panel (d)) coefficient on allowance is larger than the coefficients on allowance for each of the subsamples (Panels (a)-(c)). The reason for this is that the coefficient on allowance is the raw difference in mortality between those allowed and denied. Older individuals have higher mortality and are more likely to be allowed. Thus allowed individuals are older, higher mortality individuals, and denied individuals are younger, lower mortality individuals. By separating the full sample into age subgroups, we condition away some of the age-related differences in mortality rates between those allowed and those denied.

The next rows show OLS and IV estimates of de-meaned (by hearing office and day) mortality on similarly de-meaned allowance and the associated standard error. De-meaning the data has very little effect on the OLS estimates. In Panel (b), the IV estimates show that being allowed benefits increases the 5 year and 10 year mortality rate by 1.81 and

	Panel (a): Aged 55-64 Mortality (Percent)				Panel (b): Aged 45-54 Mortality (Percent)			
	5 years		10 years		5 years		10 years	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Without Covariates:								
Allowed	9.71		21.95		8.02		17.41	
Denied	8.35		19.99		6.14		14.78	
Coef on allowance	1.35		1.97		1.88		2.64	
(Std. Error)	(0.11)		(0.19)		(0.08)		(0.13)	
Coef on demeaned allowance*	1.35	1.81	1.87	1.93	1.87	1.47	2.71	2.59
(Std. Error)	(0.12)	(0.44)	(0.19)	(0.76)	(0.08)	(0.63)	(0.13)	(0.99)
With Covariates:								
Coef on demeaned allowance*	1.94	2.30	2.77	2.81	2.29	1.49	3.45	2.60
(Std. Error)	(0.12)	(0.50)	(0.18)	(0.91)	(0.08)	(0.62)	(0.12)	(0.94)
With Covariates and Underreporting Correction:								
Coef on demeaned allowance*	1.76	2.12	2.51	2.54	2.05	1.26	3.02	2.17
(Std. Error)	(0.12)	(0.50)	(0.18)	(0.90)	(0.08)	(0.62)	(0.12)	(0.94)

	Panel (c): Aged 25-44 Mortality (Percent)				Panel (d): All Ages (25-64) Mortality (Percent)			
	5 years		10 years		5 years		10 years	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Without Covariates:								
Allowed	5.23		10.91		7.47		16.29	
Denied	3.56		8.47		5.11		12.24	
Coef on allowance	1.67		2.44		2.36		4.05	
(Std. Error)	(0.06)		(0.11)		(0.10)		(0.20)	
Coef on demeaned allowance*	1.63	1.18	2.46	2.26	2.27	2.17	3.93	4.30
(Std. Error)	(0.06)	(0.42)	(0.09)	(0.70)	(0.07)	(0.80)	(0.14)	(1.64)
With Covariates:								
Coef on demeaned allowance*	1.78	1.10	2.76	2.16	2.21	1.64	3.55	2.96
(Std. Error)	(0.05)	(0.38)	(0.09)	(0.64)	(0.05)	(0.58)	(0.10)	(1.05)
With Covariates and Underreporting Correction:								
Coef on demeaned allowance*	1.80	1.02	2.70	1.94	1.97	1.39	3.09	2.49
(Std. Error)	(0.05)	(0.51)	(0.09)	(0.81)	(0.06)	(0.58)	(0.10)	(1.05)

Notes: N= 610,231 in Panel (a), N= 1,048,344 in Panel (b), N=1,101,332 in Panel (c), and N= 2,759,907 in Panel (d). Instrument is judge leniency. Covariates are those in Table 2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. Standard errors clustered by judge. *For de-meaned allowance, all variables are de-meaned from the hearing office-day average.

Table 4: Estimated Effect of DI Recipiency on Mortality

1.93 percentage points, respectively. Surprisingly, the IV estimates are close to the OLS estimates.

What can we learn from the similarity of the OLS and IV estimates? Less than one might think. The average allowed applicant is likely in worse health than the average denied applicant; thus, higher overall mortality for those allowed benefits in OLS, without covariates, is expected. The OLS estimate also assumes homogeneous treatment effects across all applicants, regardless of health. This seems unlikely. IV, in contrast, estimates the average effect of allowance for the subsample of applicants who are on the margin for being allowed or denied, and hence affected by the judge leniency instrument. Given the 84% average allowance rate, this subsample is likely healthier than the average for all applicants. The IV estimate is based on random assignment, so the marginal allowed and denied applicants should be in similar health. IV provides a credible estimate of the effect of allowance on mortality, but only for those on the margin to be allowed or denied.

The next rows provide OLS and IV estimates which include the covariates listed in Table 2. Adding covariates to this specification has only a small effect on the IV estimates. Recall that our IV estimation procedure should deliver consistent estimates, with or without covariates. Thus, it is reassuring to see that adding covariates has only a small effect on the IV estimates. The IV estimates are strongly statistically significant at both 5 and 10 years.

More surprisingly, adding covariates increases the estimated OLS effect of benefit receipt on mortality. On closer look, adding some covariates increases the estimated effect of benefit receipt on mortality, whereas adding others decreases this estimate. Some groups with higher mortality rates (shown in Table 5) also have high allowance rates (shown in Table 3). For example, those with cancer, and older (age 60+) individuals have both higher mortality rates and higher allowance rates. Conditioning on these variables moves the OLS estimates closer to 0. However, other groups with higher mortality, such as blacks and those with low prior earnings, have lower allowance rates. Conditioning on these variables produces larger OLS estimates.

The OLS estimates with covariates would have a causal interpretation only under two strong assumptions: that unobservables do not predict both allowance and mortality (no omitted variable bias); and treatment effects are homogeneous. Since accounting for selection on observables somewhat increases the estimated mortality effect, it is plausible that inability to account for unobservables does not necessarily lead to upward biased estimates. However, below, we find evidence of heterogeneous treatment effects.

The final rows in each panel of Table 4 display the estimates with covariates, after including the underreporting correction described in section 5.3. As expected, the estimates fall slightly but remain broadly the same.

For the full sample, in Panel D, the IV estimates with covariates show that being allowed benefits increases the 5 year and 10 year mortality rate by 1.64 and 2.96 percentage points, respectively. For the full sample, the IV estimates with covariates are somewhat smaller than the OLS estimates, but this comparison should be made cautiously, because these estimates apply to different populations.

6.4 Heterogeneity in the Mortality Effect Based on Judge Leniency: Marginal Treatment Effects

Using the Marginal Treatment Effects approach described in Section 4.3 and the Appendix C.1, this section shows how the predicted effect of DI benefit allowance varies with predicted de-meaned allowance.

Figure 4 presents how the MTE (i.e., the mortality response for the marginal case allowed) varies with predicted de-meaned allowance for different age groups. In each age panel the left figure displays 5 year mortality, and the right figure displays 10 year mortality, controlling for covariates.

We use third order polynomials for both the instrument and the endogenous variable (de-meaned allowance) when estimating equations (6) and (7). The cubic specification is flexible, although visual inspection of Figure 4, as well as both the Akaike and Bayesian information criterion show that there is little gain from going beyond the quadratic specification. In Appendix B.3 we show that these results change only modestly when excluding covariates or when using a local polynomial smoother following Maestas et al. (2013).

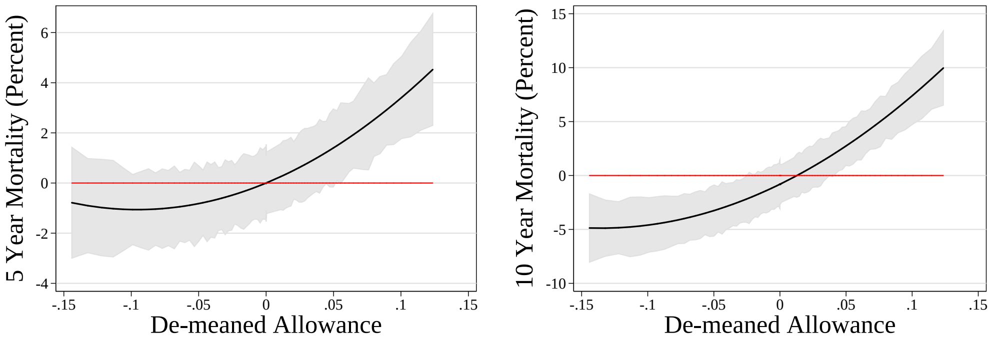
Since polynomial smoothers have poor endpoint properties, we show estimated MTEs over the middle 90% of the distribution of de-meaned allowance rates. In Monte Carlo experiments, we found our procedure produced little bias over this range. Figure 4 also shows bootstrapped 95% confidence intervals.

Consider first panel (a), which shows estimates for our main estimation sample, ages 55-64. The estimated MTE is close to zero at the average predicted allowance rate, at both 5 and 10 years, but there is strong heterogeneity in the responses. Being allowed benefits reduces 5 year mortality by an estimated 0.8 percentage points for the marginal applicant heard by an ALJ who is stricter than 95% of all judges. These judges have allowance rates that are twelve percentage points below the average (de-meaned by hearing office and day). However, allowance increases mortality by 4.5 percentage points for the marginal applicant heard by an ALJ who is more lenient than 95% of all judges. These judges have allowance rates that are ten percentage points above the average. The 10 year mortality response of those 55-64 is qualitatively similar to the 5 year response. The magnitudes are larger, which is unsurprising, given that the impacts of allowance have more years to accumulate.

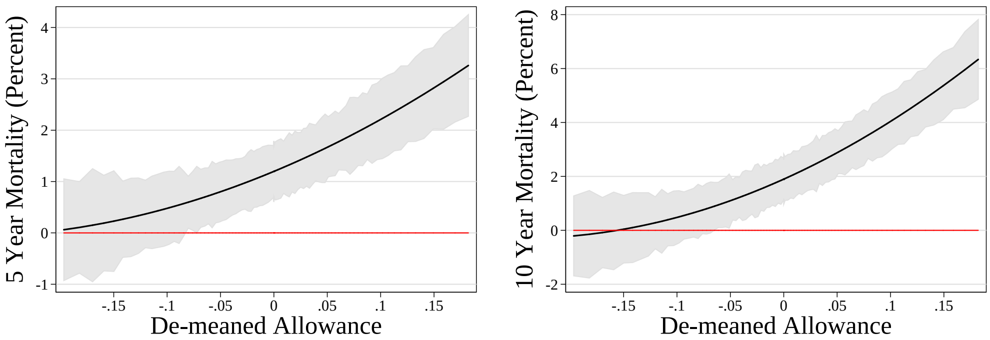
Estimates for the other age groups follow the same basic patterns as for those aged 55-64: greater leniency implies higher recipient mortality for the marginal applicant.

In summary, our results suggest that making the DI screening threshold significantly less strict (and thus increasing the allowance rate) will increase mortality of the marginal applicants, at least for those assigned to more lenient judges. Interestingly however, for the 55-64 year olds, our 5 year mortality estimates suggest that increasing the screening threshold for the strictest judges would not increase mortality. This provides some evidence, at least for the 55-64 year olds, that current screening thresholds are about right. Our evidence also suggests that the screening threshold could be made stricter for younger age groups without worsening – and likely increasing – their longevity.

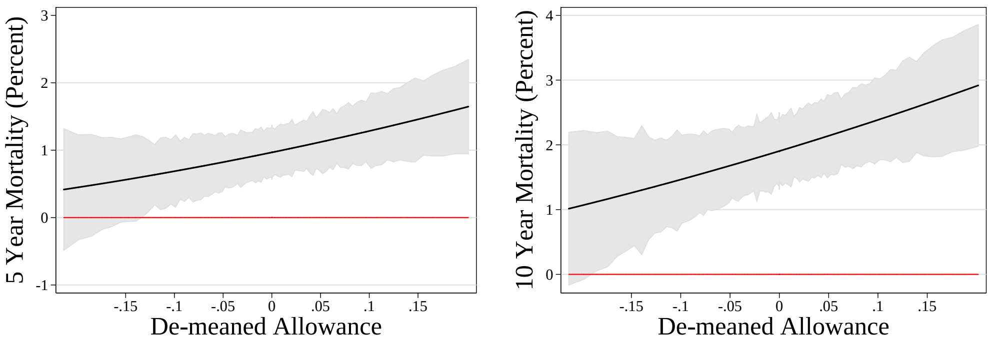
Our results are consistent with the notion that as allowance rates rise, more healthy individuals are allowed DI. Healthier individuals benefit less from Medicare and Medicaid insurance from DI allowance. These individuals also have a bigger decline in labor supply in response to DI receipt (see French and Song, 2014). This is due to more of them being



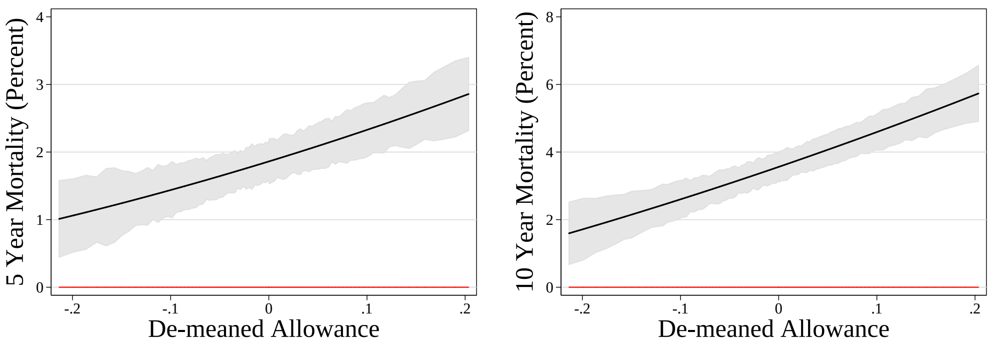
(a) Aged 55-64



(b) Aged 45-54



(c) Aged 25-44



(d) All Ages (25-64)

Notes: This figure displays the estimated mortality response as a function of predicted de-meaned allowance. We control for the covariates listed in Table 2. Within each panel: the left figure displays 5 year mortality, and the right figure displays 10 year mortality. Mean allowance rate is 0.84 for those aged 55-64, 0.71 for those aged 45-54, 0.63 for those aged 25-44, and 0.71 for those aged 25-64.

Figure 4: Marginal Treatment Effects: Mortality Response by De-meaned Allowance

able to work in the absence of DI receipt. Since working potentially has beneficial effects on mortality, any adverse effect of not working will be stronger; younger recipients will also lose more working years. Thus, any beneficial effect of DI allowance is smaller, and the adverse effect larger, for healthier (and younger) individuals.

Given that the average allowance rate is 71% for those aged 25-64 (and is 84% for those aged 55-64), and this excludes the sickest individuals who are allowed at the initial level, the average recipient is substantially less healthy than the marginal applicant. For this reason we would expect the average recipient to be positioned well off to the left of the MTE graphs. If the MTE curve continues to slope down and to the left – which is plausible, but unprovable – this suggests that receiving disability benefits reduces mortality, perhaps strongly so, for the average applicant. This is true even though DI receipt increases mortality, on average, for the applicants who are affected by our judge leniency instrument.

6.5 Heterogeneity in the Mortality Effect Based on Observables

Table 5 disaggregates the 5 and 10 year mortality response by demographics, prior earnings, and health conditions. The left panel shows 5 year mortality estimates and the right panel shows the 10 year mortality estimates, for applicants aged 55-64. Each panel reports the unadjusted mean mortality for allowed and denied individuals, the OLS estimate of allowance on mortality with covariates, the IV estimate of allowance on mortality with covariates, and the standard error. Table 5 shows that the effect of DI allowance on 10 year mortality does not vary in a dramatic way across subgroups. Other than the subgroups for specific health conditions (bottom rows), all subgroup IV estimates are positive, most are statistically significant, and the 95% confidence interval for related subgroups generally overlap. The principal difference across subgroups is that the higher mortality for whites is smaller at both 5 and 10 years than for other racial groups.

		Panel (a): 5 Year Mortality (Percent)						Panel (b): 10 Year Mortality (Percent)					
		Mortality Rates		OLS		IV		Mortality Rates		OLS		IV	
	Obs	Allowed	Denied	Diff.	SE	Diff.	SE	Allowed	Denied	Diff.	SE	Diff.	SE
All groups	610,231	9.71	8.35	1.94	(0.12)	2.30	(0.50)	21.95	19.99	2.77	(0.18)	2.81	(0.91)
Sex													
Male	291,994	12.42	10.96	2.20	(0.19)	2.73	(0.98)	27.14	25.65	2.59	(0.27)	3.46	(1.50)
Female	318,237	7.23	5.91	1.67	(0.15)	1.88	(0.58)	17.22	14.67	2.89	(0.22)	2.16	(0.83)
Race													
White	415,125	9.64	8.78	1.67	(0.15)	1.60	(0.64)	22.07	20.72	2.59	(0.22)	2.32	(1.28)
Black	98,698	11.09	9.33	2.70	(0.28)	3.39	(1.50)	23.83	21.83	3.45	(0.40)	4.17	(2.09)
Other	96,408	8.55	6.04	2.29	(0.24)	4.05	(0.93)	19.48	15.92	3.17	(0.37)	3.90	(1.33)
Education Group													
Less than high school	218,871	9.76	8.40	1.73	(0.20)	2.63	(0.72)	22.51	20.44	2.30	(0.29)	2.12	(1.04)
High school graduate	267,634	9.64	8.24	2.18	(0.18)	2.54	(1.00)	21.73	19.86	3.03	(0.27)	3.45	(1.70)
Some college	77,685	9.78	8.45	1.86	(0.32)	0.33	(1.29)	21.89	20.00	3.34	(0.46)	2.85	(2.08)
College graduate	46,041	9.72	8.50	1.72	(0.40)	2.70	(1.87)	20.71	18.69	2.75	(0.55)	2.84	(2.24)
Income													
Average earnings $<$ \$10000	283,146	10.78	9.29	1.80	(0.17)	2.64	(0.62)	23.90	21.79	2.28	(0.24)	2.26	(0.98)
Average earnings $\geq$ \$10000	327,085	8.89	6.77	2.23	(0.17)	1.99	(0.79)	20.47	16.95	3.61	(0.25)	3.74	(1.43)
Health conditions													
Neoplasms	20,000	25.80	28.26	-1.55	(1.02)	0.33	(8.15)	40.93	43.22	-0.43	(1.21)	-1.59	(9.66)
Respiratory system	30,066	14.53	12.09	2.84	(0.57)	-3.72	(2.70)	32.71	30.10	3.75	(0.82)	0.49	(3.55)
Endocrine system	39,331	14.19	10.51	3.38	(0.49)	0.56	(2.99)	32.11	25.42	5.97	(0.68)	1.89	(5.75)
Circulatory system	103,725	11.95	9.53	2.38	(0.30)	4.69	(1.28)	27.73	23.76	2.99	(0.43)	3.96	(2.12)
Mental retardation	3,193	11.00	9.98	4.49	(1.72)	4.76	(7.26)	22.26	21.63	3.89	(2.37)	9.59	(12.13)
Nervous system	34,444	9.41	8.60	2.38	(0.47)	2.04	(2.16)	21.88	21.20	2.69	(0.66)	0.71	(2.75)
Mental disorders	61,508	8.33	7.26	1.81	(0.32)	3.56	(1.27)	19.20	17.88	2.58	(0.45)	4.22	(1.91)
Injuries	27,091	7.62	6.22	1.51	(0.16)	1.82	(1.11)	17.95	15.92	2.32	(0.25)	2.59	(1.26)
Musculoskeletal disorders	231,391	5.94	5.01	2.48	(0.48)	2.29	(1.87)	15.02	13.72	3.67	(0.73)	4.11	(3.26)
All other	59,482	12.11	11.00	2.26	(0.39)	3.08	(2.16)	25.03	24.00	2.84	(0.54)	2.95	(2.34)

Notes: This table displays the estimated effect of DI recipiency on mortality (with covariates) for those aged 55-64. Mortality Rates do not adjust for covariates. OLS and IV estimates are demeaned and include covariates. Covariates are those in Table 2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. IV estimates use demeaned variables and judge leniency as the instrument. Standard errors clustered by judge.

Table 5: Estimated Effect of DI Recipiency on Mortality, Disaggregated, Aged 55-64

The subgroups based on health condition listed in the disability application are listed in order of decreasing 5 year mortality rates. Sample sizes are generally much smaller and standard errors are much larger, but there are some suggestive differences. Individuals diagnosed with neoplasms (e.g. cancer) have the highest overall mortality rates, and have *higher* mortality rates when denied, in both the OLS estimates and the 10 year IV estimates (the 5 year IV estimate is close to zero). This is potentially evidence that DI, and the associated health care benefits, are more valuable to those with cancer than other disabilities. Perhaps health insurance is of special value to this group, given both the high cost of treating cancer, and the high mortality of those with cancer. Note too that the second highest mortality group, with respiratory disease, has a negative IV estimate at 5 years, and a near-zero estimate at 10 years. We investigate these hints of differential effects based on health condition, and the cost of treating that condition, in the next section.

7 The Channels by which DI Affects Mortality

We do not find any adverse effect of being denied benefits on mortality for the marginal applicant. Yet, as we show below, cash income and health insurance transfers to the disabled are large, which would suggest lower mortality, other factors equal. This leaves the effect of receiving benefits on labor supply as a potential offsetting effect. As noted previously in Section 2.3, many studies have shown that DI receipt reduces employment, and other studies suggest that employment reductions can increase mortality. In this section we discuss some channels by which allowance could impact mortality.

We summarize the quantitative magnitude of these channels in Table 6. In this table we display several outcomes for individuals denied by an ALJ and calculate the difference in these outcomes between those allowed versus denied.

	(1) Outcomes for denied			(2) Difference in outcomes between those allowed versus denied at ALJ stage						
	Years after ALJ stage			1 year later		3 years later		5 years later		Total discounted
	1	3	5	OLS	IV	OLS	IV	OLS	IV	benefits up to age 65*
A. Prob of being allowed at future times	0.24	0.42	0.54	0.76	0.71	0.58	0.51	0.46	0.40	
B. Prob of earnings > 0	0.22	0.20	0.16	-0.12	-0.12	-0.11	-0.10	-0.08	-0.07	
C. Prob of earnings > SGA	0.08	0.08	0.06	-0.06	-0.06	-0.06	-0.06	-0.05	-0.04	
D. Cash income	4,265	5,733	6,471	5,969		4,204		2,958		24,039
D(i). Cash benefits	2,313	4,139	5,433	7,866		6,046		4,804		33,799
D(ii). Average earnings, before taxes	2,402	1,950	1,314	-2,182	-2,365	-2,111	-2,256	-1,943	-1,586	-11,103
D(iii). Average earnings, net of tax	1,953	1,594	1,037	-1,897		-1,843		-1,846		-9,760
E. Prob of receiving Medicare/Medicaid	0.12	0.40	0.50	0.16		0.48		0.39		
F. Annual Medicare/Medicaid payments	1,457	4,809	6,037	1,926		5,784		4,736		23,038
G. Total dollar value	5,723	10,541	12,507	7,665		8,793		6,182		47,077

Notes: Panel (1) displays the predicted average outcomes for those denied at the ALJ stage. Panel (2) displays the difference in outcomes between those allowed versus denied at ALJ stage. OLS and IV esimtates in panel (2) control for covariates. The average predicted outcomes for those allowed at the ALJ stage is the sum of the relevants cells in panels (1) and (2). The calculations assume individuals were under age 60 at assignment. All dollar amounts in 2014 dollars.
Row A: probability of being allowed at future times. Source: French and Song (2014).
Row B: probability of having positive earnings. Source: Our data.
Row C: probability of earning above the Substantial Gainful Activity Level ($12,480 per year in 2014). Source: Our data.
Row D: predicted cash benefits plus after tax income. Source: Our data.
Row D(i): predicted cash benefits received after deducting the average reduction in benefits due to work. Source: French and Song (2014).
Row D(ii): predicted average earnings before tax. Source: Our data.
Row D(iii): predicted average earnings after tax. Source: Source: Our data and French and Song (2014).
Row E: probability of receiving Medicare and/or Medicaid. Source: Rupp and Riley (2012) and the appendix.
Row F: average annual medical payments from Medicare and/or Medicaid. Source: Appendix of De Nardi et al., 2016.
Row G: total dollar value difference of predicted cash income, benefits, taxes, and medical payments from Medicare and/or Medicaid.
* The total discounted values assume that an individual is first seen by an ALJ at age 58, which is the median age at assignment in our sample. Benefits are cumulated through age 65 (7 years later), and discounted using an interest rate of 3% and the observed mortality rate for those allowed by an ALJ in our sample.

Table 6: Key Outcome Differences Between Those Allowed versus Denied

7.1 Allowance

We estimate the effect of ALJ allowance on mortality. However, many individuals who are initially denied are eventually allowed upon reapplication or appeal. In this sense we have an "intent to treat" estimate, rather than a "treatment effect on the treated" estimate. We estimate the impact of initial allowance by an ALJ, rather than final allowance, because final allowance depends on mortality: only still-living persons can receive benefits after appeal. However, appeals and re-applications are important for understanding the magnitude of the effect of allowance on benefits received.

Panel (1) of Table 6 shows outcomes for those denied by an ALJ 1, 3, and 5 years after assignment to an ALJ. Row A shows that 54% of those denied by an ALJ are allowed within 5 years.

Panel (2) displays the difference in outcomes between those allowed versus denied at the ALJ stage. For example, virtually 100% of those allowed benefits are still receiving benefits 5 years later, whereas 54% of those denied by an ALJ are allowed 5 years later, therefore the difference is 100-54=46%. This can be seen in the 5 year OLS estimate of row A.

The results of Table 6 take into account that many persons who are denied benefits at the ALJ stage are later allowed. We present calculation details in appendix section D, and provide more information on the data sources behind our estimates.

7.2 The Income Benefit and Labor Supply Incentives

One potentially important determinant of mortality is income. There are many possible channels through which income can affect health, including through investment in health through better food, shelter, and health care. In this section we discuss how income responds to benefit allowance. Specifically, we focus on the response of taxable earnings and DI/SSI benefits to benefit allowance.

Both income effects (through the high replacement rate) and substitution effects (beneficiaries will lose benefits if they earn above the SGA amount) causes DI recipients to reduce labor supply. DI/SSI benefits likely also reduce labor supply through a third channel – health insurance, which greatly reduces the value of employer-provided health insurance, which can be an important work incentive (French and Jones, 2011).

Row B of Table 6 presents estimates of the employment response to being allowed disability benefits. Panel (1) shows that 16% of all individuals denied by an ALJ have positive earnings 5 years after assignment to an ALJ. The OLS estimates in Panel (2) show that being allowed benefits by an ALJ reduces employment rates by 8 percentage points after 5 years, with similar IV estimates. The OLS estimates in row C show that being allowed by an ALJ reduces the probability that earnings exceed the SGA limit (of $12,480 in 2014) by 5 percentage points after 5 years; IV estimates are again similar. These reductions in employment lead to significant declines in earnings: pre-tax earnings fall when allowed by $1,943 after 5 years (see row D(ii)), although the post-tax earnings loss is somewhat smaller (see row D(iii)).

Total cash income rises after allowance, since the cash value of DI/SSI benefits exceed the decline in income. The average extra value of these benefits for those allowed at the ALJ stage averages $5,969 1 year after being allowed by an ALJ, but falls to $2,958 5 years after. This fall occurs because many of those initially denied are later allowed upon appeal or re-application or because they are old enough to receive Social Security benefits.

We should note that we cannot assess all channels by which DI/SSI receipt may affect household income. For example, Autor et al. (2015) show that show that in Norway disability benefit receipt also leads to reductions in spouse's earnings and other benefits (such as unemployment insurance).

7.3 Health Insurance Benefits

Individuals receiving DI benefits are eligible for Medicare after a two year waiting period. Individuals drawing SSI are often also immediately eligible for Medicaid, the government health insurance program for the poor. Livermore et al. (2011) show that federal and state governments spend more on health care than on cash benefits for the disabled.

Rupp and Riley (2012) report the percentage of DI beneficiaries receiving either Medicare or Medicaid over a period covering 12 months before they were awarded DI until 6 years after. They show that immediately following DI/SSI benefit receipt, 24.7% receive either Medicaid or Medicare, the majority being SSI beneficiaries who receive Medicaid. The total jumps to 89.7% just after 2 years when DI beneficiaries become eligible for Medicare, and reaches 96.8% after 6 years.

Using the values from Rupp and Riley (2012) and the calculations explained in Appendix D we calculate the difference in the probability of receiving Medicare or Medicaid between those allowed versus denied at ALJ stage, taking into account that many of those denied by an ALJ are later allowed. These results are shown in row E of Table 6. The higher probability of receiving Medicare and/or Medicaid is fairly small 1 year later at only 16 percentage points, peaks at 3 years later when almost everyone allowed by the ALJ is receiving Medicare, and then decline as many of the initially denied are later allowed.

Using data from the appendix of (De Nardi et al., 2016) we calculate that the average Medicare and/or Medicaid recipient receives $12,012 worth of medical transfers from Medicare/Medicaid per year. Row F of Table 6 calculates the difference in the average annual medical payments by multiplying $12,012 by the difference in probability of receiving Medicare/Medicaid (row E). This means that 1 year later those allowed are receiving on average $7,665 more in medical transfers. After 5 years this difference is $6,182 per year.

7.4 Total Discounted Value of Income and Benefits

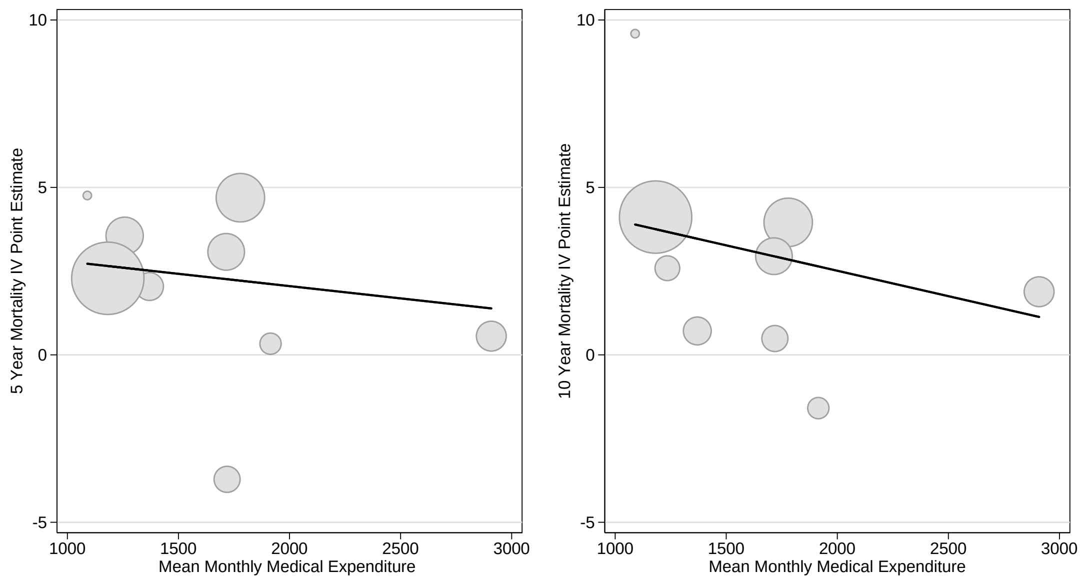
The final column in Table 6 shows the present discounted value of all income and benefits that arise from being allowed DI by an ALJ up to age 65, when everyone should become eligible for Medicare and Social Security benefits

To calculate this we assume that everyone in the age 55-64 group is age 58, which is the median age for this group in our sample. We discount future benefits and income using an interest rate of 3%, taking into account that not everyone lives to age 65, using the mortality rates for those allowed by an ALJ in our sample. We estimate that the average total discounted value of income and benefits of being awarded DI by an ALJ is $47,077. Of this, 51% is in cash income and 49% in medical transfers. These are substantial amounts which, other factors equal, would be expected to reduce mortality.

7.5 Effects Disaggregated by Health Condition

In Table 5 we find some evidence that mortality responses vary for different reported applicant health conditions. In Figure 5, we investigate further the hints from that table that the effect of benefit allowance on mortality is more favorable (less adverse) for more expensive health conditions and for conditions that predict higher near-term mortality.

Figure 5 plots the 5 and 10 year mortality point estimates by health condition from Table 5 against mean monthly health care spending for that condition (in thousands of 2014 US dollars) by health condition from the Medicare Current Beneficiary Survey (MCBS).[14] We calculate mean medical spending for disabled Medicare beneficiaries under age 65.[15] The size of the circles represents the number of observations.

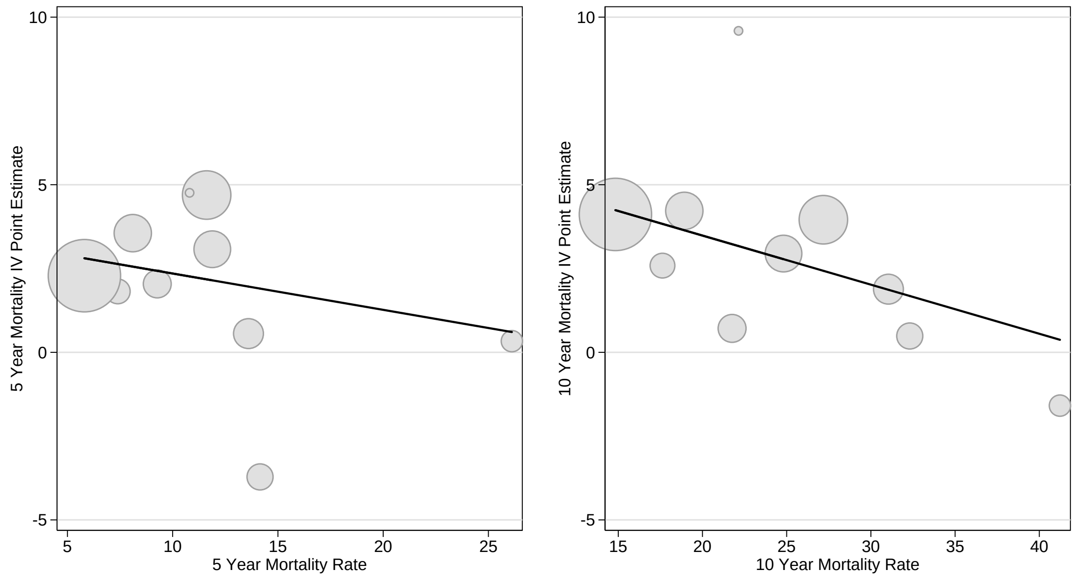


Notes: This figure displays a scatter plot of the 5 year (left graph) and 10 year (right graph) mortality IV point estimates by health condition from Table 5 plotted against mean monthly spending for that condition (in thousands of 2014 US dollars), from the MCBS. Circle size is proportional to number of disability applicants with that condition. The line represents predicted mortality from a regression of the health condition specific mortality point estimates against mean monthly spending, weighted by the number of individuals in each condition group from the SSA.

Figure 5: Estimated 5 year and 10 year mortality effect of allowance by mean monthly medical expenditure for each health condition

[14] We use estimates from the appendix in De Nardi et al. (2016).

[15] More precisely, we use those receiving Medicare benefits who are younger than 65. Virtually everyone under age 65 who receives Medicare also receives disability benefits. The MCBS has high quality medical spending data since it uses administrative Medicare records for Medicare spending and a mixture of survey data and reconciliation of survey, Medicaid participation, and Medicare records to infer payments by other payors. De Nardi et al. (2016) find that the MCBS captures approximately 80% of total medical spending for its target population and French et al. (2017) find that out of pocket spending and private insurance information match up well between MCBS and the Health and Retirement Study. An attractive aspect of the MCBS data is that respondents are asked about the main health condition that caused them to be eligible for Medicare benefits. Thus we can match the condition that led to allowance in both the Social Security data and the MCBS data.



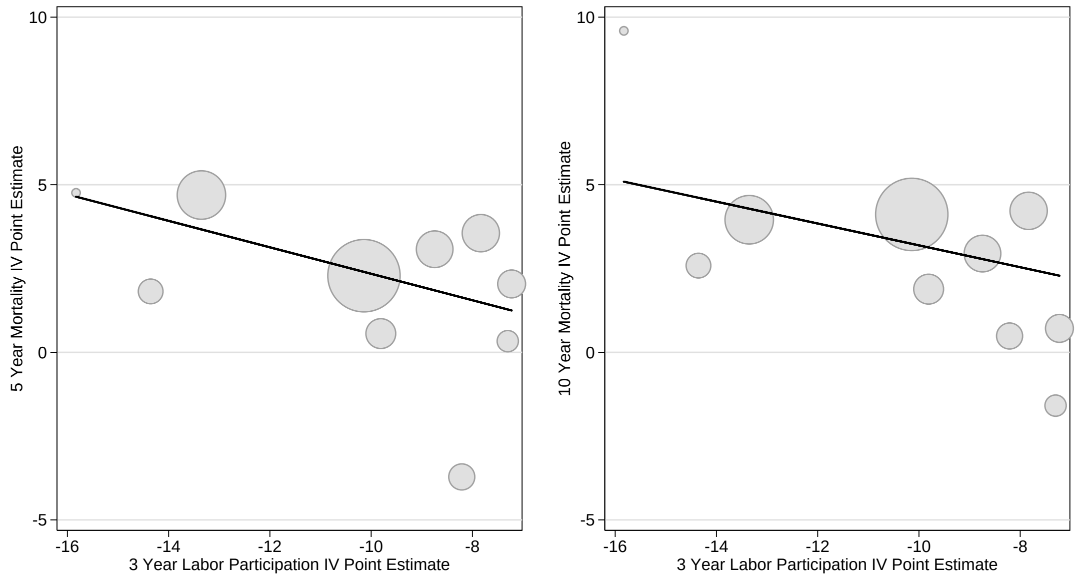
Notes: This figure displays a scatter plot of the 5-year and 10-year mortality IV point estimates by health condition from Table 5 plotted against the unconditional mortality rate of the individuals with each condition from SSA mortality records.

Figure 6: Estimated 5 year and 10 year mortality effect of allowance by the mortality rate for each health condition

Over both 5 year and 10 year periods, we find a general tendency, albeit with substantial scatter, for benefit allowance to be less adverse to mortality (averaged over the range of judge leniency we observe) for higher-cost medical conditions. This is consistent with the view that access to health insurance, and thus potential access to better healthcare, reduces mortality for those with more expensive conditions, and can offset any adverse effect of work disincentives.

Figure 6 plots the 5 and 10 year mortality point estimates by health condition from Table 5 against the average 5 and 10 year mortality rate for that condition. Over 5 years, we find either no increase or a predicted decline in mortality, among those with neoplasms (e.g., cancer), respiratory conditions and problems with the endocrine system, which are the highest mortality rate conditions. These conditions are also amongst the most expensive in terms of medical treatment. Conversely, conditions with relatively low mortality, which also tend to have lower medical spending, such as mental retardation, mental disorders, and musculoskeletal disorders, have increased average marginal mortality following benefit receipt. Similar to Figure 5, Figure 6 is consistent with the view that improved access to health care can reduce mortality for expensive, high mortality conditions. The negative slopes in Figures 5 and 6 are not statistically significant, however, and should only be taken as suggestive evidence.

Any effects due to improved access to health care for high cost conditions are likely offset by increases in mortality associated with reduced employment. This can be seen in the relationship plotted in Figure 7, which shows that the health conditions where receipt of benefits increases mortality by the most are also the conditions where the receipt of benefits



Notes: This figure displays a scatter plot of the 5 year mortality IV point estimates by health condition from Table 5 plotted against the 3 year labor supply IV point estimates for each condition group from the SSA.

Figure 7: Estimated 5 year and 10 year mortality effect of allowance by the labor supply response for each health condition

decreases labor supply the most. This is consistent with the view that the work disincentive from receiving benefits could increase mortality.

8 Robustness

Our results for the main estimation sample (those aged 55-64 at time of application) are robust to a number of other modifications to sample selection and functional form. Table 7 provides robustness checks. The left panel shows 5 year mortality estimates and the right panel shows the 10 year mortality estimates. In each panel, odd-numbered columns reports the estimates with no covariates; even columns present estimates with covariates.

The first two rows display OLS estimates. In the second row, we include the 10,006 individuals who died within a year of seeing a judge. As discussed in Section 5, we exclude these individuals as we are concerned that some of our sample who are denied are likely just those who die before being heard by a judge. Including these individuals decreases the coefficients by a small amount without covariates, but increases the coefficients slightly with covariates; thus, our choice to exclude these individuals does not meaningfully affect our overall findings.

The remaining rows provide IV estimates in various different specifications or with different sample selections. Including persons who died within a year of seeing a judge has only a small effect on our IV estimates. The next two rows change the number of judges we exclude due to them seeing a limited number of cases. Whereas in our baseline specification

	Panel A: 5 Year Mortality (Percent)		Panel B: 10 Year Mortality (Percent)	
	No Covariates	With Covariates	No Covariates	With Covariates
OLS				
Baseline	1.35	1.94	1.87	2.77
	(0.12)	(0.12)	(0.19)	(0.18)
Inc. those who die year of app.	1.15	1.99	1.69	2.82
	(0.14)	(0.13)	(0.19)	(0.18)
IV				
Baseline	1.81	2.30	1.93	2.81
	(0.44)	(0.50)	(0.76)	(0.91)
Inc. those who die year of app.	1.73	2.30	1.82	2.81
	(0.44)	(0.50)	(0.75)	(0.90)
Drop Judges who saw $<$ 50 cases	1.80	2.28	1.91	2.78
	(0.45)	(0.51)	(0.78)	(0.93)
Drop Judges who saw $<$ 500 cases	1.76	2.27	1.80	2.73
	(0.46)	(0.54)	(0.79)	(0.95)
Drop Middle Third of Judges	1.81	2.29	1.94	2.79
	(0.43)	(0.49)	(0.76)	(0.91)
Doyle's Instrument	1.41	1.88	1.53	2.38
	(0.45)	(0.51)	(0.70)	(0.84)
Demean by hearing office-year	1.94	2.43	2.35	3.18
	(0.49)	(0.63)	(0.96)	(1.24)
Underreporting Correction				
Baseline Correction	1.62	2.12	1.66	2.54
	(0.44)	(0.50)	(0.76)	(0.90)
Double Size of Base Correction	1.43	1.92	1.18	2.07
	(0.44)	(0.49)	(0.74)	(0.89)
Set p = .94 for All Denied	1.38	1.87	0.86	1.75
	(0.44)	(0.50)	(0.76)	(0.91)
Set p =.88 for All Denied	0.73	1.24	-0.73	0.18
	(0.44)	(0.50)	(0.75)	(0.90)

Notes: Baseline instrument is judge leniency. Covariates are those in Table 2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. For details on how the correction for underreporting of mortality is calculated, see the discussion in section 5.2. Standard errors clustered by judge.
In the Drop Middle Third of Judges row we only keep judges in the top and bottom thirds of the distribution of judge leniency.
In the Doyle's Intrument row we replace the baseline instrument with the one constructed in the appendix section C.2
In the baseline we demean by hearing office-day and drop judges who saw less than 200 cases, which gives N=610,231.
In the rows where we include those who die within 1 year of seeing a judge, N = 620,237.
In the rows where we drop judges who saw <50 cases, N = 616,599.
In the rows where we drop judges who saw <500 cases, N =601,042.
In the rows where we drop the middle third of judges, N = 408,853.

Table 7: Robustness Checks

we exclude judges who heard less than 200 cases, here we consider a lower threshold of 50 cases and a higher threshold of 500 cases - neither of which change our estimates by much. In the next row we see if we can increase the strength of our instrument by only keeping judges in the top and bottom third in the distribution of judge leniency, our instrumental variable. Our estimates and standard errors are almost unchanged. The next row displays the results from the IV regression where we use the instrument proposed by Doyle Jr (2007) instead of our judge leniency instrument. With Doyle's instrument, our coefficients are somewhat smaller than the baseline but inference is similar. In the next row, instead of demeaning by hearing-office and day as in the baseline, we demean by hearing-office year. The estimates tend to be a bit higher, but standard errors also increase. The bottom four rows all adjust for underreporting, using different values for the correction, p. Our baseline underreporting correction uses individual specific values for p (which average 0.95 and 0.96 at 5 and 10 years, respectively), as described in section 5.2. Next, we assign individual values for p, assuming the undercount is twice as large as for the baseline correction. This adjustment modestly decreases our point estimates. The next underreporting correction assumes that the value of p is the one estimated for individuals aged 25-44, whose mean value of $p = 0.94$, which is much lower than for the baseline. We assign this value of p for all denied individuals. Our final underreporting correction takes an extreme value for p by assuming the undercount for the individuals aged 25-44 is twice as large and assigning this value for everyone. As expected, with the correction the estimates fall but only slightly. It would take a very large undercount of deaths of denied individuals – much larger than is plausible – to change our core inferences.

These robustness checks, taken together, increase our confidence in our estimation strategy. In every case, barring implausibly large underreporting corrections, our estimates are positive, statistically significant, and similar in magnitude to our main estimates.

9 Conclusion and Policy Implications

This paper estimates the effect of Disability Insurance receipt on mortality, for persons on the margin of receiving benefits or not. Those receiving benefits receive large cash transfers, and health insurance from Medicare or Medicaid. However, beneficiaries also face important work disincentives. Each of these factors could affect mortality, but not necessarily with the same sign. We would expect higher income and access to health insurance to cause lower mortality. However, reduced employment may increase mortality. Identifying this combined effect is difficult, however, because those allowed benefits are likely to be less healthy than those not allowed, perhaps in ways observed by the ALJ, but not fully captured by our covariates. We rely for causal inference on the effectively random assignment of judges to disability cases, and on an instrumental variable that measures the tendency for each judge to allow benefits, relative to other judges in the same hearing office on the same day.

We find that benefit receipt increases mortality on average, for those on the margin to receive benefits or not, after both 5 and 10 years. This is consistent with the view that benefit receipt lowers labor supply, which in turn increases mortality. However, Marginal Treatment Effects estimates reveal strong heterogeneity in the response to benefit allowance, even within the range of leniency that we observe. For those aged 55-64, allowance reduces mortality for less healthy applicants who would be allowed by all but the strictest judges, but increases mortality for healthier applicants who would only be allowed by the most lenient judges. These results suggest that significant changes in the DI screening threshold may increase mortality of the marginal applicants. This provides some evidence that current screening thresholds are about right. For younger age groups we find that a modest tightening of the screening thresholds (i.e. making them more strict) will not decrease the longevity of applicants and might increase the longevity of marginal applicants.

Our estimates show that among the healthier individuals, DI receipt increases mortality, but among the less healthy individuals, DI receipt tends to reduce estimated mortality. All of our estimates are for *marginal* recipients, who would receive benefits if seen by a lenient ALJ, but be denied by a stricter one. However, the majority of benefit recipients are *inframarginal* cases who are less healthy than our marginal cases. Thus, our findings are consistent with the view that DI receipt reduces mortality on average.

We also find evidence that for certain expensive, high mortality health conditions such as cancer and respiratory conditions, benefit receipt is relatively more favorable for future mortality, whereas for lower cost, lower mortality conditions such as musculoskeletal disorders, benefit receipt predicts higher mortality for marginal applicants.

Our findings have important policy implications. Given the extreme cost of the disability insurance program, many reform proposals have been put forward, including making the disability criteria more stringent. Our results speak directly to how increasing stringency might impact applicants' health. In general our findings suggest that for maximizing the longevity of current DI applicants, the current disability thresholds are at about the right level. However, a modest increase in program stringency should not increase mortality, and might decrease it, especially for younger applicants.

We also find evidence for the value of health insurance for selected high-cost, high-mortality conditions. This suggests that persons with these conditions who receive benefits might gain from not being subject to the current 2 year waiting period to receive Medicare coverage.

We provide evidence that working appears to reduce mortality, at least on average, for marginal recipients. Thus, reforms of the disability insurance rules to reduce the strong work disincentives of the current rules may improve recipient mortality.

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of econometrics*, 113(2):231–263.

Autor, D. and Duggan, M. (2006). The growth in the social security disability rolls: A fiscal crisis unfolding. *The Journal of Economic Perspectives*, 20(3):71–96.

Autor, D., Kostol, A. R., and Mogstad, M. (2015). Disability benefits, consumption insurance, and household labor supply.

Autor, D. H. and Duggan, M. (2010). Supporting work: a proposal for modernizing the us disability insurance system. *Center for American Progress and The Hamilton Project.*

Benitez-Silva, H., Buchinsky, M., Chan, H. M., Rust, J., and Sheidvasser, S. (1999). An empirical analysis of the social security disability application, appeal, and award process. *Labour Economics*, 6(2):147–178.

Bjorklund, A. and Moffitt, R. (1987). The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics*, pages 42–49.

Black, B., Espin-Sanchez, J.-A., French, E., and Litvak, K. (2017). The long-term effect of health insurance on near-elderly health and mortality. *American Journal of Health Economics*.

Black, B., French, E., McCauley, J., Faria de Carvalho Nunes, L., and Song, J. (2016). Comparing death data from the social security administration and the national center for health statistics, 1990-2013.

Brook, R., Ware, J. E., Rogers, W. H., Keeler, E. B., Davies, A. R., Donald, C. A., Goldberg, G. A., Lohr, K. N., Masthay, P. C., and Newhouse, J. P. (1983). Does free care improve adults' health? results from a randomized trial. *New England Journal of Medicine*, 309(23):1426–1434.

Burkhauser, R. V. and Daly, M. (2011). *The declining work and welfare of people with disabilities: What went wrong and a strategy for change*. AEI Press.

Burkhauser, R. V., Daly, M. C., McVicar, D., and Wilkins, R. (2014). Disability benefit growth and disability reform in the us: lessons from other oecd nations. *IZA Journal of Labor Policy*, 3(1):4.

Card, D., Dobkin, C., and Maestas, N. (2009). Does medicare save lives? *The Quarterly Journal of Economics*, 124(2):597–636.

Dahl, G. B., Kostol, A. R., and Mogstad, M. (2014). Family welfare cultures. *The Quarterly Journal of Economics*, 129(4):1711–1752.

De Nardi, M., French, E., Jones, J. B., and McCauley, J. (2016). Medical spending of the us elderly. *Fiscal Studies*, 37(3-4):717–747.

Doyle Jr, J. J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *The American Economic Review*, 97(5):1583–1610.

Eliason, M. and Storrie, D. (2006). Lasting or latent scars? swedish evidence on the long-term effects of job displacement. *Journal of Labor Economics*, 24(4):831–856.

Finkelstein, A. and McKnight, R. (2008). What did medicare do? the initial impact of medicare on mortality and out of pocket medical spending. *Journal of Public Economics*, 92(7):1644–1669.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., and Group, O. H. S. (2012). The oregon health insurance experiment: Evidence from the first year. *Quarterly Journal of Economics*, 127(3):1057–1106.

Fitzpatrick, M. and Moore, T. (2016). The mortality effects of retirement: Evidence from social security eligibility at age 62. mimeo.

French, E. and Jones, J. B. (2011). The effects of health insurance and self-insurance on retirement behavior. *Econometrica*, 79(3):693–732.

French, E., Jones, J. B., and McCauley, J. (2017). The accuracy of economic measurement in the health and retirement study. In *Forum for Health Economics and Policy*. De Gruyter.

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

French, E. and Taber, C. (2011). Identification of models of the labor market. *Handbook of Labor Economics*, 4:537–617.

GAO (2013). "Preliminary Observations on the Death Master File", Testimony Before the Comittee on Homeland Security and Governmental Affairs.

Gelber, A., Moore, T., and Strand, A. (2017). Disability insurance income saves lives. mimeo.

Hausman, J. A., Newey, W. K., Woutersen, T., Chao, J. C., and Swanson, N. R. (2012). Instrumental variable estimation with heteroskedasticity and many instruments. *Quantitative Economics*, 3(2):211–255.

Heckman, J. J., Urzua, S., and Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, 88(3):389–432.

Hernandez-Pizarro, H. (2016). The effect of long-term care benefits on mortality. mimeo.

Hill, M. E. and Rosenwaike, I. (2002). The social security administration's death master file: The completeness of death reporting at older ages. *Social Security Bulletin*, 64:45–51.

Hu, J., Lahiri, K., Vaughan, D. R., and Wixon, B. (2001). A structural model of social security's disability determination process. *Review of Economics and Statistics*, 83(2):348–361.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Kuhn, A., Wuellrich, J.-P., and Zweimuller, J. (2017). Fatal attraction? early retirement and mortality.

Livermore, G., Stapleton, D. C., and O'Toole, M. (2011). Health care costs are a key driver of growth in federal and state assistance to working-age people with disabilities. *Health Affairs*, 30(9):1664–1672.

Maestas, N., Mullen, K. J., and Strand, A. (2013). Does disability insurance receipt discourage work? using examiner assignment to estimate causal effects of ssdi receipt. *The American Economic Review*, 103(5):1797–1829.

Myers, R. J. (1993). *Social security.* Univ of Pennsylvania Pr.

Paletta, D. (2011). Disability-claim judge has trouble saying 'no'; near perfect approval record; social security program strained. *Wall Street Journal.*

Parsons, D. O. (1991). Self-screening in targeted public transfer programs. *Journal of Political Economy*, pages 859–876.

Rege, M., Telle, K., and Votruba, M. (2009). The effect of plant downsizing on disability pension utilization. *Journal of the European Economic Association*, 7(4):754–785.

Rupp, K. and Riley, G. F. (2012). Longitudinal patterns of medicaid and medicare coverage among disability cash benefit awardees. *Soc. Sec. Bull.*, 72:19.

Snyder, S. E. and Evans, W. N. (2006). The effect of income on mortality: Evidence from the social security notch. *The Review of Economics and Statistics*, 88(3):482–495.

Social Security Advisory Board (2006). Improving the social security administration's hearing process.

Sommers, B. D., Long, S. K., and Baicker, K. (2014). Changes in mortality after massachusetts health care reform: a quasi-experimental study. *Annals of internal medicine*, 160(9):585–593.

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

Taylor, M. H. (2007). "refugee roulette" in an administrative law context: The" déjà vu" of decisional disparities in agency adjudication. *Stanford Law Review*, pages 475–501.

United States (2007). US Code Title 5, Part III, Subpart B, Chapter 31, Subchapter I, Section 3105.

U.S. Census Bureau (2015). Annual estimates of the resident population for selected age groups by sex for the united states, states, counties, and puerto rico commonwealth and municipios: April 1, 2010 to july 1, 2014.

U.S. Social Security Administration (2007). Annual statistical report on the social security disability insurance program.

U.S. Social Security Administration (2008). Social security bulletin annual statistical supplement.

U.S. Social Security Administration (2009). Hearings, appeals and litigation law manual (HALLEX).

U.S. Social Security Administration (2014a). Annual statistical report on the social security disability insurance program.

U.S. Social Security Administration (2014b). Ssi annual statistical report.

Weathers, R. R. and Stegman, M. (2012). The effect of expanding access to health insurance on the health and mortality of social security disability insurance beneficiaries. *Journal of health economics*, 31(6):863–875.

A Data

We use the universe of all DI or SSI appeals heard by ALJs, 1995-2004. We merge data from the following sources: the Office of Hearings and Appeals Case Control System (OHACCS), the Hearing Office Tracking System (HOTS), the Appeals Council Automated Processing System (ACAPS), the Litigation Overview Tracking System (LOTS), the SSA 831 file, SSA Master Earnings file (MEF), the Master Beneficiary Record (MBR), the Supplemental Security Record (SSR), and the SSA Numerical Identification (NUMIDENT) file.

The OHACCS data contain details of Social Security DI and SSI cases adjudicated at the ALJ level (plus limited information on cases heard at the Appeals Council or in federal court). The OHACCS data also include cases involving Retirement and Survivors Insurance and Medicare Hospital insurance. We keep only the SSI and DI cases. The OHACCS data are used for administering DI and SSI cases, and are thus very accurate. They include information on the judge assigned to the case, the hearing office, the date of assignment, and the case outcome (such as allowed or denied), the claimant's Social Security number and type of claim (DI versus SSI). Because the SSA mortality data is less complete prior to 1995, we use OHACCS data only for 1995-2014.

Until 2004, individual hearing offices maintained their own data, called the Hearing Office Tracking System (HOTS). These data were then uploaded to the OHACCS system. We found some missing cases in the OHACCS system, apparently the result of HOTS data not being properly uploaded. The problem occurs in about 1% of all cases. For these cases we augment the OHACCS data with HOTS. After 2004, all uploading of data is automatic, and thus there are no problems with missing data.

Although OHACCS also contains Appeals Council records, Appeals Council decisions are sometimes missing from OHACCS. Thus we use the Appeals Council Automated Processing System (ACAPS) data, which the Appeals Council uses to administer cases, to track outcomes for cases heard at the Appeals Council level.

The Litigation Overview Tracking System (LOTS) data are used by SSA to administer cases that were denied by the Appeals Council but then reach federal courts. We combine the LOTS data with information provided by the Federal Court to determine whether the cases was eventually allowed or denied. The SSA 831 data have information on the details of the DI application received by the Disability Determination Service. The data include the type of application (whether DI or SSI or concurrent) and whether the claim is based on one's own earnings history or on the history of a spouse or parent. It also has all the information relevant for determining whether the application should be allowed at the initial level, before reaching an ALJ, based on the applicant having a listed medical condition or the vocational grid. Thus we have detailed information on applicants' health at time of application. Because of the vocational grid, we have information on age, education, industry and occupation. We also have some other demographic information such as sex. Since a new 831 record is established whenever a new application is filed and adjudicated, we use information in the 831 file to identify those who reapplied for benefits.

The Master Earning File (MEF) includes annual longitudinal earnings data for the US population, taken directly from W-2 filings, starting from 1978. Wage earnings are not top-coded. Self-employment earnings are top coded until 1992. Our earnings measure is the sum of wage earnings and self employment earnings, which we topcode at $200,000 per year, the topcoding affects only a very small percentage of applicants.

The Master Beneficiary Record (MBR) includes beneficiary and payment history data for the entire Social Security OASDI program. The Supplemental Security Record (SSR) contains information on individuals applying for SSI benefits. We use the MBR and SSR to identify disability benefit award status of individuals.

Lastly, we use the SSA NUMIDENT file for information on date of death. The NUMIDENT file includes information from the Social Security Number application form such as name, date of birth and Social Security number, and once the individual dies, the date of death.

For Figure 1 we use all cases filed 1989-1999. For all other figures and tables, we begin with the universe of all cases adjudicated by an ALJ and make the sample restrictions, described in Table A1. We drop a relatively small number of cases who died within the year of assignment to the judge, had missing education data, or where the judge handled fewer than 200 cases. This leaves an estimation sample with 2,759,907 cases. In many analyses we further restrict the sample to persons age 55-64 at application, which is 610,231 cases.

	Sample Size
Original data	3,368,017
Number of drops	
Age at application <25 or >64	339,515
Died the year of application	30,807
Missing education data	204,859
Judge handled fewer than 200 cases	32,929
Remaining sample (Aged 25-64)	2,759,907
Age at application: 25-44	1,101,332
Age at application: 45-54	1,048,344
Age at application: 55-64 (Main Sample)	610,231

Notes: The original sample excludes those with missing judge or hearing office information, pre-viewed cases, DI Child cases, and Survivor cases.

Table A1: Sample Selection

B Additional Tables and Figures

B.1 Main Tables: All Ages

Table 2 in the text provides evidence for random assignment for our main estimation sample (age 55-64 at time of application). Table A3 provides a similar table for the full sample. The last two columns show whether the instrument (judge leniency) significantly predicts our

Year	All (20+)	55+	55-59	60-64	55-64	65+
1995	96.6	96.9	96.0	96.5	96.3	97.0
1996	96.8	97.0	96.1	97.1	96.6	97.0
1997	97.0	97.2	96.8	97.1	96.9	97.2
1998	97.1	97.3	96.6	97.2	96.9	97.4
1999	97.5	97.7	97.5	97.9	97.7	97.7
2000	97.7	98.0	97.9	98.2	98.1	97.9
2001	97.9	98.2	98.7	98.8	98.8	98.1
2002	98.1	98.4	98.7	99.4	99.1	98.3
2003	98.2	98.4	98.8	99.5	98.1	98.3
2004	98.6	99.0	98.9	99.6	99.3	99.0
2005	98.8	99.2	98.8	99.6	99.2	99.2
2006	98.8	99.3	98.6	99.6	99.1	99.3
2007	99.1	99.6	98.8	99.7	99.3	99.6
2008	99.4	99.8	99.3	99.6	99.5	99.8
2009	99.4	99.8	98.8	99.7	99.3	99.9
2010	99.7	100.0	99.3	100.0	99.7	100.1
2011	99.7	100.0	99.1	99.9	99.5	100.1
2012	99.7	100.1	98.7	99.9	99.4	100.2
2013	99.7	100.0	98.5	99.5	99.0	100.2
2014	98.6	99.1	96.5	97.7	97.2	99.4
Average	98.4	98.8	98.1	98.8	98.5	98.8

Notes: Estimated ratio of deaths in the SSA Numident data to adjusted National Death Index deaths over 1995-2014, by age group. Total (20+) column excludes children (age 0-19). The estimated ratio is calculated as 100 $\times$ D_{kt}/O_{kt} where D_{kt} represents the number of deaths reported in the SSA data for age group k occurring in year t and O_{kt} represents the official number of deaths of US residents reported in the NDI for age group k during year t.

Table A2: Estimated Percentage of US Deaths Included in the SSA Death Data, 1990-2014, by Age Group

covariates which it should not, if assignment is random. Of the 20 covariates in the table, only one takes a coefficient with a t-statistic $>$ 2.0, and only mildly so (female t-statistic = 2.2). This is consistent with random assignment.

Table 3 in the main text provides first-stage results for our main estimation sample (age 55-64 at time of application), disaggregated by gender, income, health, and other subgroups of our 55-64 sample. Table A4 provides a similar table for the full sample. The allowance rates are lower for the full sample (70.8%) than for our 55-64 subsample (84.1%). The full sample coefficient from regressing allowance on judge leniency is 0.966, and thus close to 1, as it should be since we estimate judge leniency using the full sample. The comparable estimate for applicants aged 55-64 is 0.676; thus, judge leniency has a larger effect on allowance rates for younger applicants. The monotonicity assumption again cannot be rejected, with most relative likelihoods close to 1.

Covariate	Dependent Variable: Allowed		Dependent Variable: Judge Leniency	
	Coefficient	t-stat	Coefficient	t-stat
	(1)	(2)	(3)	(4)
Sex				
Female	0.0175	16.1	0.0008	2.2
Age				
55 to 59	-0.1073	-53.5	-0.0109	-1.5
Race				
Black	-0.0582	-26.5	-0.0025	-1.0
Other (non-black, non-white) or unknown	-0.0085	-4.1	-0.0017	-0.9
Labor force participation and income				
Average participation rate, years -11 to -2	0.0043	6.5	0.0005	0.8
Average earnings/billion, years -11 to -2 ($2006)	0.0012	19.7	0.0000	1.2
Represented by lawyer				
Represented by lawyer	0.0479	5.5	-0.0075	-1.5
Application type				
SSDI	0.0289	20.1	0.0016	0.6
Education				
High school graduate, no college	-0.0044	-4.5	-0.0008	-1.0
Some college	-0.0154	-10.7	-0.0025	-1.6
College graduate	-0.0032	-1.7	-0.0022	-1.6
Health conditions (by diagnosis group)				
Neoplasms (e.g., cancer)	0.0436	17.2	0.0018	0.9
Mental disorders	-0.0207	-9.5	-0.0012	-1.2
Mental retardation	0.0007	0.2	0.0018	0.8
Nervous system	0.0009	0.5	-0.0008	-0.7
Circulatory system (e.g., heart disease)	0.0235	14.9	0.0024	1.1
Musculoskeletal disorders (e.g., back pain)	-0.0036	-2.3	0.0003	0.4
Respiratory system	-0.0281	-13.8	-0.0011	-1.4
Injuries	-0.0090	-4.4	-0.0007	-0.7
Endocrine system (e.g., diabetes)	0.0182	10.1	0.0008	0.7
Standard deviation of dependent variable	0.4116		0.1058	
R^2	0.0361		0.0040	
Number of Applicants = 2,759,907	Number of Judges =1,436			

Notes: Column (1) is from a regression of de-meaned allowance on all the covariates listed. Column (3) is from a regression of judge leniency on all the covariates listed. Omitted category is male, 55-64s, white, not represented by a lawyer, applying for SSI or SSI and DI concurrently, not a high school graduate, with a health condition other than those listed above. The sample includes all applicants aged 25 to 64, and we exclude applicants who died the year of application. Standard errors clustered by judge.

Table A3: Predictors of Allowance and Judge Leniency, All Ages

	Obs.	Allowance Rate at ALJ Stage	Coefficient on Judge Leniency	Std. Error	T-Ratio	Relative Likelihood*
	(1)	(2)	(3)	(4)	(5)	(6)
All groups						
All groups	2,759,907	0.708	0.966	(0.019)	52	1.000
Sex						
Male	1,322,817	0.704	0.947	(0.023)	41	0.980
Female	1,437,090	0.711	0.984	(0.015)	66	1.019
Age						
25 to 54	2,149,676	0.670	1.022	(0.020)	51	1.058
55 to 64	610,231	0.841	0.705	(0.012)	60	0.730
Race						
White	1,738,652	0.737	0.939	(0.011)	84	0.972
Black	546,125	0.637	1.007	(0.039)	26	1.042
Other or unknown	475,130	0.682	1.001	(0.021)	47	1.036
Income						
Average earnings $<$ \$10000	1,587,843	0.644	1.035	(0.032)	33	1.071
Average earnings $\geq$ \$10000	1,172,064	0.794	0.839	(0.007)	121	0.868
Represented by lawyer						
Represented by lawyer	1,802,345	0.732	0.946	(0.007)	130	0.979
Not represented by lawyer	957,562	0.663	1.023	(0.051)	20	1.059
Application type						
SSDI	1,144,427	0.774	0.869	(0.008)	103	0.899
SSI or Concurrent (both SSDI and SSI)	1,615,480	0.661	1.023	(0.026)	39	1.059
Education						
Less than high school	918,011	0.693	0.981	(0.025)	39	1.015
High school graduate, no college	1,287,621	0.712	0.964	(0.017)	57	0.997
Some college	399,954	0.708	0.978	(0.015)	63	1.012
College graduate	154,321	0.763	0.865	(0.014)	61	0.896
Health conditions (by diagnosis group)						
Neoplasms (e.g., cancer)	55,935	0.791	0.763	(0.023)	34	0.790
Mental disorders	506,499	0.660	1.072	(0.023)	47	1.109
Mental retardation	32,893	0.645	1.022	(0.049)	21	1.058
Nervous system	172,606	0.708	0.922	(0.022)	42	0.954
Circulatory system (e.g., heart disease)	267,349	0.765	0.853	(0.018)	47	0.883
Musculoskeletal disorders	1,008,542	0.722	0.966	(0.013)	73	1.000
Respiratory system	96,781	0.686	0.967	(0.035)	27	1.001
Injuries	159,977	0.687	0.969	(0.026)	37	1.003
Endocrine system (e.g., diabetes)	144,969	0.723	0.913	(0.024)	39	0.945
All other	314,356	0.694	0.946	(0.028)	34	0.979

Notes: Column (3) displays the first stage estimate of the coefficient λ from the regression of de-meaned allowance rates on judge leniency for the full sample. Average earnings is calculated on income between 11 and 2 years before application. Standard errors clustered by judge.
*Relative likelihood is the ratio of the group specific coefficient on judge leniency (presented in column 4) to the full sample coefficient

Table A4: First Stage Estimates: Regression of Allowance Rates on Judge Leniency, By Demographics, All Ages

		Panel (a): 5 Year Mortality (Percent)						Panel (b): 10 Year Mortality (Percent)					
		Mortality Rates		OLS		IV		Mortality Rates		OLS		IV	
	Obs	Allowed	Denied	Diff.	SE	Diff.	SE	Allowed	Denied	Diff.	SE	Diff.	SE
All groups	610,231	9.71	8.35	1.94	(0.12)	2.30	(0.50)	21.95	19.99	2.77	(0.18)	2.81	(0.91)
Age Band													
55 to 59	390,600	9.13	7.71	1.96	(0.14)	2.56	(0.60)	20.60	18.88	2.64	(0.21)	2.79	(1.03)
60 to 64	219,631	10.72	9.60	1.94	(0.20)	1.91	(0.85)	24.32	22.13	3.11	(0.29)	3.03	(1.28)
Represented by lawyer													
Represented by lawyer	385,118	9.18	7.99	1.76	(0.15)	2.16	(0.78)	21.19	19.34	2.75	(0.22)	3.13	(1.20)
Not represented by lawyer	225,113	10.65	8.85	2.22	(0.20)	2.78	(0.64)	23.31	20.89	2.86	(0.29)	2.78	(1.10)
Application type													
SSDI	352,991	8.49	8.04	1.73	(0.15)	2.40	(0.67)	19.67	18.94	2.73	(0.23)	3.61	(1.03)
SSI or SSI/SSDI concurrent	257,240	11.44	8.69	2.57	(0.19)	2.45	(0.76)	25.22	21.15	3.54	(0.27)	2.37	(1.38)

Notes: This table displays the estimated effect of DI recipiency on mortality (with covariates) for those aged 55-64. Mortality Rates do not adjust for covariates. OLS and IV estimates are demeaned and include covariates. Covariates are those in Table 2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. IV estimates use demeaned variables and judge leniency as the instrument. Standard errors clustered by judge.

Table A5: Estimated Effect of DI Recipiency on Mortality, Disaggregated cont'd, Aged 55-64

		Panel (a): 5 Year Mortality (Percent)						Panel (b): 10 Year Mortality (Percent)					
		Mortality Rates		OLS		IV		Mortality Rates		OLS		IV	
	Obs	Allowed	Denied	Diff.	SE	Diff.	SE	Allowed	Denied	Diff.	SE	Diff.	SE
All groups	2,759,907	7.47	5.11	2.21	(0.05)	1.64	(0.58)	16.29	12.24	3.55	(0.10)	2.96	(1.05)
Sex													
Male	1,322,817	9.65	6.94	2.53	(0.07)	1.85	(0.60)	20.35	15.97	3.75	(0.12)	3.25	(1.04)
Female	1,437,090	5.48	3.38	1.92	(0.06)	1.43	(0.55)	12.59	8.72	3.34	(0.10)	2.66	(1.03)
Race													
White	1,738,652	7.35	5.45	1.82	(0.06)	1.56	(0.52)	16.29	12.81	3.15	(0.11)	3.04	(1.02)
Black	546,125	8.65	5.30	3.00	(0.10)	1.70	(0.91)	17.81	12.44	4.46	(0.16)	3.00	(1.51)
Other	475,130	6.67	3.83	2.55	(0.08)	1.85	(0.62)	14.65	10.27	3.78	(0.13)	2.66	(0.97)
Education Group													
Less than high school	918,011	7.81	5.23	2.23	(0.07)	1.55	(0.50)	17.40	12.64	3.74	(0.13)	2.84	(1.05)
High school graduate	1,287,621	7.31	5.06	2.21	(0.07)	1.57	(0.65)	15.88	12.03	3.51	(0.11)	2.91	(1.11)
Some college	399,954	7.35	4.97	2.33	(0.09)	1.61	(0.68)	15.60	11.97	3.56	(0.15)	3.30	(1.22)
College graduate	154,321	7.16	5.19	1.98	(0.16)	3.06	(1.14)	15.08	12.18	2.83	(0.23)	3.38	(1.39)
Income													
Average earnings $<$ \$10000	1,587,843	7.92	5.33	2.30	(0.06)	1.58	(0.54)	17.03	12.70	3.65	(0.11)	3.02	(0.96)
Average earnings $\geq$ \$10000	1,172,064	6.97	4.59	2.04	(0.07)	1.76	(0.69)	15.47	11.17	3.34	(0.11)	2.87	(1.27)
Health conditions													
Neoplasms	55,935	23.74	24.47	-0.16	(0.65)	2.71	(6.73)	36.91	40.75	-2.42	(0.82)	2.94	(7.62)
Respiratory system	96,781	11.43	6.57	3.77	(0.21)	1.00	(1.15)	25.57	17.59	5.73	(0.32)	3.04	(2.19)
Endocrine system	144,969	12.53	7.06	5.03	(0.18)	2.96	(1.34)	27.72	17.93	8.76	(0.26)	4.72	(2.58)
Circulatory system	267,349	11.57	7.99	3.79	(0.15)	3.36	(1.02)	26.00	19.63	6.18	(0.25)	5.28	(1.90)
Mental retardation	32,893	5.60	3.51	1.76	(0.25)	1.02	(0.88)	12.29	8.87	2.73	(0.38)	0.56	(1.53)
Nervous system	172,606	6.58	5.15	1.93	(0.15)	0.94	(0.88)	14.85	12.27	3.17	(0.21)	1.12	(1.38)
Mental disorders	506,499	5.61	4.15	1.52	(0.07)	1.61	(0.29)	12.47	10.25	2.27	(0.11)	2.54	(0.42)
Injuries	159,977	5.37	3.77	1.65	(0.05)	1.12	(0.21)	12.15	8.93	2.92	(0.08)	2.41	(0.46)
Musculoskeletal disorders	1,008,542	4.59	3.10	1.55	(0.12)	1.11	(0.50)	11.12	8.25	2.92	(0.18)	3.04	(0.80)
All other	314,356	10.89	8.07	3.31	(0.18)	2.27	(1.70)	21.03	17.22	4.57	(0.28)	4.30	(2.45)

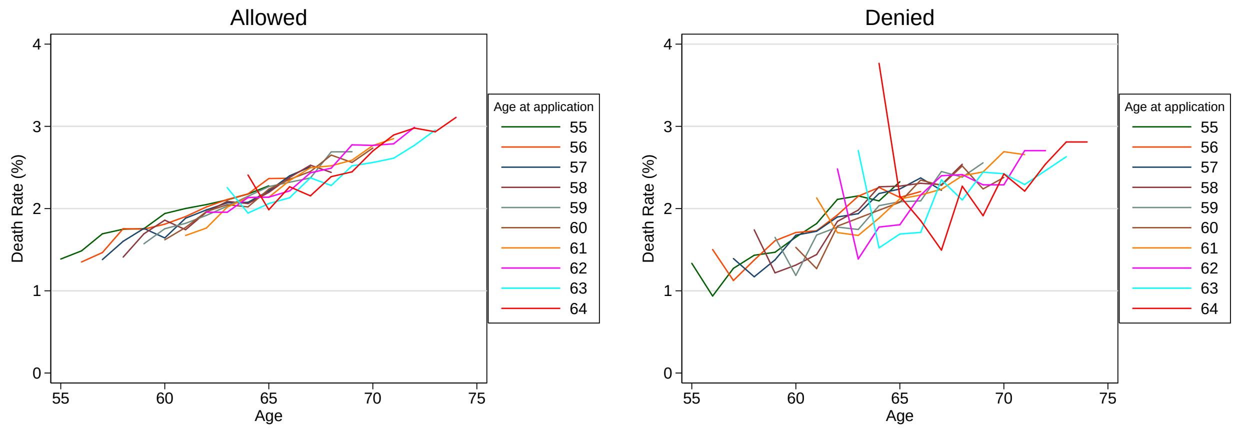
Notes: This table displays the estimated effect of DI recipiency on mortality (with covariates) for the full sample (all ages). Mortality Rates do not adjust for covariates. OLS and IV estimates are demeaned and include covariates. Covariates are those in Table 2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. IV estimates use demeaned variables and judge leniency as the instrument. Standard errors clustered by judge.

Table A6: Estimated Effect of DI Recipiency on Mortality, Disaggregated, All Ages

		Panel (a): 5 Year Mortality (Percent)						Panel (b): 10 Year Mortality (Percent)					
		Mortality Rates		OLS		IV		Mortality Rates		OLS		IV	
	Obs	Allowed	Denied	Diff.	SE	Diff.	SE	Allowed	Denied	Diff.	SE	Diff.	SE
All groups	2,759,907	7.47	5.11	2.21	(0.05)	1.64	(0.58)	16.29	12.24	3.55	(0.10)	2.96	(1.05)
Age Band													
25 to 44	2,149,676	6.67	4.67	2.25	(0.06)	1.54	(0.54)	14.27	11.18	3.66	(0.10)	3.05	(0.88)
45 to 64	610,231	9.71	8.35	2.21	(0.12)	2.55	(0.64)	21.95	19.99	3.48	(0.23)	3.70	(1.56)
Represented by lawyer													
Represented by lawyer	1,802,345	6.97	4.99	1.90	(0.05)	1.40	(0.45)	15.51	12.04	3.16	(0.08)	2.55	(0.78)
Not represented by lawyer	957,562	8.50	5.29	2.75	(0.09)	2.05	(0.75)	17.89	12.55	4.25	(0.17)	3.68	(1.44)
Application type													
SSDI	1,144,427	6.29	4.72	1.55	(0.07)	1.61	(0.67)	14.25	11.26	2.66	(0.12)	2.96	(1.21)
SSI or SSI/SSDI concurrent	1,615,480	8.45	5.29	2.61	(0.06)	1.67	(0.57)	17.98	12.70	4.08	(0.11)	2.99	(1.04)

Notes: This table displays the estimated effect of DI recipiency on mortality (with covariates) for the full sample (all ages). Mortality Rates do not adjust for covariates. OLS and IV estimates are demeaned and include covariates. Covariates are those in Table 2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. IV estimates use demeaned variables and judge leniency as the instrument. Standard errors clustered by judge.

Table A7: Estimated Effect of DI Recipiency on Mortality, Disaggregated cont'd, All Ages



Notes: Top panels: all years of data. Bottom panels: drops observations year of assignment to the ALJ..

Figure A1: Mortality of those Allowed and Denied

B.2 Quality of SSA Mortality Data

In section 5 we described some of the reasons why mortality rates of those denied benefits might be undercounted in the SSA data. By showing that aggregate mortality rates are very similar in both the SSA data and the National Death Index, we provided evidence that this undercount was not a serious issue. In this appendix we provide further evidence that the SSA data accurately measures mortality of those denied benefits.

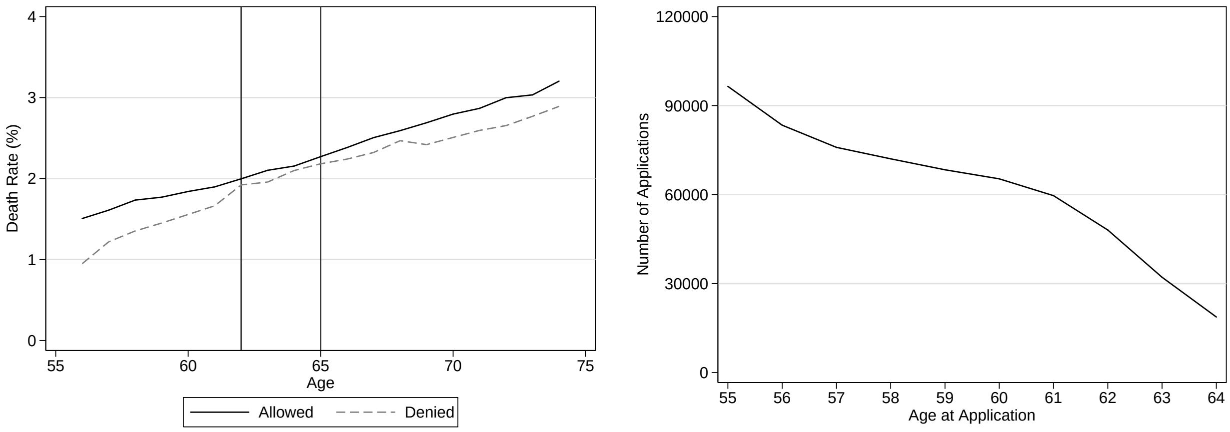
To provide an alternative approach to measuring mortality undercounts, we note that most individuals who are denied DI benefits will take regular retirement benefits at either ages 60, 62, or 65.[16] And once individuals are receiving benefits, we should expect SSA data to have high accuracy in recording deaths, both for those allowed DI benefits, and those who are receiving regular retirement benefits. Thus, we should expect any undercount of mortality for those who are denied benefits to occur principally prior to these ages. If there is significant undercounting prior to these ages, we should also expect to see a jump in mortality rates at these ages.

In Figure A1 we plot mortality rates at different ages, separately for those allowed and denied by an ALJ, by age at application, for our main estimation sample (age 55-64 at application). The figure shows mortality for up to 10 years after assignment and includes data for the year of assignment. The left panel shows mortality of those allowed by an ALJ, whereas the right panel shows mortality of those denied.

Mortality rates of those allowed rise from approximately 1.3% at age 56 to 3.0% by age 74. The lines are not perfectly smooth, but this is mostly due to sampling variability. There is no noticeable jump in mortality rates after any particular age. Among those denied, mortality rates are slightly lower than for those allowed. However, unlike the allowed, the denied appear to have spikes in mortality in the year of application.

The size of the spikes in mortality rates for the denied rises progressively for those who apply between ages 61 through 64, and there is an increase for those allowed also at ages

[16] Many widows and widowers can draw benefits at 60. The Social Security Early Retirement Age is 62. The Normal Retirement Age is 65 or 66, depending on the year.



Notes: Top panels: all years of data. Bottom panels: drops observations year of assignment to the ALJ..

Figure A2: Mortality of those Allowed and Denied

63 and 64. This high first year mortality rate is potentially due to the sample becoming increasingly selected towards individuals with high near term mortality. Note that the number of applicants drops sharply after age 60, as shown in the right panel of Figure A1. The applicants who apply shortly before the regular retirement age of 65, and at or after the early retirement age of 62, are self-selected, potentially in different ways than those who apply at earlier ages.[17]

The high first year mortality among that we observe for the denied (but less so for the allowed) is potentially due to mismeasurement. Although we drop those whose death was recorded before an ALJ decision was made, there is the possibility that individuals who die prior to having their case heard might errantly be defined as being denied. It is for this reason that we drop sample members who die in the year of assignment in all of our main results. However, in Section 8 we show that whether or not we include these individuals has virtually no effect on our estimates.

Following the approach of our main analysis, in the left panel of Figure A2 we drop individuals who die within the year of assignment. The left panel of Figure A2 pools the mortality data from Figure A1 by age to present a clearer pattern. While there is no jump in mortality rates at ages 60 or 65, it shows a small jump in mortality at age 62 for denied individuals, which is consistent with potential underreporting of mortality amongst these people before 62. We provide evidence on the size of the potential underreport below, giving us an alternative estimate of the underreporting correction p. However, we should also point out that Fitzpatrick and Moore (2016), using data from the National Center for Health Statistics' Multiple Cause of Death (MCOD) files, document a two percent increase in male mortality immediately after age 62, and argue that the jump in mortality is caused by the fall in labor supply at this age. This could explain the jump that we observe, which is also consistent with the findings in this paper.

[17] One reason for applying, even though one would receive benefits for a short period of time, is if applicants have expensive high-mortality conditions, and hope to receive Medicaid (available to SSI applicants). The conditions that prompt application could also lead to high first-year mortality, relative to those who apply when younger.

Here we make the assumption that the jump in mortality at age 62 is caused entirely by measurement error, and estimate the potential size of any undercount based on that jump. Let m^* denote the observed mortality rate, which is a function of age, and can be seen in the left panel of Figure A2. Let $f(age) = \sum_{k=0}^{K} \gamma_k age^k$ denote the true age-specific mortality rate for the denied and as before p is the underreporting rate for those who are not claiming either disability or Social Security benefits. Let b_{age} denote the percentage of the population at that age that are claiming social security benefits.[18] Then

$$m^* = p(1 - b_{age})f(age) + b_{age}f(age) + \epsilon_{it}.$$

We estimate p using nonlinear least-squares estimation using different functional forms for $f(age)$. The estimates are sensitive to the ages used and also the order of the polynomial K. If we use ages 59-68 and $K = 2$, we estimate $\hat{p} = .90$. If instead we use ages 56-74 and $K = 2$, then we estimate $\hat{p} = 1.02$. This approach provides often lower estimates of p than in our approach in the main text. However, this provides additional evidence that p is in the range of 0.9 to 1.0. Table 7 in Section 8 presents estimates which accounts for p being in this range. The key results of this paper do not markedly change when using these values of p.

B.3 Marginal Treatment Effect Estimates

The estimates displayed in Figure A3 are calculated the same as the MTE estimates displayed in Figure 4, with the only difference being that Figure A3 does not controls for covariates. When we control for covariates the slopes of the graphs change. For the sample aged 55-64 the slopes become more steep, and for those aged 25-64 the slope becomes less steep. The results for the aged 55-64s remain similar.

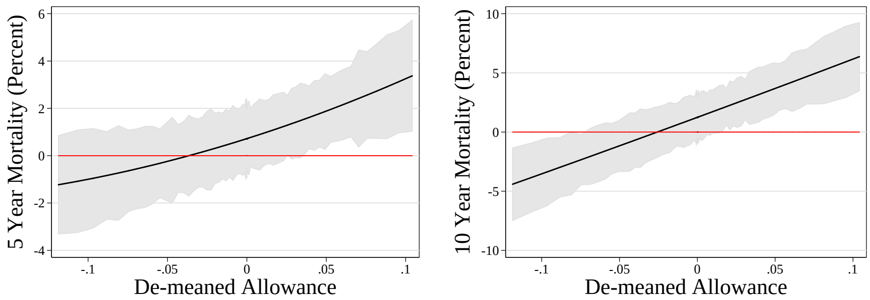
Figure A4 presents MTE graphs using local polynomial smoothed estimates. While the first stage is calculated in the same way as in the MTE figures displayed in the main text (where allowance is regressed on a cubic of the instrument), the second stage is calculated using the local linear smoother from Maestas et al. (2013). Specifically, estimate a local quadratic regression of mortality (de-meaned by hearing offfice and day) and compute numerical derivatives to estimate the MTE $(\partial E[\tilde{y}]/\partial \hat{\tilde{A}})$. Despite the difference in methodology, the estimated MTEs presented in Figure A4 are similar to the ones presented in Figures A3 and 4.

C Derivations

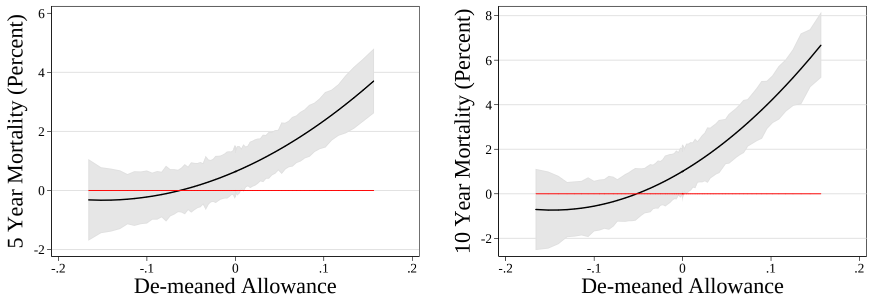
C.1 Marginal Treatment Effects

All derivations in this appendix are purely for completeness – they are straightforward adaptations of the results in Heckman et al. (2006) and French and Taber (2011). Define A_i as a 0-1 indicator =1 if individual i is allowed benefits, y_i is mortality. Other variables are

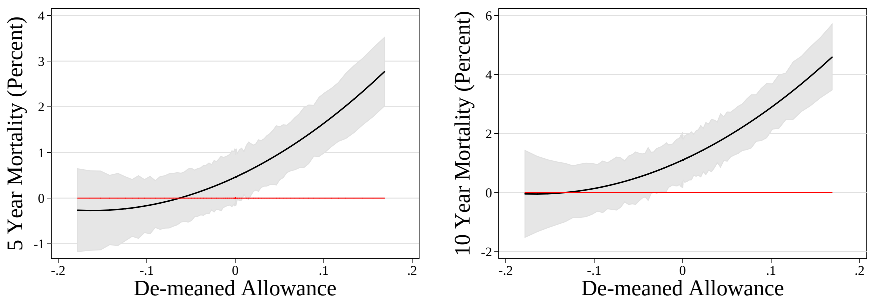
[18] Note that before age 62, b_{age} is small, but jumps to 50% at 62, and is close to 95% by age 66. We are grateful to Timothy Moore for providing this data for each age.



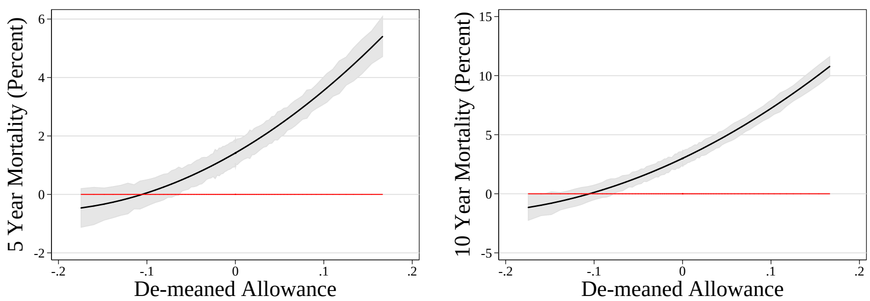
(a) Aged 55-64



(b) Aged 45-54



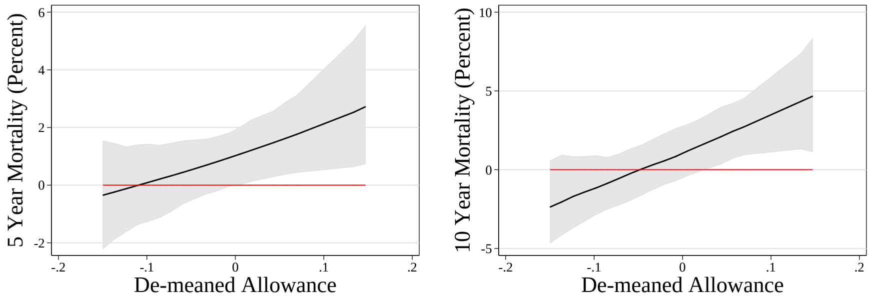
(c) Aged 25-44



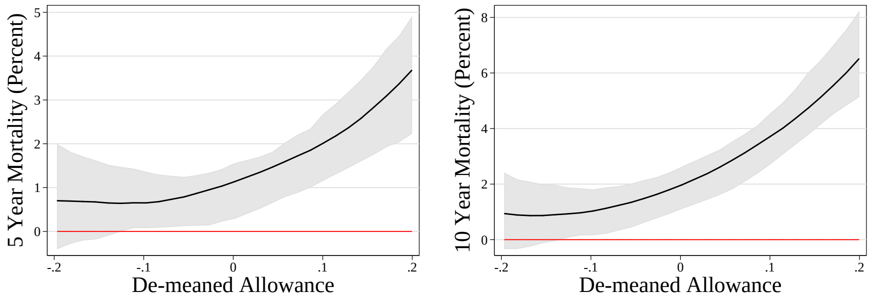
(d) All Ages (25-64)

Notes: This figure displays the estimated mortality response as a function of predicted de-meaned allowance. This is the same as Figure 4, but the estimates were calculated without covariates. Within each panel: the left figure displays 5 year mortality, and the right figure displays 10 year mortality. Mean allowance rate is 0.84 for those aged 55-64, 0.71 for those aged 45-54, 0.63 for those aged 25-44, and 0.71 for those aged 25-64.

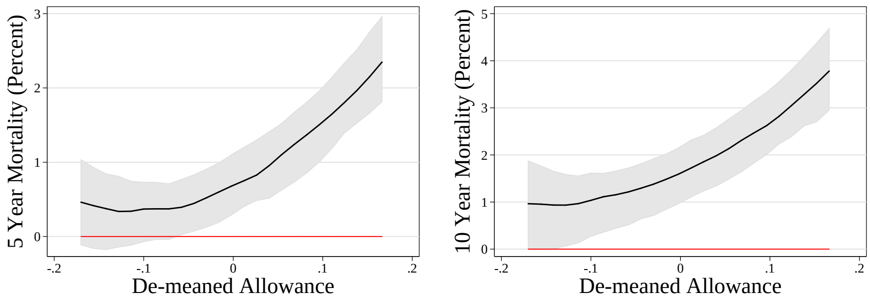
Figure A3: Marginal Treatment Effects: Mortality Response by De-meaned Allowance, Without Covariates



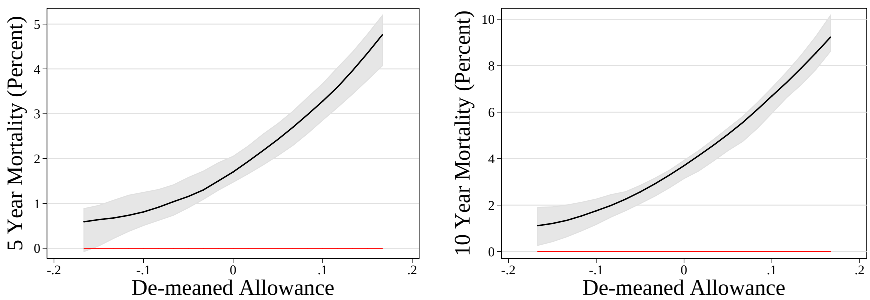
(a) Aged 55-64



(b) Aged 45-54



(c) Aged 25-44



(d) All Ages

Notes: This figure displays the estimated mortality response as a function of predicted de-meaned allowance. This is the same as Figure 4, but the estimates were calculated using local polynomial smoothed estimates without covariates. Within each panel: the left figure displays 5 year mortality, and the right figure displays 10 year mortality. Mean allowance rate is 0.84 for those aged 55-64, 0.71 for those aged 45-54, 0.63 for those aged 25-44, and 0.71 for those aged 25-64.

Figure A4: Marginal Treatment Effects: Mortality Response by De-meaned Allowance, Using Local Polynomial Smoothed Estimates defined in text. The outcome variable for individual i is:

$$y_i = \begin{cases} y_{1i} & \text{if} \quad A_i = 1 \\ y_{0i} & \text{if} \quad A_i = 0 \end{cases} \tag{10}$$

where

$$\begin{aligned} y_{1i} &= \phi + X_i\delta_y + u_{1i} \\ y_{0i} &= X_i\delta_y + u_i \end{aligned} \tag{11}$$

Combining equations (10) and (11) yields:

$$y_i = A_i\phi_i + X_i\delta_y + u_i. \tag{12}$$

where $\phi_i = \phi + u_{1i} - u_i$. Allowance is determined by

$$A_i = 1\{g(Q_i) - V_i > 0\} \tag{13}$$

where $1\{g(Q_i) - V_i > 0\}$ is the indicator function and is equal to 1 when $g(Q_i) - V_i > 0$, $g(Q_i)$ is an arbitrary function of $Q_i = (X_i, Z_i)$, where Z_i is our judge leniency measure described in the text, and V_i can be interpreted as a measure of the health of individual i. The variables u_i and ϕ_i are potentially correlated with each other but by assumption V_i is independent of Z_i and X_i. The Marginal Treatment Effect is

$$MTE(X_i = x, V_i = p) \equiv E[y_{1i} - y_{0i}|X_i = x, V_i = p] \tag{14}$$

where $P(Q_i) \equiv \Pr(A_i = 1|Q_i)$. Given equation (11), $MTE(X_i = x, V_i = p) = \phi + u_{1i} - u_{0i} = \phi_i$. Using equation (12), we estimate the conditional expectation function

$$\begin{aligned} E[y_i|X_i = x, P(Z_i) = p] &= E[A_i\phi_i + X_i\delta_y + u_i|X_i = x, P(Q_i) = p] \\ &= E[A_i(\phi + u_{1i} - u_i)|X_i = x, P(Q_i) = p] + X_i\delta_y + E[u_i|X_i = x, P(Q_i) = p] \\ &= E[A_i\phi|X_i = x, P(Q_i) = p] + E[(u_{1i} - u_i)|A_i = 1, X_i = x, P(Q_i) = p]p + X_i\delta_A \\ &\quad + E[u_i|X_i = x, P(Q_i) = p] \end{aligned} \tag{15}$$

where the step $E[A_i(u_{1i} - u_i)|X_i = x, P(Q_i) = p] = E[(u_{1i} - u_i)|A_i = 1, X_i = x, P(Q_i) = p] \Pr[A_i = 1|X_i = x, P(Q_i) = p]$ follows from the Law of Total Probability, and noting that $\Pr[A_i = 1|X_i = x, P(Q_i) = p] = p$. Continuing with the simplifications, and noting that we have already assumed that u_{1i}, u_i are independent of X_i we have:

$$\begin{aligned} E[y_i|X_i = x, P(Q_i) = p] &= \phi p + E[(u_{1i} - u_i)|A_i = 1, P(Q_i) = p] + X_i\delta_A + E[u_i|P(Q_i) = p] \\ &= X_i\delta_A + \phi p + E[(u_{1i} - u_i)|A_i = 1, P(Q_i) = p]p + E[u_i|P(Q_i) = p] \\ &= X_i\delta_A + K(p) \end{aligned} \tag{16}$$

where $K(p) \equiv \phi p + E[(u_{1i} - u_i)|A_i = 1, P(Q_i) = p]p + E[u_i|P(Q_i) = p]$. Differentiating equation (16) with respect to p yields

$$\frac{\partial E[y_i|X_i = x, P(Q_i) = p]}{\partial p} = K'(p) \tag{17}$$

This derivative is equal to the Marginal Treatment Effect. To see this, note that as a normalization we can let the distribution of V_i be uniform $[0, 1]$, so

$$\begin{aligned}
\frac{\partial E[y_i|X_i = x, P(Q_i) = p]}{\partial p} &= \frac{\partial \left[\int_0^p E[y_{1i}|X_i = x, V_i]dV_i + \int_p^1 E[y_{0i}|X_i = x, V_i]dV_i \right]}{\partial p} \\
&= E[y_{1i}|X_i = x, V_i = p] - E[y_{0i}|X_i = x, V_i = p] \\
&\equiv MTE(X_i = x, V_i = p). \qquad (18)
\end{aligned}$$

Thus estimation of equation (16) and taking $K'(p)$ yields the MTE. In the text we refer to $P(Q_i)$ as the plim of $\widehat{A}_i$.

C.2 De-Meaning the Data and Doyle's Instrument

In our estimation procedure, we have just under 200,000 hearing office-day interactions as covariates, so directly estimating equations (1) and (2) is not computationally feasible. To simplify the problem we de-mean variables by hearing office and day, and construct variables $\widetilde{A}_i = A_i - \bar{A}_i$, $\widetilde{y}_{i\tau} = y_{i\tau} - \bar{y_{i\tau}}$ where $\bar{A}_i$ and $\bar{y_{i\tau}}$ are the mean values of $A_i, y_{i\tau}$ for the hearing office-day on which case i was assignment.

Our instrument, from equation (3) of the text, which we rewrite below, is:

$$Z_i = \frac{1}{N_j - 1} \sum_{s \in \{J\}, s \neq i} (A_s) \tag{19}$$

which we then de-mean by hearing office and day, constructing $\widetilde{Z}_i$

As an alternative to this instrument, we also use Doyle Jr (2007) estimation procedure, also used in French and Song (2014), described below. This instrumental variable (which we term the judge allowance differential), is:

$$\widetilde{Z}_i^2 = \frac{1}{N_j - 1} \sum_{s \in \{J\}, s \neq i} (A_s - \overline{A_s}) \tag{20}$$

where $\overline{A_s}$ is the mean allowance rate by ALJs at case s's hearing office on the day case s was heard. This instrument is equivalent to the predicted allowance rate from OLS estimation of equation (1) where A_{i1} (the ALJ decision) is the dependent variable, controlling for a full set of hearing office$\times$ time interactions, and leaving observation i out, as in a jackknife estimator.

The instrument is $j_i\hat{\gamma}_1$ from the equation

$$A_i = j_i\hat{\gamma}_1 + X_i\delta_A + e_i \tag{21}$$

which implies

$$E[A_s|X_s] = E[j_s\hat{\gamma}|X_s] + X_s\delta_A \tag{22}$$

for any given s and so

$$E[j_s\hat{\gamma} - E[j_s\hat{\gamma}|X_s]] = E[A_s - E[A_{s1}|X_s]] \tag{23}$$

where the left-hand side object is $E[j_s\hat{\gamma} - E[j_s\hat{\gamma}|X_s]]$, the de-meaned instrumental variable. We approximate the right-hand side object, but using the sample analog and leaving observation i out, as in a jackknife estimator, so the constructed instrument is:

$$\widetilde{j_i\hat{\gamma}_{-i}} = \frac{1}{N_j - 1} \sum_{s\in\{J\}, s\neq i} (A_s - \overline{A_s}) \tag{24}$$

where N_j is the number of cases heard by judge j_i over the sample period, $\{J\}$ is the set of cases heard by judge j_i, $\overline{A_s}$ is the mean allowance rate by ALJs at case s's hearing office on the day case s was heard.

We then estimate equations (25) and (26):

$$\begin{aligned} \widetilde{A}_i &= \lambda(\widetilde{j_i\hat{\gamma}_{-i}}) + \eta_i, && (25) \\ \widetilde{y}_{it} &= \varphi_t(\widetilde{\widehat{\widetilde{A}}_i}) + \mu_{it} && (26) \end{aligned}$$

where "$\sim$" represents a de-meaned variable, e.g., $\widetilde{A}_{it} = A_{it} - \bar{A}_{it}$ and $\bar{A}_{it}$ is the mean allowance rate at the hearing office and on the day that case i was assigned and $\widetilde{j_i\hat{\gamma}_{1,-i}} = j_i\hat{\gamma}_{1,-i} - \overline{j_i\hat{\gamma}_{-i}}$ and $\overline{j_i\hat{\gamma}_{-i}}$ is the mean value of $j_i\hat{\gamma}_{1,-i}$ at the hearing office and on the day that case i was assigned. Because we remove case i from $\widetilde{j_i\hat{\gamma}_{-i}}$, as in a jackknife estimator, it should be independent of η_i and μ_{it}, even in a small sample. Based on Monte Carlo experiments with what seemed reasonable parameters, the procedure produced accurate approximations.

C.3 Econometric Procedures to Address Missing Mortality Information

In section 5.1 we showed that about 98.5% of deaths among those ages 55-64 are captured in the SSA mortality data and described some of the reasons for this discrepancy. Nevertheless, there is likely an under-count of those that die in our sample. Furthermore, this under-count is unlikely to be random. Because the SSA has a less of a financial incentive to measure deaths of non-DI/SSI recipients than non-recipients, the SSA data likely captures more than 98% of the deaths of those receiving benefits, but less than 98% of those not receiving benefits. This may make it look like non-beneficiaries are less likely to die than they are, and thus might lead us to infer that benefits do not reduce mortality when in fact they do. The larger the discrepancy between underreporting of beneficiaries relative to non-beneficiaries, the greater the potential bias in our estimates.

We assess how serious this problem is for our estimates. To construct the most extreme case, we assume that all deaths of beneficiaries are measured, but only a fraction p of non-beneficiaries' deaths that are measured. Define individual i's measured mortality at time τ as $y^*_{i\tau}$. Given the undercount of mortality amongst those denied, this will be

$$y^*_{i\tau} = \begin{cases} y_{i\tau} & \text{if} \quad A_i = 1 \\ y_{i\tau} & \text{with probability } p \text{ if} \quad A_i = 0 \\ 0 & \text{with probability 1-}p \text{ if} \quad A_i = 0 \end{cases} \tag{27}$$

where by assumption the probability p is independent of any of the variables that determine mortality. To address with problem we create the variable

$$\tilde{y}_{i\tau} = \begin{cases} y^*_{i\tau} & \text{if} \quad A_i = 1 \\ \frac{1}{p} y^*_{i\tau} & \text{if} \quad A_i = 0 \end{cases} \tag{28}$$

Writing our new variable this way, suppose that the true model is a modified version of equation (2)

$$y_{i\tau} = A_i \phi + X_i \delta_{y\tau} + u_{i\tau} \tag{29}$$

so that the coefficient on allowance is common to everyone. Using the adjusted mortality measure $\tilde{y}_{i\tau}$ the OLS estimate in equation identifies the conditional expectation of $\tilde{y}_{i\tau}$ given X_i

$$\begin{aligned} \mathbb{E}[\tilde{y}_{i\tau}|A_i, X_i] &= \mathbb{E}[\tilde{y}_{i\tau}|A_i = 1, X_i] \Pr(A = 1|X) + \mathbb{E}[\tilde{y}|A = 0, X] \Pr(A = 0|X) \\ &= \mathbb{E}[y_{i\tau}|A_i = 1, X_i] \Pr(A = 1|X) + \mathbb{E}[\frac{1}{p} y^*|A = 0, X] \Pr(A = 0|X) \\ &= \mathbb{E}[y_{i\tau}|A_i = 1, X_i] \Pr(A = 1|X) + \mathbb{E}[y|A = 0, X] \Pr(A = 0|X) \qquad (30) \\ &= \mathbb{E}[y_{i\tau}|A_i, X_i] \qquad (31) \end{aligned}$$

which is the conditional expectation of $y_{i\tau}$ given X_i, which is what OLS recovers.

Defining the instrument we use as Z_i, instrumental variables, using $\tilde{y}_{i\tau}$ as the left hand side variable, the conditional expectation of $\tilde{y}_{i\tau}$ given Z_i, X_i is

$$\begin{aligned} \mathbb{E}[\tilde{y}|Z, X] &= \mathbb{E}[\tilde{y}|A = 1, Z, X] \Pr(A = 1|Z, X) + \mathbb{E}[\tilde{y}|A = 0, Z, X] \Pr(A = 0|Z, X) \\ &= \mathbb{E}[\phi + X_i \delta_{y\tau} + u_{i\tau}|A = 1, Z, X] \Pr(A = 1|Z, X) \\ &+ \mathbb{E}[\frac{1}{p} p (X_i \delta_{y\tau} + u_{i\tau}) |A = 0, Z, X] \Pr(A = 0|Z, X) \\ &= (\phi + X_i \delta_{y\tau}) \Pr(A = 1|Z, X) + (X_i \delta_{y\tau}) \Pr(A = 0|Z, X) \qquad (32) \\ &+ \mathbb{E}[u_{i\tau}|A = 1, Z, X] \Pr(A = 1|Z, X) + \mathbb{E}[u_{i\tau}|A = 0, Z, X] \Pr(A = 0|Z, X) \end{aligned}$$

Note that, by the Law of Iterated Expectations,

$\mathbb{E}[u_{i\tau}|Z, X] = \mathbb{E}[u_{i\tau}|A = 1, Z, X] \Pr(A = 1|Z, X) + \mathbb{E}[u_{i\tau}|A = 0, Z, X] \Pr(A = 0|Z, X)$

and also recall the usual IV assumption that $\mathbb{E}[u_{i\tau}|Z, X] = 0$. Thus

$$\mathbb{E}[\tilde{y}|Z, X] = (\phi + X_i\delta_{y\tau}) \Pr(A = 1|Z, X) + (\phi + X_i\delta_{y\tau}) \Pr(A = 0|Z, X) = (X_i\delta_{y\tau}) + \phi Pr[D = 1|Z, X]$$

and $\mathbb{E}[A|Z, X] = \Pr[A = 1|Z, X]$. Likewise, the conditional expectation of $y_{i\tau}$ given Z_i can be derived using the same formula as in equation (32):

$$\mathbb{E}[y|Z, X] = (X_i\delta_{y\tau}) + \ \phi Pr[A = 1|Z, X]$$

and so the IV estimator using $\tilde{y}$ as the left hand side variable should (asymptotically) yield the same values as the IV estimator using y as the left hand side variable:

$$\frac{\mathbb{E}[y|Z, X]}{\mathbb{E}[A|Z, X]} = \frac{\mathbb{E}[\tilde{y}|Z, X]}{\mathbb{E}[A|Z, X]} = \frac{(X_i\delta_{y\tau}) + \phi Pr[A = 1|Z, X]}{Pr[A = 1|Z, X]}$$

which is the standard formula for a IV estimator with binary endogenous variable [need to check this].

Next, we describe how to measure p. Using the Law of Total Probability, the assumption $\Pr(y^*_{i\tau} = 1|y_{i\tau} = 1, A_i = 1)$ and the definition $p \equiv (y^*_{i\tau} = 1|y_{i\tau} = 1, A_i = 0)$ we get:

$$\begin{aligned} \Pr(y^*_{i\tau} = 1|y_{i\tau} = 1) &= \Pr(y^*_{i\tau} = 1|y_{i\tau} = 1, A_i = 1) \Pr(A_i = 1|y_{i\tau} = 1) \\ &\quad + (y^*_{i\tau} = 1|y_{i\tau} = 1, A_i = 0) \Pr(A_i = 0|y_{i\tau} = 1) \\ &= \Pr(A_i = 1|y_{i\tau} = 1) + p \Pr(A_i = 0|y_{i\tau} = 1) \end{aligned} \tag{33}$$

Using Bayes rule we know that:

$$\Pr(A_i = 1|y_{i\tau} = 1) = \frac{\Pr(y_{i\tau} = 1|A_i = 1) \Pr(A_i = 1)}{\Pr(y_i = 1)}, \tag{34}$$

$$\Pr(A_i = 0|y_{i\tau} = 1) = \frac{\Pr(y_{i\tau} = 1|A_i = 0) \Pr(A_i = 0)}{\Pr(y_i = 1)}, \tag{35}$$

Combining equations (33)- (35) yields

$$p = \frac{[\Pr(y^*_{i\tau} = 1|y_{i\tau} = 1) \Pr(y_i = 1)] - [\Pr(y_{i\tau} = 1|A_i = 1) \Pr(A_i = 1)]}{\Pr(y_{i\tau} = 1|A_i = 0) \Pr(A_i = 0)}. \tag{36}$$

Using the Law of Total Probability and straightforward algebra shows that

$$\Pr(y_{i\tau} = 1|A_i = 0) = \frac{\Pr(y_{i\tau} = 1) - \Pr(y_{i\tau} = 1|A_i = 1) \Pr(A_i = 1)}{\Pr(A_i = 0)} \tag{37}$$

Combining equations (36) and (37) yields:

$$p = \frac{[\Pr(y^*_{i\tau} = 1|y_{i\tau} = 1) \Pr(y_i = 1)] - [\Pr(y_{i\tau} = 1|A_i = 1) \Pr(A_i = 1)]}{[\Pr(y_i = 1)] - [\Pr(y_{i\tau} = 1|A_i = 1) \Pr(A_i = 1)]}. \tag{38}$$

Since, using the definition of a joint probability and the fact that anytime a death is observed in the SSA data are also observed in the National Death Index, $[\Pr(y^*_{i\tau} = 1|y_{i\tau} = 1)\Pr(y_i = 1)] = [\Pr(y^*_{i\tau} = 1, y_{i\tau} = 1)] = \Pr(y^*_{i\tau} = 1)$. Thus equation (38) can be rewritten as

$$p = \frac{[\Pr(y^*_{i\tau} = 1)] - [\Pr(y_{i\tau} = 1|A_i = 1)\Pr(A_i = 1)]}{[\Pr(y_i = 1)] - [\Pr(y_{i\tau} = 1|A_i = 1)\Pr(A_i = 1)]}. \tag{39}$$

Assuming that $\widehat{\Pr(y^*_{i\tau} = 1)}$ = #of deaths in the SSA data/population and $\widehat{\Pr(y_{i\tau} = 1)}$ = #of deaths in the NDI data/population, equation 40 can be estimated as

$$\begin{aligned} p &= \frac{\text{\#of deaths in the SSA data/population} - \text{\#of deaths of beneficiaries in SSA data/population}}{\text{\#of deaths in the NDI data/population} - \text{\#of deaths of beneficiaries in SSA data/population}} \\ &= \frac{\text{\#of deaths in the SSA data} - \text{\#of deaths of beneficiaries in SSA data}}{\text{\#of deaths in the NDI data} - \text{\#of deaths of beneficiaries in SSA data}}. \end{aligned} \tag{40}$$

We can estimate all the probabilities in equation (40). For those ages 55-64, $\Pr(y^*_{i\tau} = 1|y_{i\tau} = 1) = .98$ as we calculated previously, $\Pr(y_i = 1)$ is the annual mortality rate of all members in this age group, which we take from aggregate life tables, $\Pr(y_{i\tau} = 1|A_i = 1)$ we calculate from internal Social Security Administration documents. adjudication We calculate $\Pr(A_i = 1)$, the probability of receiving benefits, again using Social Security Administration data.

D Calculations of the Impact of ALJ Allowance on Subsequent Allowance, Income, and Benefits

In section 7 we present evidence on how receipt of DI benefits affects labor supply, earnings, health insurance, and the dollar value of those health care benefits. In this appendix we further document the calculations in that section.

D.1 Allowance

Many denied applicants continue to appeal and reapply for benefits until they are allowed. Figure 1 shows that 35% of all applicants denied by an ALJ were allowed benefits within three years. French and Song (2014) show both IV and OLS estimates of subsequent allowance rates, where the IV estimates use our judge leniency instrument. The difference between OLS and IV estimates are that the IV estimates measure the subsequent allowance rates for the marginal individual, whereas the OLS estimates measures subsequent allowance rates for the average. French and Song find that IV estimates are slightly higher than OLS. For example, the IV estimate of allowance is 42% three years after assignment, versus 35% from the OLS estimates. This finding is consistent with the view that those affected by the instrument are likely the marginal cases who have a better chance of final allowance than others denied benefits.

D.2 Income Benefits

If an applicant is allowed DI benefits, the dollar amount of benefits depends on previous labor earnings. Disabled worker benefits averaged $1,004 per month among DI beneficiaries in 2007

(U.S. Social Security Administration (2008)). Because the benefit schedule is progressive, disability benefits replace 60% and 40% of labor income for those at the 10th and 50th percentile of the earnings distribution, respectively (Autor and Duggan (2006)). Those receiving benefits can earn up to the Substantial Gainful Activity level (SGA), which was $500 per month (in current dollars) during the 1990s and $900 per month in 2007. Those earning more than this amount for more than a nine month Trial Work Period lose their benefits. Disabled individuals with especially weak earnings histories and low asset levels are eligible for a related program called Supplemental Security Income (SSI). SSI benefits are not a function of previous labor income. The Federal Maximum SSI benefit level was $386 per month in 1990 and $623 in 2007. Some states supplement this benefit. Benefits are reduced by 50 cents for every dollar of labor income. Many people draw both DI and SSI benefits concurrently. We take DI/SSI benefit calculations from French and Song (2014), which use the distribution of post-tax wages plus DI/SSI benefits for everyone in our data using the federal, state, and local tax schedule shown in French and Jones (2011). Detailed information on earnings histories and state of residence allow for accurate measurement of individual benefits. Our main limitation on these measurements is that ideally we should know family structure and all sources of income to calculate taxes. Unfortunately, we do not have this information, so we assume that the individual can claim no dependants for the DI/SSI.

D.3 Health Insurance

DI/SSI beneficiaries usually receive either Medicare or Medicaid health insurance. DI beneficiaries almost always receive Medicare benefits after a 2 year waiting period. For SSI beneficiaries, things are more complicated. If they meet certain requirements, SSI beneficiaries are immediately eligible for Medicaid. In certain states all SSI beneficiaries receive Medicaid benefits, whereas in other states the Medicaid eligibility criteria are more stringent.[19] Thus some SSI beneficiaries never get health insurance benefits. See Rupp and Riley (2012) for more information. We know whether the individual is applying for DI versus SSI benefits, and also state of residence. Thus we can exploit these variables and the estimates in Rupp and Riley (2012) in whether an individual has Medicaid and/or Medicare. They estimate the share of DI and SSI beneficiaries with Medicaid or Medicare benefits at different points in time.

D.4 Employment and Earnings

Both income effects (through the high replacement rate) and substitution effects (beneficiaries will lose benefits if they earn above the SGA amount) causes DI recipients to reduce labor supply. Likewise, the income benefit and the clawback of benefits for SSI beneficiaries also causes SSI beneficiaries to reduce labor supply. Furthermore, DI/SSI benefits likely reduce labor supply through a third channel – health insurance eligibility. Medicare and

[19] 32 states and DC, SSI beneficiaries are automatically eligible for Medicaid. In another seven states, SSI beneficiaries are eligible for Medicaid but must file a separate application. The remaining states have rules for Medicaid eligibility that differ from the eligibility rules for SSI.

Medicaid largely eliminate the value of employer-provided health insurance. For those working at a firms providing health insurance, the health insurance from work is potentially a powerful incentive to stay at that job. The employment and earnings losses for our sample are reported in Table 6.