Journal of Public Economics 229 (2024) 105033

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

The effect of disability insurance receipt on mortality ☆

Bernard Black [a], Eric French [b,*], Jeremy McCauley [c], Jae Song [d]

[a] Northwestern University, Pritzker School of Law and Kellogg School of Management, United States of America
[b] University of Cambridge and Institute for Fiscal Studies, United Kingdom
[c] University of Bristol, United Kingdom
[d] Social Security Administration, United States of America

ARTICLE INFO

JEL classification:
H51
H55
I12
I13
J14
J22

Keywords:
Disability
Benefits
Mortality

ABSTRACT

This paper estimates the effect of Disability Insurance and Supplemental Security Income receipt on mortality for individuals on the margin of being allowed versus denied benefits. Exploiting the random assignment of administrative law judges to disability insurance cases, we find that benefit allowance increases 10-year mortality rates by 2.8 percentage points for marginal beneficiaries. However, using a Marginal Treatment Effects approach, we find evidence that benefit receipt reduces mortality for inframarginal beneficiaries, who are typically less healthy than marginal beneficiaries. Furthermore, we find suggestive evidence that allowance reduces mortality among those with expensive health conditions such as cancer.

1. Introduction

Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI) for the disabled are among America's largest social insurance programs. Over 6% of the working age population now receive either SSDI or SSI benefits.[1] Most SSDI and SSI beneficiaries also receive health insurance benefits through Medicare (for SSDI beneficiaries) or Medicaid (for SSI beneficiaries). The total amount spent on both the cash and health insurance benefits makes these programs several times more expensive than unemployment insurance. These programs have grown rapidly over time, in part due to changes in the disability criteria (Autor and Duggan, 2006). Given the increasing cost of these programs, many reform proposals have been put forward, including making the disability criteria more stringent (Waddell, 2020).

This paper estimates the effect of SSDI and SSI benefits on mortality for those who appealed an initial denial. Since we focus on appellants who are on the margin of being allowed or denied benefits, our findings are informative for how changes in disability criteria might impact mortality. To the best of our knowledge, ours is the first paper to study this issue. In general our results suggest that for maximizing the longevity of current appellants, the current disability thresholds are at about the right level.

We compare mortality rates of those who were allowed benefits to the mortality rates of those who were denied. Although beneficiaries receive cash transfers and health insurance benefits that likely reduce mortality, they also face work disincentives which potentially increase mortality. Identifying the overall effect on mortality is difficult, however, because those allowed benefits may be unobservably less healthy than those denied.

☆ This paper was begun when French was an IPA at SSA; French thanks the hospitality of the SSA during his stay. We are grateful for excellent research assistance from Sreevidya Ayyar, and comments from the Editor (Sarah Miller), three referees, Camille Landais, Magne Mogstad, Timothy Moore, Owen O'Donnell, Francesca Salvati, and participants from the IFS-ONS conference on administrative data; Conference on The Welfare State and the Labour Market: Micro meets Macro; UBC; Columbia Law School; ETH-Zurich; NBER Conference on Working Longer; the Labour Market with an Ageing Population Conference, Uppsala; Haifa; Erasmus-Rotterdam; Health and Labor Workshop, Nantes; Bocconi-CEPR workshop; 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.

* Corresponding author.
E-mail addresses: bblack@northwestern.edu (B. Black), eric.french.econ@gmail.com, ebf26@cam.ac.uk (E. French), jeremymccauley@gmail.com, jeremy.mccauley@bristol.ac.uk (J. McCauley), jae.song@ssa.gov (J. Song).

[1] See Section 3 for more details.

[2] We focus on this age group due to their data being particularly high quality. We show results for younger ages in the Appendix.

https://doi.org/10.1016/j.jpubeco.2023.105033
Received 12 December 2022; Received in revised form 17 November 2023; Accepted 18 November 2023

Using Social Security administrative data, we exploit the essentially random assignment of those appealing an initial denial to Administrative Law Judges (ALJs), who allow or deny benefits. We document differences in allowance rates across judges, and show that these differences are unrelated to the health or earnings potential of appellants. We use judge-specific allowance rates to construct an instrumental variable, which we call "judge leniency", and use this instrument to predict benefit allowance and its effect on mortality. We study persons aged 55–64 who were assigned to an ALJ between 1995 and 2004.[2] As our sample includes all disability appellants in this age group, we obtain precise estimates, even for relatively small subgroups of this population.

Our IV estimates show that disability benefit allowance increases the 10-year mortality rate by a statistically significant 2.8 percentage points, relative to a mean rate of 20% for those denied benefits. It is likely that cash benefits, health insurance, and work disincentives all impact mortality. As these effects differentially impact different parts of the population, benefit allowance produces heterogeneous effects, improving mortality outcomes for some and worsening them for others.

We find heterogeneous effects based on appellants' prior observable health conditions, although estimates are not precise. Benefit allowance lowers mortality among those with high mortality respiratory and nervous system conditions and cancer. Conversely, benefit allowance increases mortality for appellants with musculoskeletal disorders, which is a low average mortality condition where benefit receipt strongly reduces labor supply.

Using a Marginal Treatment Effects (MTE) approach, we identify heterogeneity in mortality responses across the judge leniency distribution. This lets us assess the mortality response amongst those who were likely in worse health in ways observable to the judge but unobservable to us. Among appellants who would only be allowed by lenient judges (and are likely healthier), benefit allowance increases mortality. However, among inframarginal appellants, who would be allowed by strict judges (and are likely less healthy), allowance reduces mortality. This suggests that inframarginal individuals, who would be allowed benefits by any judge, likely benefit from disability benefit receipt. As those allowed benefits at the initial stage are less healthy than those allowed at the ALJ stage (see, e.g., Parsons, 1991), it is likely that those receiving disability benefits at the initial stage benefit from allowance as well.

Both the results disaggregated by health condition and the MTE estimates suggest that for maximizing the longevity of appellants, the current disability thresholds are about right. Increasing allowance rates by making the threshold less strict will likely increase mortality for marginal appellants, although a significant reduction in allowance rates would likely increase mortality as well.

The rest of the paper is organized as follows. Section 2 provides a literature review. Section 3 describes the disability system in the US. Section 4 describes our estimation methods. 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 benefit 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 SSDI and SSI, relatively little research has been done on how these programs affect the health and mortality of the disabled population. Benefit allowance might impact health and mortality since receiving benefits impacts health insurance, income, and employment of beneficiaries.[3] In this section we briefly review the literature on assessing the separate effects of health insurance, income, and work on health; we return to this literature in Section 7, where we discuss how our study, which assesses the joint effects of all three, fits within this literature.

2.1. Income effect from receiving disability insurance

To our knowledge, the only other paper to estimate the effect of US disability benefits on adult mortality is Gelber et al. (2023). They estimate the effect of Disability Insurance benefit income on mortality rates by exploiting the kinks in the SSDI benefit formulas. They find evidence that higher SSDI benefit income leads to lower mortality at the lower bend point of the SSDI benefit formula, but find no robust evidence of an effect at the upper bend point. It is important to note they study a different margin than the one we study. They measure the effect of benefit generosity on mortality. In contrast, we measure the effect of being allowed benefits versus being denied, which jointly affects income, health insurance, and labor supply.

Silver and Zhang (2022) use an examiner-generosity design, similar to ours, to study the effect of receipt of Veterans Administration (VA) disability benefits for veterans with mental conditions, and estimate a precise zero effect of receiving benefits on mortality. However, unlike SSDI, the VA program does not affect eligibility for health insurance and continued receipt of benefits does not depend on future employment.

Focusing on the impact of SSI receipt for low birth weight infants, Guldi et al. (2022) exploit a discontinuity in infant eligibility to study a range of outcomes for both the infant and their parents. They find no evidence that infant mortality changes discontinuously at the SSI eligibility threshold.

2.2. Health insurance

Several important studies find that for the adult and elderly population, the near-term effect of health insurance on subsequent health outcomes is small. These studies include the RAND Health Insurance Experiment (Brook et al., 1983), and analyses of both Medicare (Finkelstein and McKnight, 2008) and Medicaid (the Oregon Health Insurance Experiment) (Finkelstein et al., 2012). Two recent studies do find that health insurance expansion under the Affordable Care Act reduces mortality (Goldin et al., 2021; Miller et al., 2021); but see the critique of these studies as underpowered by Kaestner (2021).

Chatterji and Meara (2010) study the consequences of eliminating disability benefits for substance abusers, and find no measurable effect on health insurance status or health care use, however, as the authors note, they have limited power to detect effects on these outcomes. To our knowledge, Weathers and Stegman (2012) is the only study that focuses specifically on the mortality effect of health insurance for the disabled. They exploit a randomized experiment that reduced the wait time for receiving Medicare benefits from 2 years to 0 years. They find no significant effect of immediate versus delayed receipt of health insurance on mortality, although their point estimates imply higher mortality among those who received Medicare immediately.

Most studies focus on short-run effects, have limited sample sizes, or both. Thus, it is difficult to determine whether health insurance has little effect on mortality, or if estimates are insignificant because the sample size or the sample period is too small. Black et al. (2017) study longer-term effects for the near-elderly (initial age 50–61), but also have a limited sample and use pure observational study methods, rather than a true or natural experiment; they find no significant effect of health-insurance on near-elderly mortality.

[3] The onset of disability negatively affects a wide range of economic outcomes (Meyer and Mok, 2019), although receipt of disability benefits can mitigate financial distress (Deshpande et al., 2021).

2.3. Income and employment

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

Multiple papers have found that government policies that induce early retirement lead to poorer health and higher mortality among beneficiaries. Fitzpatrick and Moore (2018) 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). Those who received higher benefits due to their year of birth had higher mortality rates and lower employment levels. They conclude that greater work effort has beneficial health impacts, which more than offset any mortality gains from greater Social Security income.[4] Several papers exploit European retirement reforms. While the evidence is mixed, several studies suggest that early retirement decreases cognitive and physical activity (Banks et al., 2019; Rohwedder and Willis, 2010). Furthermore, Kuhn et al. (2020) 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 changes in health behavior as a causal mechanism for higher mortality. Several papers using European administrative data, such as Rege et al. (2009) and Eliason and Storrie (2006), find similar results. While the aforementioned papers focus primarily on voluntary early retirement, Sullivan and von Wachter (2009) also find that involuntary job loss significantly increases mortality, potentially due to loss of health insurance and loss of income.

3. The disability insurance system

Both SSDI and disabled SSI applicants face the same medical criteria and application process. Whereas individuals with a previous strong attachment to the labor market receive SSDI benefits, individuals with low income additionally receive SSI benefits. In 2014, 6.4% of people aged 18–64, and 16.3% of those aged 55–64, were receiving either SSDI or SSI benefits (U.S. Social Security Administration, 2014a).[5] Most SSDI and SSI beneficiaries also receive health insurance benefits through Medicare (for SSDI beneficiaries) or Medicaid (for SSI beneficiaries). The combined cost of these programs was $428 billion in 2008 (Livermore et al., 2011), making them several times more expensive than unemployment insurance. Costs have risen rapidly, generating many policy proposals to reform the system (Autor and Duggan, 2010; Burkhauser and Daly, 2011; Burkhauser et al., 2014). In what follows we will collectively refer to the SSDI and SSI programs as SSDI.

3.1. Exit rates from the SSDI program

Relatively few people lose disability benefits for reasons other than death.[6] For example, of the 7.1 million individuals drawing SSDI benefits in 2007, 0.5% had benefits terminated because they earned above the Substantial Gainful Activity (SGA) limit for an extended period of time. 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 SSDI 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.

3.2. Determining eligibility for SSDI 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,[7] 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 and multiple sclerosis. These individuals are not in our sample.

The second condition is inability to work, either at their past work or other work, for at least one year or for their life expectancy, if shorter. Eligibility under this condition depends on a combination of medical impairment and vocational factors such as education, work experience, and age. These cases can be 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!"

Fig. 1 shows the percent of applicants who are allowed at different stages of the disability determination process. After an initial 5-month waiting period, applicants have their case reviewed by a Disability Determination Service review board. At this initial stage, clear-cut cases are allowed, such as those with a listed impairment, while cases that are harder to judge (such as musculoskeletal problems) are usually denied. Fig. 1(a) shows that around 38% of applicants are allowed and 62% are denied at this stage. Just under half of all applicants who are initially denied appeal at the Disability Determination Service reconsideration stage. About 16% of those appeals are allowed benefits at this stage. Rejected applicants can appeal to an Administrative Law Judge (ALJ), including those who did not seek reconsideration. A typical time from appeal to hearing is about one year.[8] Over 60% of the appealed claims

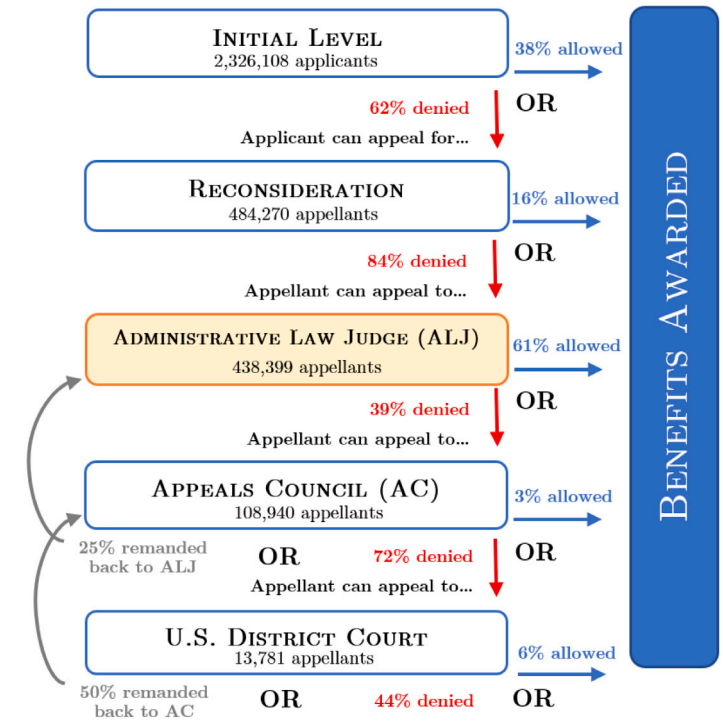
[4] Several papers also consider the timing of benefit receipt and mortality (e.g. Evans and Moore, 2011; and Dobkin and Puller, 2007). They find a jump in mortality in the days immediately following benefit receipt.

[5] 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 SSDI, 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 SSDI 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).

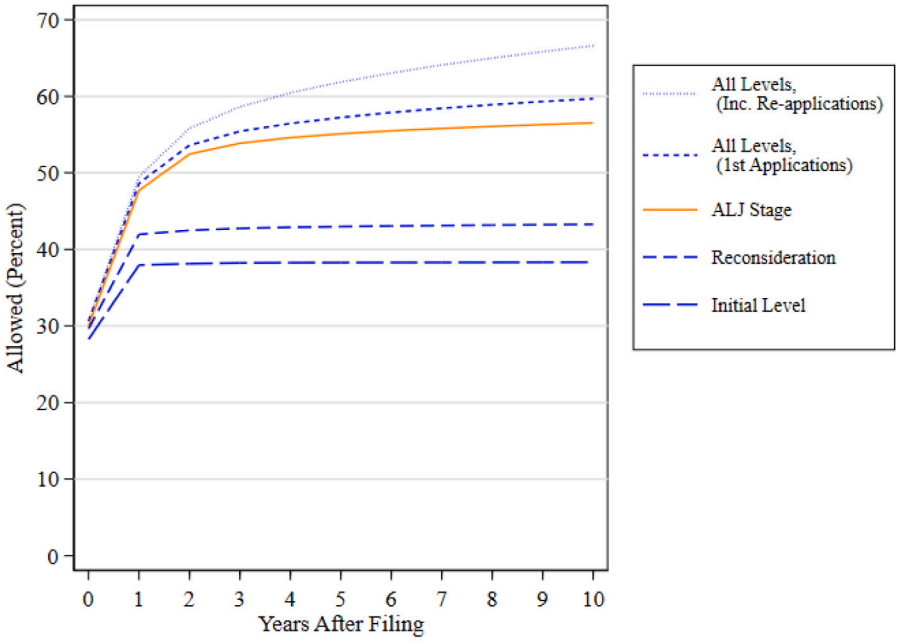
[6] SSDI benefits are converted into retiree benefits once the beneficiary turns the normal retirement age. The statistics in this sub-section are for SSDI benefits before the conversion to retiree benefits.

[7] See Hu et al. (2001) or Benitez-Silva et al. (1999) for details. The guidelines provided by SSA imply all ALJs are required to consider the same factors, including health, age, education, and work history, and use the same sequential process to evaluate these factors. If all ALJs provide the same weights of each factor (although they may have different thresholds for allowance), this implies that "strict monotonicity" as described in Section 4.3 is plausible.

[8] 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. Because we only consider allowance at the ALJ stage, we consider a case as denied even if they are eventually allowed, but only on appeal or reapplication after the ALJ stage.



(a) Application Process, All Ages, 2002



(b) Cumulative Allowance Rates, All Ages, 1995-2004

Fig. 1. Applications and appeals process of SSDI.
Notes: Panel (a) shows data for the 2002 fiscal year for applicants of all ages. The number of appellants for the U.S. District Court include those on Continuing Disability Reviews (CDRs), who are not included in the counts at other stages. For ALJ, AC, and District Court, the percent denied includes cases that are dismissed. The data in panel (a) relate to workloads processed, but not necessarily received, in fiscal year 2002, therefore stages may include cases received at one or more of the earlier stages prior to 2002. Panel (b) shows cumulative allowance rates at all stages of the application process for all new applications between 1995–2004. Those denied benefits at any stage can choose to reapply at the Initial Level instead of appealing: Panel (b) shows that 7% of all initial applicants are eventually allowed through a reapplication.

(10% of all claims) are allowed at the ALJ stage.[9] 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 to U.S. District Court. However, Fig. 1(b) shows that appeals at the higher levels are rarely successful: only about 2% of all initial claimants are allowed benefits at the Appeals Council or U.S. District Court level. Lastly, denied applicants can re-apply for benefits. The top line on Fig. 1(b) includes those who re-apply for benefits. Another 7% of all initial claims are eventually allowed benefits through a re-application. Around 33% do not get benefits at any point during the 10 years after initial application.

Because we identify the causal effect of SSDI on mortality using variation at the ALJ level, the estimated effect applies only to these marginal cases. The least healthy individuals, such as those with listed impairments, will almost always be allowed at the initial level and never reach an ALJ. The healthiest individuals will almost always be denied, whichever ALJ they see. Thus, our results are not generalizable to all SSDI 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 levels of the SSDI system.

3.3. Assignment of SSDI cases to judges

Judicial independence means that judges have a great deal of latitude to determine whether an appellant is allowed or denied (Taylor, 2007). As a result, two different judges can have very different allowance rates even though they see similar appellants.

ALJs are assigned to hearing offices, and within hearing office, are assigned cases on a rotating basis.[10] When a judge finishes a case, that judge receives the oldest pending case at his or her hearing office. Therefore, for appellants who apply at a 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, nor are they assigned cases based on their expertise. Furthermore, an appellant cannot choose an alternate judge after being assigned a judge.

In around 4% of cases, the initially assigned judge is not the judge who decides the case. In these instances, the deciding judge may not be randomly assigned.[11] Fortunately, we have information on the assigned judge in addition to the deciding judge. Since the initially assigned judge is randomly assigned, we use initial assignment to a judge as our source of exogenous variation.

As we confirm below, the assigned judge is for all practical purposes randomly assigned conditional on hearing office and day of assignment to ALJ. 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. Appellant 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 SSDI 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, which should be essentially random.

4. Estimating equations

To estimate the effect of SSDI allowance on mortality, we use a two-step procedure. In the first step we generate an instrumental variable that is a measure of judge leniency. We show this variable is correlated with the probability of allowance, but is independent of appellant health and other characteristics. In the second step we use instrumental variables procedures to estimate the effect of SSDI on mortality.

4.1. Basic specification

Our basic estimating approach is a modified instrumental variables regression. In the first stage, we estimate:

$$A_i = \mathbf{j}_i\gamma + \mathbf{X}_i\delta_A + e_i. \tag{1}$$

where A_i is a 0–1 indicator equal to 1 if individual i is allowed benefits by the ALJ, $\mathbf{j}_i$ is a set of judge indicator variables (equal to 1 if individual i's case is assigned to a given judge), $\mathbf{X}_i$ is a vector of hearing office-day indicators (equal to 1 if individual i's case is assigned on that hearing office-day pair), and γ and δ_A are parameters. 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):

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

where $y_{i\tau}$ is mortality (or another outcome variable such as earnings, participation, appeals or allowance) τ years after assignment to an ALJ. We allow for heterogeneity in the parameter $\phi_{i\tau}$ to capture heterogeneity in the effect of benefit allowance 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 Fig. 1(b). We use ALJ allowance at first hearing rather than eventual allowance, because those who die 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 initial allowance.

4.2. Estimating equations

When estimating Eq. (2) we address three issues. First, we wish to allow for heterogeneity in the parameter $\phi_{i\tau}$. Second, we have 1436 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.

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

[10] 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 remove them from our sample. We study the remaining cases where there was a hearing.

[11] 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.

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

$$Z_i = \frac{1}{N_{j_i} - 1} \sum_{s \in \{J_{j_i}\}, s \neq i} A_s \tag{3}$$

where N_{j_i} is the number of cases heard by judge j_i over the sample period, and $\{J_{j_i}\}$ 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. (2019), 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 over 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_i with the corresponding average probability for all other judges in the same office-day. We refer to our 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 small samples.[12]

Finally, we jointly estimate the equations:

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

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

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

Given the above assumptions, the estimated effect can (approximately) be interpreted as a Local Average Treatment Effect (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), French and Taber (2011), and Frandsen et al. (2023) for more details).[13] We identify this weighted average if four 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 all judges rank appellants 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. Fourth, if judges differ only in leniency then the "exclusion restriction" should hold as only the allowance decision, and the effects of that decision, should impact mortality. Sections 6.1 and 6.2 provide evidence on the extent to which these assumptions hold.

[12] De-meaning based on office-day is equivalent to regressing on the full set of office-day indicators. Because we have almost 200,000 office-day indicators, our procedure potentially suffers from finite-sample bias. To assess this problem, in our robustness checks we consider using only office-year indicators, of which we have fewer than 1000. An alternative approach is the UJIVE procedure of Kolesar (2013) which consistently handles many covariates and addresses the finite-sample bias from many instruments.

[13] Blandhol et al. (2022) show that when adding covariates to Eqs. (1) and (2), 2SLS will not deliver a weighted average of $\phi_{i\tau}$ for the individuals affected by the instrument unless the specification is fully non-parametric. While our specification is not fully non-parametric (we assume separability), by including a full set of hearing office-day indicators we attempt to minimize this problem. See also Sloczynski (2022) on the negative weights with covariates problem.

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 of judges. 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 as well as other procedures.[14] 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 Eq. (7) is the predicted value of $\widetilde{A}_i$ from Eq. (6), and K is the order of the polynomial.

As shown by Heckman et al. (2006) and French and Taber (2011), as well as in 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 (de-meaned) proneness to be given allowance, due to judge-observed health or other characteristics (e.g., prior work history and skills), of the appellants who would be allowed by a fraction a of all judges. Eq. (8) shows that MTE(a) is the mean value of $\phi_{i\tau}$ for those who would be allowed if their assigned judge allowed a share a of cases, but denied if assigned to a slightly stricter judge. As a increases, so does the leniency of the marginal judge. As Frandsen et al. (2023) show, recovery of MTEs hinges upon "strict monotonicity", although identification of the weighted average of treatment effects in Eq. (5) only relies on "average monotonicity". Appendix section C.1 provides more details on interpretation and estimation of the MTE.

5. Data

Our data consists of all individuals aged 25–64 who appealed either a SSDI or SSI initial benefit denial, and were assigned to an ALJ during 1995–2004. Using Social Security Numbers, we match 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 runs 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 observations with missing education information. Table A1 in Appendix A.1 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

[14] 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.

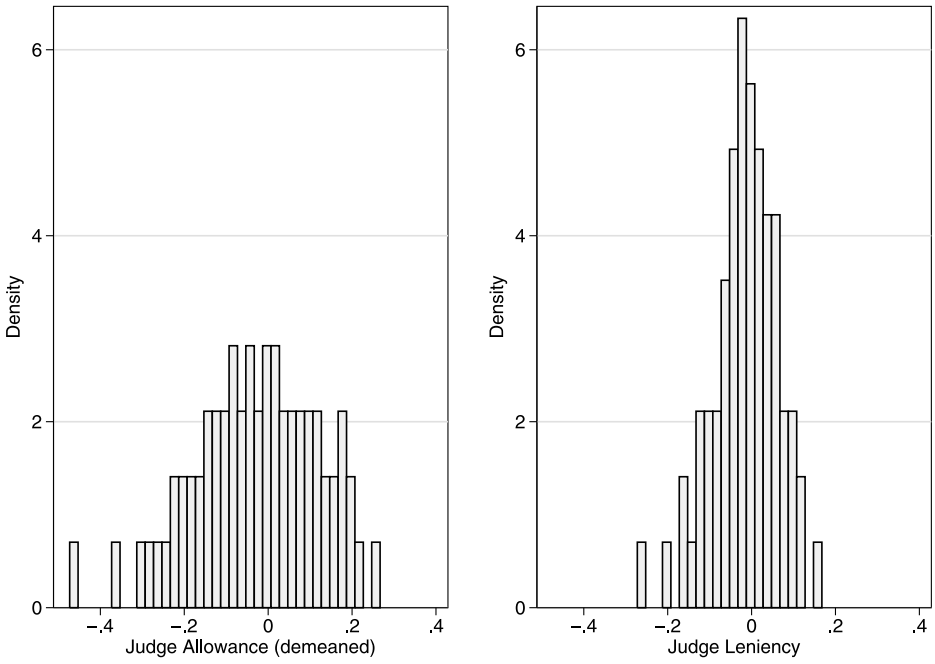


Fig. 2. Allowance rate of ALJs and Judge leniency.
Notes: The left panel displays the de-meaned allowance rates of ALJs. The right panel displays the Judge Leniency variable, which is de-meaned by hearing office and day.

individual died before their case was heard. In addition, to address any mismeasurement in whether a case was heard before death, we also drop 30,807 cases where the individual died in the year of assignment to an ALJ. This selection decision has only a modest effect on our estimates, as is shown by robustness checks in Section 8. Our full estimation sample has 2,759,907 SSDI or SSI cases heard by 1,436 judges, with a mean allowance rate at the ALJ stage of 70.8%. Our main estimation subsample of those aged 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.

Fig. 2 plots the distribution of unconditional judge-specific allowance rates (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 0.149, but the standard deviation of the judge leniency variable is 0.096 (weighted by the number of cases handled by each judge). This means that a judge one standard deviation more lenient has a probability of allowance 9.6 percentage points higher than that of their office-day average.

5.1. SSA mortality data

Our mortality data come from the SSA's NUMIDENT file. The SSA use these data to process SSDI, SSI, and Social Security benefits. The SSA obtains death records from various sources, including states, family members, funeral directors, post offices, financial institutions, and other federal agencies. The SSA has a financial incentive to record deaths, especially if it is paying benefits to an individual.

One potential concern is that because the SSA has a greater financial incentive to record deaths of beneficiaries than non-beneficiaries, it will be more likely to record their deaths, or at least record them more quickly (Autor et al., 2017). Any underreporting of deaths of those denied benefits will bias down their estimated mortality, making it appear that receiving benefits causes higher mortality. Previous studies have shown that older versions of the SSA mortality data understate mortality relative to the National Death Index (NDI) data from the Centers for Disease Control and Prevention (CDC), which is considered the "gold standard" of US mortality data (Hill and Rosenwaike, 2002).[15]

As we show below, however, underreporting of deaths is not an issue in recent years, especially for those over 55. Finlay and Genadek (2021) show that, due to recent improvements in data collection including retroactive updating for prior years, the SSA NUMIDENT data is of high quality for deaths post-1990. Here we perform a similar exercise to Finlay and Genadek (2021) but extend their analysis to adjust for differences in sample frame of the two datasets as in Black et al. (2022).[16]

The top panel of Table 1 shows the ratio of deaths recorded in the SSA data to those in the NDI data by age and year group. The table confirms that the SSA data during our sample period are of high quality, especially for those 55 and older. Given that the ratio of SSA/NDI deaths is lower for those under 55, we restrict our sample to those who appealed between the ages of 55–64 in the remainder of the paper. In the appendices we provide estimates for younger ages, which provide results consistent with those in the main (55 and older) sample. We provide additional results on the quality of the SSA mortality data in Appendix Table A2 and Appendix A.3 and perform robustness checks to allow for the possibility that under-reporting is twice as large as we estimate.

[15] The NDI is maintained by CDC and is used to construct the Vital Statistics data for the US.

[16] The NDI records deaths of foreigners living in the US whereas the SSA only records deaths of those with Social Security numbers, largely excluding foreigners. Likewise, the SSA records deaths of those with Social Security numbers living abroad whereas the NDI only record deaths in the US. Thus, we adjust the NDI data to include deaths of Americans living abroad (in US territories) and exclude deaths of foreigners who died in the US.

Table 1
Estimated percentage of U.S. deaths included in the SSA death data and underreporting correction, by age group.

	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	0.984	0.948	0.973	0.985	0.988
Estimated Fraction 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 des 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 Eq. (9).

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, which is the probability that a denied individual's death is observed. 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 as there are other reasons why SSA may count fewer deaths than NDI. However, making this assumption gives a sense of how important underreporting among those denied could be for our results.[17] This is likely a worst case bound since if both those allowed and those denied had the same underreporting probabilities then the bias would only come from usual attenuation bias. In Appendix C.3 we show that p can be estimated as

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

Each individual in our sample is assigned a value of p based on their age and year of application which we denote p_i.[18] This approach allows us to capture the higher quality of the mortality data at older ages and in more recent years.

In Appendix C.3 we show how we use 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 $\frac{1}{p_i}$ and using the estimation procedures shown in Section 4.

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. Fig. 3 shows cumulative mortality rates for those assigned to an ALJ. For those aged 55–64 at the time of assignment to an ALJ, the cumulative mortality rates in the following year are 1.3% for those denied, versus 1.7% for those allowed. 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 addresses this issue. The mortality rates for those denied, with and without the correction for underreporting described in the previous section, can been seen in the figure. The underreporting correction has only a modest effect on our estimates: the estimated difference in 10 year cumulative mortality rates for appellants aged 55–64 between those allowed and denied falls from 2.0% to 1.8%.

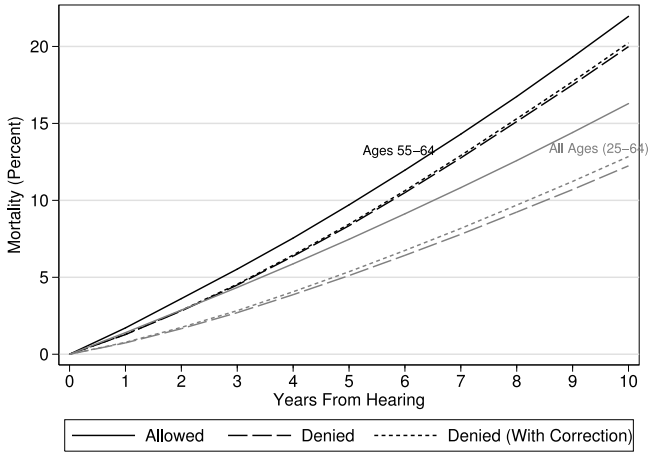


Fig. 3. Cumulative mortality rates, allowed versus denied.
Notes: Cumulative mortality rates for appellants aged 55–64, and aged 25–64, at time of assignment with separate mortality rate curves for those allowed benefits, those denied benefits, and those denied benefits with correction for underreporting of mortality.

Our estimated mortality rates are lower than (Parsons, 1991). He reports a six-year mortality rate of 12.9% for those denied versus 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. For individuals aged 55–64, mortality rates are: 10.0% for those denied versus 11.1% for those allowed. Our estimates are likely lower because

[17] 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, unauthorized immigrants who lack a Social Security number should still 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.

[18] In practice we first calculate $p_{i,t}$ for each year for the following age groups: 25–44, 45–54 and 55–64. We assume that there is no under-count after 65 since those denied will draw Social Security benefits after this age. Using these values, we then calculate p_i as the average of $p_{i,t}$ for each individual over the period we observe their mortality (5 or 10 years).

Parsons' cohort is from 1970, whereas ours is from 1995–2004. We also find a much smaller gap than Parsons between mortality for those allowed and denied. More recent SSDI beneficiaries tend to be healthier than older ones and have primary diagnoses less related to mortality, as shown in Autor and Duggan (2006).

We estimate the effect of benefit receipt on mortality for those who are initially denied and appeal at the ALJ stage. Parsons (1991) shows that average mortality rates of those allowed at the initial stage are much higher than mortality rates of those who are allowed at the ALJ stage.[19] This highlights that appellants who reach the ALJ stage are likely healthier than those allowed at the initial stage (but may be less healthy than individuals denied at the initial stage, some of whom do not appeal to an ALJ). This can also be seen by comparing our sample with that in Gelber et al. (2023), whose sample consists of allowed individuals aged 21–61 (with 72% of their sample allowed before the ALJ stage). After 1 year, the mortality rate for their sample is 7.0%, versus 1.6% for our main sample (aged 55–64) and 1.2% for our full sample (aged 25–64).[20]

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 appellants who appear before that judge. Table 2 presents a test of this conditional random assignment hypothesis for individuals ages 55–64. For similar tests on the full sample, see Table A4 in the appendix.

First, we consider which variables predict allowance. Column 1 of Table 2 presents estimates from a multivariate regression of 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 all more likely to be allowed benefits. Column 2 shows 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.

Column 3 presents estimates from a regression of judge leniency $\tilde{Z}_i$ on the same covariates. Column 4 provides $t-statistics$ with standard errors clustered by judge.[21] 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 ages 25–64, we again only find one statistically significant covariate (see Table A4). All the estimated coefficients are small in comparison to the coefficients on their counterparts in the allowance equation, suggesting that our inability to reject the null of 0 coefficient is not the result of low power of these tests. 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. The next section provides some evidence on whether rank (instrument strength) and monotonicity conditions hold.

6.2. First stage: The effect of judge leniency on allowance

Table 3 presents estimates of the effect of judge leniency on allowance rates for the main estimation sample. These can be interpreted as tests of "average monotonicity" as described in Frandsen et al. (2023). 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.[22] In terms of 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 intellectual disabilities 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 Eq. (4). The estimated value of $\hat{\lambda}$ for the main estimation sample is 0.68, meaning that the probability that case i is allowed at assignment rises 0.68 percentage points for every 1 percentage point increase in judge leniency. The judge leniency instrument is calculated on the full sample (all ages) and has a standard deviation of 0.096, meaning that a judge who is estimated to be one standard deviation more lenient than average in the full sample on average allows 9.6% more cases for the full sample. Since our main sample are aged 55–64, and these individuals are more likely to be allowed by all judges, the judge leniency instrument predict a smaller change in allowance percentage for this group. The coefficient on the instrument when predicting allowance in Table 3 is 0.68; therefore a judge who is one standard deviation more lenient in the full sample has a predicted allowance rate that is $0.096 \times 0.68 = 0.065$ (6.5%) higher for new appellants aged 55–64. The estimates of $\hat{\lambda}$ are highly statistically significant for all subgroups. 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. For all subgroups, the t-statistic is well over the minimum level recommended by Lee et al. (2022) for unbiased standard error estimates.

Abadie (2003) shows that the ratio of the group specific estimate of $\hat{\lambda}$ to the full sample estimate of $\hat{\lambda}$ (this ratio is shown in Column 6) is informative of 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 prior participation and earnings, those not represented by a lawyer, and those with mental disorders more than for other groups, holding the appellant 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 appellants.

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, since all t-statistics are

[19] Parsons (1991) finds that for all applicants at the initial stage in 1970, the mortality rate is 27.4% within 6 years (37.3% for those allowed). However, for those initially denied who appeal at the ALJ stage, the mortality rate is only 15.2% (17.5% for those allowed).

[20] Our sample also differs on other observables: notably fewer high mortality conditions (e.g., cancer), and more lower mortality conditions (e.g., musculoskeletal conditions). See Appendix Table A3 for more comparisons.

[21] We follow common practice in the judge fixed effects literature (e.g., Frandsen et al., 2023) and cluster at the level of a judge. Abadie et al. (2023) recommend clustering at the level of treatment assignment, which would suggest no clustering, but it is unclear how to implement this advice for a continuous instrument. Standard errors are less than half as large when not clustering, therefore throughout this paper we go with the more conservative approach used in most of the literature and cluster by judge.

[22] 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. However, Hoynes et al. (2022) find that representation ultimately leads to no detectable change in the probability of final allowance.

Table 2
Predictors of allowance and judge leniency, aged 55–64.

Covariate	Dependent variable: Allowed		Dependent variable: Judge leniency	
	Coefficient	T-Stat	Coefficient	T-Stat
	(1)	(2)	(3)	(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
Intellectual disability	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	
F−statistic	64.5		5.4	
R^2	0.0127		0.0022	
Number of Appellants = 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, age 60–64, white, not represented by a lawyer, applying for SSI or SSI and SSDI concurrently, not a high school graduate, with a health condition other than those listed above. The sample includes appellants aged 55 to 64, and we de appellants who died the year of application. Standard errors are clustered by judge.

large, instrument strength (the rank condition) is not in question. Table A5 in the Appendix presents the same estimates for all ages.

6.3. Second stage: The effect of benefit receipt on mortality

Table 4 presents estimates of the effect of disability benefit allowance on mortality 5 and 10 years after assignment to an ALJ, for our main estimation sample. The bottom row shows that 19.99% of those denied benefits die within 10 years and the first row displays the coefficient of allowance on mortality, 1.97%, which is the difference in mortality between those denied and allowed benefits. Therefore 21.95% of those allowed benefits die within 10 years, which suggests that those allowed benefits are more likely to die. The coefficient on allowance is highly statistically significant. However, these are simple OLS estimates, without covariates, which do not address selection effects and do not provide causal estimates.

The next rows show the OLS and IV estimates of de-meaned (by hearing office and day) mortality on similarly de-meaned allowance, as well as the associated standard error. De-meaning the data has very little effect on the OLS estimates. The IV estimates show that being allowed benefits increases the 5 year and 10 year mortality rate by 1.81 and 1.93 percentage points, respectively. Surprisingly, the IV estimates are close to the OLS estimates.

Despite the apparent similarity of the OLS and IV estimates, it is important to remember that they identify different objects. The OLS estimate is for the full sample and potentially suffers from selection bias. Our IV estimates, in contrast, are based on the random assignment of judges, and thus identifies the average effect of allowance on mortality, but only for those on the margin of being allowed or denied depending on the judge they are assigned.

Next, we provide OLS and IV estimates which include the covariates listed in Table 2. Adding covariates to this specification increases both the OLS and IV estimates. The two estimates remain similar, and the IV estimates remain strongly statistically significant at both 5 and 10 years. Surprisingly, adding covariates increases the estimated OLS coefficient on benefit allowance. Some groups with higher mortality (shown in Table 5) such as those with cancer and older individuals have higher allowance rates (shown in Table 3). For these groups, the OLS estimate is closer to 0. However, other groups with higher mortality rates, such as black individuals and those with low prior earnings, have lower allowance rates. These groups have larger OLS estimates. On

Table 3
First stage estimates: Regression of allowance rates on judge leniency variable, by demographics, aged 55–64.

	Obs.	Allowance rate at ALJ stage	Coefficient on judge leniency	Std. error	T-Stat	Relative likelihood[a]
	(1)	(2)	(3)	(4)	(5)	(6)
All groups						
All persons age 55–64	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
Intellectual disability	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 are clustered by judge.
[a] Relative likelihood is the ratio of the group specific coefficient on judge leniency (presented in column 3) to the full sample coefficient.

net, conditioning on all our observables somewhat increases the OLS estimates. This means that it is not obvious whether conditioning on unobservables should lead OLS to over- or under-estimate the impact of allowance on mortality. The final rows in each panel of Table 4 display the estimates with covariates, after including the underreporting correction described in Section 5.3. The estimates fall, but only slightly. Table A6 in the Appendix presents the same estimates for different age groups.

6.4. Heterogeneity in the mortality effect based on judge leniency: marginal treatment effects

Fig. 4 presents how the MTE (i.e., the mortality response for the marginal case allowed) varies with predicted de-meaned allowance for our main sample, where the prediction is based on the leniency of the judge who heard a specific case. Subfigure (a) displays 5 year mortality, and subfigure (b) displays 10 year mortality, both controlling for covariates. See Section 4.3 and Appendix C.1 for details on constructing MTEs.

We use third order polynomials for both the instrument and the endogenous variable (de-meaned allowance) when estimating Eqs. (6) and (7), although visual inspection of Fig. 4 as well as both the Akaike and Bayesian information criterion show that there is little gain from going beyond a quadratic specification. In Appendix B.3, we show that these results change only modestly when adding covariates or when using a local polynomial smoother as in 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. Fig. 4 also shows bootstrapped 95% confidence intervals.

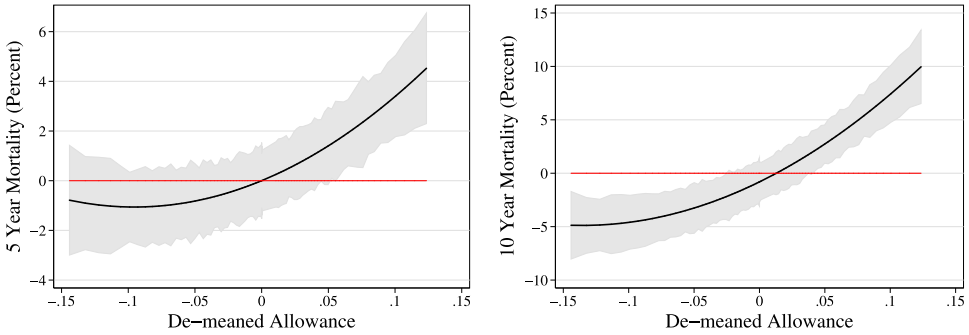
Fig. 4 shows the estimated MTE is close to zero at the average allowance rate, at both 5 and 10 years, but there is heterogeneity in the responses. Allowance reduces 5 year mortality by an estimated 0.8 percentage points for the marginal appellant heard by an ALJ who is stricter than 95% of all judges. These judges have allowance rates that are twelve percentage points below the average (de-meaned by hearing office and day). However, allowance increases 5 year mortality by 4.5

Table 4
Estimated effect of SSDI recipiency on mortality, Aged 55–64.

	Mortality (Percent)			
	5 years		10 years	
	OLS	IV	OLS	IV
Without Covariates:				
Coef on allowance	1.35		1.97	
(Std. Error)	(0.11)		(0.19)	
Coef on de-meaned allowance[a]	1.35	1.81	1.87	1.93
(Std. Error)	(0.12)	(0.44)	(0.19)	(0.76)
With Covariates:				
Coef on de-meaned allowance[a]	1.94	2.30	2.77	2.81
(Std. Error)	(0.12)	(0.50)	(0.18)	(0.91)
With Covariates and Underreporting Correction:				
Coef on de-meaned allowance[a]	1.76	2.12	2.51	2.54
(Std. Error)	(0.12)	(0.50)	(0.18)	(0.90)
Mean mortality for those denied	8.35		19.99	

Notes: Sample is based on 610,231 appellants aged 55–64. Instrument is judge leniency. Covariates are those in Table 3; 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 are clustered by judge.

[a] For de-meaned allowance, all variables are de-meaned from the hearing office-day average.



(a) 5 Year Mortality Response (b) 10 Year Mortality Response

Fig. 4. Marginal treatment effects, Aged 55–64.
Notes: This figure displays the estimated mortality response as a function of predicted de-meaned allowance, where the prediction is based on the leniency of the judge who heard a specific case. We control for the covariates listed in Table 2.

percentage points for the marginal appellant 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 is qualitatively similar to the 5 year response, although the magnitudes with covariates, perhaps reflecting that the impacts of allowance have more years to accumulate.

In Appendix C.1 we display estimates for the other age groups, which follow the same basic pattern as for those aged 55–64: greater leniency implying higher mortality for the marginal appellant.

Our results suggest that increasing the allowance rate, by making the SSDI screening threshold less strict, will on average increase mortality. However, for the 55–64 year olds, increasing the allowance rate of the strictest judges would reduce mortality. This provides some evidence, at least for the 55–64 year olds, that current screening thresholds are about right.

Given that many appellants in our sample would be allowed by a high percentage of judges, the average appellant is substantially less healthy than the marginal appellant. Furthermore, the sickest SSDI applicants are allowed at the initial level and are thus not in our sample. For this reason we would expect the average SSDI applicant to be positioned well 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 being allowed SSDI benefits increases mortality, on average, for the marginal appellants for whom allowance is 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. 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 SSDI allowance on 10 year mortality does not vary in a dramatic way across gender, education, race, or prior income. 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

Table 5
Estimated effect of SSDI recipiency on mortality, disaggregated, Aged 55–64.

	Obs	Panel (a): 5 year mortality (Percent)						Panel (b): 10 year mortality (Percent)					
		Mortality rates		OLS		IV		Mortality rates		OLS		IV	
		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)
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 ≥ $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)
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)
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)
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)
Intellectual disability	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 SSDI recipiency on mortality for those aged 55–64. IV estimates use de-meaned variables and judge leniency as the instrument. Standard errors are clustered by judge.

estimates for white individuals are smaller in magnitude at both 5 and 10 years than for other racial groups.

The bottom panel of Table 5 shows estimates by health condition listed in the disability application. These health conditions are sorted by 5 year mortality rates. There are suggestive differences by health condition, although standard errors are large. Individuals diagnosed with neoplasms (e.g. cancer) have the highest overall mortality rates, and have *lower* mortality rates when allowed, in both the OLS estimates and the 10 year IV estimates. This is potentially evidence that SSDI, 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. The second highest mortality group, with respiratory disease, also has a negative IV estimate at 5 years and a near-zero estimate at 10 years. Note too that while the OLS and IV estimates are similar overall, they can be very different when subdivided by health condition. 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 SSDI affects mortality

Receiving benefits jointly affects income, labor supply, and health insurance. It is important to understand the effects of each channel since different reforms will likely impact different channels. For example, many have advocated the removal of the two year wait period for Medicare health insurance without changing cash benefits, which was trialed and evaluated in Weathers and Stegman (2012). Furthermore, in order to stimulate labor supply, the SSA is currently conducting a Trial Work Period for those on SSDI, where recipients may work for up to 9 months without the loss of benefits.

We find evidence that benefit allowance causes lower mortality for marginal appellants. We discuss in this part what can be said about these three channels separately, recognizing that this effort requires inference from other studies; our own setting cannot disentangle the first two effects, and can provide only indirect evidence on the potential effect of health insurance. As we show below, cash income and health insurance transfers to the disabled are large, which would suggest lower mortality, other factors equal. A potential offsetting effect is that receiving benefits reduces labor supply. As noted previously in Section 2.3, many studies have shown that SSDI receipt reduces employment, and other studies suggest that employment reductions can increase mortality. In this section we discuss these three channels by which allowance could impact mortality. We summarize the dollar effect of receiving benefits in Table 6.

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

Table 6
Key outcome differences between those allowed versus denied.

	(1) Outcomes for denied			(2) Coefficient on allowed						
	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]
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 estimates 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 for Medicare recipients aged 55–64. 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.
[a] 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.

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. The results in 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.

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. Since almost 100% of those allowed benefits are still receiving benefits but 54% of those denied by an ALJ are allowed 5 years later, the difference is 100−54=46%. This can be seen in the 5 year OLS difference estimate of row A. The IV estimate is smaller at 40%, which means that marginal denied applicants are more likely to be allowed on appeal than the average appellant denied by an ALJ.

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 affects health, including through investments in health via better food, shelter, and healthcare. In this section we discuss how taxable earnings and SSDI/SSI benefit income responds to benefit allowance. We show this income gain is significant.

Both income effects (through the high replacement rate) and substitution effects (beneficiaries will lose benefits if they earn above the SGA amount) cause SSDI recipients to reduce labor supply. Furthermore, SSDI/SSI beneficiaries receive Medicaid or Medicare health insurance, which reduces the value of employer-provided health insurance (French and Jones, 2011) and thus the incentive to work to obtain it.

Research on how income affects mortality for SSDI recipients is limited. In addition to the US-based studies by Gelber et al. (2023) and Silver and Zhang (2022), discussed in Section 2, Ahammer and Packham (2022) use Austrian data to find that stricter screening requirements for SSDI increase employment and do not predict worse health outcomes. Using Dutch data, García-Gómez and Gielen (2018) find that a reduction in annual benefits leads to lower mortality for men, but higher mortality for low earning women.

As discussed in Section 2, a substantial body of research finds evidence that lower employment predicts higher mortality despite the related loss of income. Row B of Table 6 presents estimated employment responses to benefit allowance.[23] While 16% of all individuals denied by an ALJ have positive earnings 5 years after assignment to an ALJ, the OLS estimates 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 dollars) 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 SSDI/SSI benefits exceed the decline in earnings. 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 later receive benefits due to successful appeal, re-application or because they become old enough to receive Social Security benefits.

We should note that we cannot assess all channels by which SSDI/SSI receipt may affect household income. For example, Autor et al. (2019) show that in Norway disability benefit receipt leads to reductions in spouse's earnings and other benefits (such as unemployment insurance). In contrast, Meyer and Mok (2019) find no evidence of spousal earnings responses to disability using US data.

[23] In estimating these responses we use the same sample of 55–64 individuals we use to estimate our mortality responses. French and Song (2014) estimate labor supply responses for more age groups and find larger labor supply responses for younger appellants.

7.3. Health insurance benefits

Individuals receiving SSDI 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. While the evidence on whether receipt of health insurance reduces mortality is mixed, the dollar value of these benefits is substantial. 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 a large effect of SSDI/SSI receipt on the likelihood that beneficiaries receive either Medicare or Medicaid over a period covering 12 months before they were awarded SSDI until 6 years after. They show that immediately following SSDI/SSI benefit receipt, 24.7% receive either Medicaid or Medicare, primarily SSI beneficiaries who receive Medicaid. The total jumps to 89.7% after 2 years when SSDI 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 the 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 increased probability of receiving Medicare and/or Medicaid is 16 percentage points after one year and 48 percentage point after 3 years.

De Nardi et al. (2016) show that the average Medicare recipient aged 55–64 receives $12,012 worth of medical transfers from Medicare/Medicaid per year. Row F of Table 6 calculates the expected value of 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 SSDI benefits are receiving on average $1,926 more in Medicare and Medicaid benefits. After 5 years this difference is $4,736 per year. Assuming that this added care from Medicare and Medicaid does not crowd out other private and public care, it represents a measure of the cost of the incremental increase in health care received.

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 SSDI by an ALJ up to age 65, when everyone should become eligible for Medicare and Social Security benefits

To calculate this value 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 SSDI 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 (Gelber et al., 2023).

7.5. Effects disaggregated by health condition

Table 5 presents evidence that mortality responses vary by health condition. We discuss here suggestive evidence that receiving benefits reduces 5-year and 10-year mortality for appellants with higher-cost medical conditions, higher-mortality conditions, and conditions for which the effect of allowance on labor supply is smaller.

Fig. 5(a) plots the 5 and 10 year mortality IV point estimates by health condition from Table 5 against mean medical spending for disabled Medicare beneficiaries under age 65 using Medicare Current Beneficiary Survey (MCBS) data.[24] The size of the circles represents the number of observations. The figure shows 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.

Over both 5 and 10 year periods, we find a general tendency for benefit allowance to be more favorable to mortality 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 the adverse effects of work disincentives.

Fig. 5(b) plots the 5 and 10 year mortality IV 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 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 expensive in terms of medical treatment. Conversely, allowance causes higher mortality among conditions with relatively low mortality, which also tend to have lower medical spending, such as intellectual disabilities, mental disorders, and musculoskeletal disorders. Similar to Figs. 5(a), 5(b) is consistent with the view that improved access to health care reduces mortality for expensive, high mortality conditions.

In Fig. 5(c) we examine evidence of a possible tradeoff between the mortality benefits from health insurance, and higher mortality from reduced employment. This figure shows that the health conditions where benefit allowance decreases employment the most are also the conditions where benefit allowance increases mortality the most (e.g., intellectual disabilities). Conversely, conditions where benefit allowance reduces employment the least are associated with falls in mortality (e.g., respiratory conditions). The negative slopes in Fig. 5 are not statistically significant, however, and should only be taken as suggestive evidence.

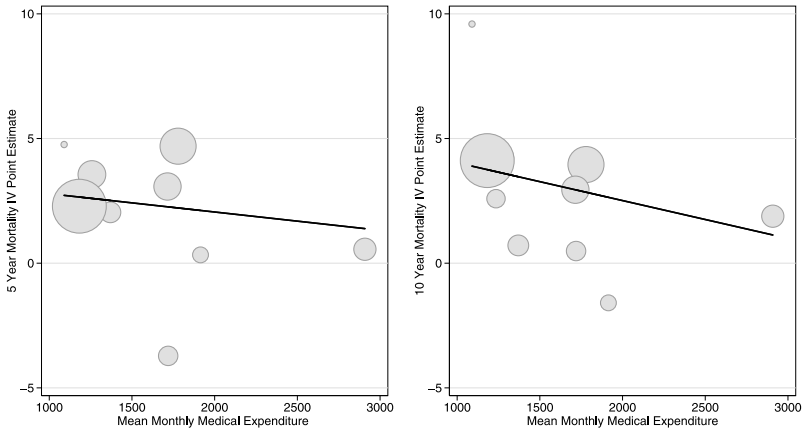
8. Robustness

Our results for the main estimation sample (those aged 55–64 at time of assignment to ALJ) are robust to a number of 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, we report estimates with no covariates as well as estimates with covariates.

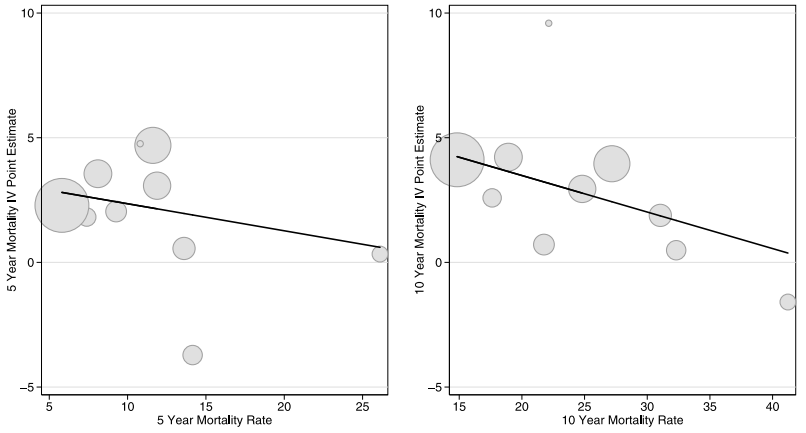
The first two rows display OLS estimates. In the second row, we include the 10,006 individuals who died in the year of application. As discussed in Section 5, we exclude these individuals because some of our sample who are recorded as denied are likely those who die before being heard by a judge.

The remaining rows provide IV estimates with different specifications. The next row reports IV results including appellants who died in the year of application. The next two rows change the minimum number of cases a judge must hear to be included in the sample to

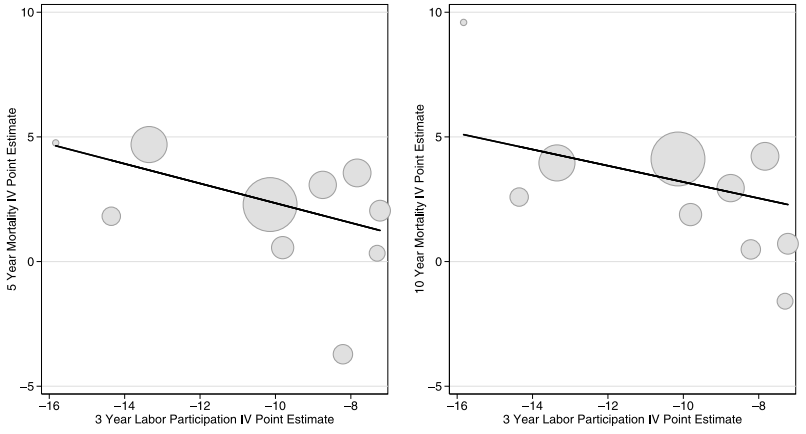
[24] We use estimates from the appendix in De Nardi et al. (2016) for those receiving Medicare benefits who are younger than 65. Virtually everyone under age 65 who receives Medicare also receives disability benefits. The MCBS has high quality medical spending data since it uses administrative Medicare records for Medicare spending and a mixture of survey data and reconciliation of survey, administrative data on Medicaid recipiency, 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.



(a) By medical expenditure



(b) By mortality rate



(c) By labor supply response

Fig. 5. Estimated mortality effect of allowance by medical expenditures, mortality rates, and labor supply for each health condition.
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 (a) mean medical expenditures, (b) unconditional mortality rate, and (c) labor supply of the individuals with each health condition from SSA mortality records.

Table 7
Robustness checks.

	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)
De-mean 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 Underreporting Correction	1.43	1.92	1.18	2.07
	(0.44)	(0.49)	(0.74)	(0.89)

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 are 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 Instrument* row we replace the baseline instrument with the one constructed in Appendix Section C.2.
In the baseline we de-mean by hearing office-day and drop judges who saw less than 200 cases, 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.

be either lower (50 cases) or higher (500 cases) than our baseline specification of 200 cases. 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 to construct our instrumental variable. The next row uses the instrument proposed by Doyle (2007) instead of our judge leniency instrument.[25] In the next row, instead of de-meaning by hearing office-day as in the baseline, we de-mean by hearing office-year. In Appendix Tables A10–A17 we repeat all our main tables, de-meaning by hearing office-year rather than hearing office-day. The final two rows adjust for any potential underreporting, using either our best estimate of the extent of underreporting p and twice this estimate, respectively.

While our main sample focuses on those aged 55–64, in Appendix B.1 we present all our main estimates for the full sample of all ages.

Taken together, these robustness checks increase our confidence in our estimation strategy. In every case our robustness check 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 and Supplemental Security Insurance allowance on mortality, for persons on the margin of being allowed or denied benefits by an ALJ. Those allowed benefits receive large cash transfers, and Medicare or Medicaid health insurance, but also face important work disincentives. Higher income and access to health insurance likely lowers mortality, while reduced employment may increase mortality. Identifying this combined effect is difficult because those allowed benefits are potentially less healthy than those denied. To circumvent this problem, we rely on the effectively random assignment of ALJs to disability cases to create an instrumental variable that measures judge leniency in allowing benefits, relative to other judges in the same hearing office and on the same day.

We find that benefit allowance increases mortality on average, for those on the margin of being allowed or denied benefits, by 2.8 percentage points after 10 years. However, there is considerable heterogeneity in the mortality response. Marginal Treatment Effects estimates suggest that allowance reduces mortality for appellants who would

[25] Doyle (2007) and French and Song (2014) construct a slightly different judge leniency variable ($\widetilde{Z}_i$) — this alternative calculation is described in detail in Appendix C.2. The main difference between these judge leniency variables is when the de-meaning takes place in the calculation.

be allowed by all but the strictest judges, but increases mortality for appellants who would only be allowed by the most lenient judges. We also find suggestive evidence that allowance reduces mortality for some expensive, high-mortality health conditions such as cancer and respiratory conditions, but increases mortality for lower cost, lower mortality conditions such as musculoskeletal disorders.

All of our estimates are for *marginal* appellants who would be allowed benefits if seen by a lenient ALJ, but be denied by a stricter one. However, the majority of disability benefit applicants are *inframarginal*, who are likely less healthy than our marginal cases and would be allowed by all judges. Thus, our findings are consistent with the view that disability benefit allowance likely reduces mortality on average.

Given the 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 disability applicants, the current disability thresholds are not far from the right level. Increasing allowance rates by making the threshold less strict will likely increase mortality for marginal appellants, although a sharp reduction in allowance rates would likely increase mortality as well.

Our findings are consistent with the view that work is good for one's health. Thus, reforms of the disability insurance rules to reduce the strong work disincentives of the current rules may improve applicant health.

However, among less healthy individuals, who are in need medical care for high-mortality conditions and would usually not work regardless of benefit receipt, disability insurance likely reduces mortality. These individuals might gain from not being subject to the current 2 year waiting period for Medicare coverage.

Declaration of competing interest

The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

Data availability

The data that has been used is confidential.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jpubeco.2023.105033.

References

Abadie, A., 2003. Semiparametric instrumental variable estimation of treatment response models. J. Econometrics 113 (2), 231–263.

Abadie, A., Athey, S., Imbens, G.W., Wooldridge, J.M., 2023. When should you adjust standard errors for clustering? Q. J. Econ. 138 (1), 1–35.

Ahammer, A., Packham, A., 2022. Disability Insurance Screening and Worker Outcomes. Mimeo.

Autor, D., Duggan, M., 2006. The growth in the social security disability rolls: A fiscal crisis unfolding. J. Econ. Perspect. 20 (3), 71–96.

Autor, D.H., Duggan, M., 2010. Supporting Work: A Proposal for Modernizing the US Disability Insurance System. Center for American Progress and the Hamilton Project.

Autor, D., Kostol, A.R., Mogstad, M., 2019. Disability benefits, consumption insurance, and household labor supply. Amer. Econ. Rev. 109 (7), 2613–2654.

Autor, D.H., Maestas, N., Mullen, K.J., Strand, A., 2017. Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants. Mimeo.

Banks, J., Cribb, J., Emmerson, C., Sturrock, D., 2019. The Impact of Work on Cognition and Physical Disability: Evidence from English Women. Technical Report, IFS Working Papers.

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

Bjorklund, A., Moffitt, R., 1987. The estimation of wage gains and welfare gains in self-selection models. Rev. Econ. Stat. 69 (1), 42–49.

Black, B., Espin-Sanchez, J.-A., French, E., Litvak, K., 2017. The long-term effect of health insurance on near-elderly health and mortality. Am. J. Health Econ. 3 (3), 281–311.

Black, B., French, E., McCauley, J., Faria de Carvalho Nunes, L., Song, J., 2022. Comparing Death Data from the Social Security Administration and the National Center for Health Statistics, 1990–2013. Mimeo.

Blandhol, C., Bonney, J., Mogstad, M., Torgovitsky, A., 2022. When is TSLS Actually Late? Technical Report, National Bureau of Economic Research.

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., Newhouse, J.P., 1983. Does free care improve adults' health? Results from a randomized trial. N. Engl. J. Med. 309 (23), 1426–1434.

Burkhauser, R.V., 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., Wilkins, R., 2014. Disability benefit growth and disability reform in the US: Lessons from other OECD nations. IZA J. Labor Policy 3 (1), 4.

Chatterji, P., Meara, E., 2010. Consequences of eliminating federal disability benefits for substance abusers. J. Health Econ. 29 (2), 226–240.

Dahl, G.B., Kostol, A.R., Mogstad, M., 2014. Family welfare cultures. Q. J. Econ. 129 (4), 1711–1752.

De Nardi, M., French, E., Jones, J.B., McCauley, J., 2016. Medical spending of the US elderly. Fiscal Stud. 37 (3–4), 717–747.

Deshpande, M., Gross, T., Su, Y., 2021. Disability and distress: The effect of disability programs on financial outcomes. Am. Econ. J.: Appl. Econ. 13 (2), 151–178.

Dobkin, C., Puller, S.L., 2007. The effects of government transfers on monthly cycles in drug abuse, hospitalization and mortality. J. Public Econ. 91 (11–12), 2137–2157.

Doyle, J.J., 2007. Child protection and child outcomes: Measuring the effects of foster care. Am. Econ. Rev. 97 (5), 1583–1610.

Eliason, M., Storrie, D., 2006. Lasting or latent scars? Swedish evidence on the long-term effects of job displacement. J. Labor Econ. 24 (4), 831–856.

Evans, W.N., Moore, T.J., 2011. The short-term mortality consequences of income receipt. J. Public Econ. 95 (11–12), 1410–1424.

Finkelstein, A., McKnight, R., 2008. What did medicare do? The initial impact of medicare on mortality and out of pocket medical spending. J. Public Econ. 92 (7), 1644–1669.

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

Finlay, K., Genadek, K.R., 2021. Measuring all-cause mortality with the Census Numident file. Am J Public Health 111 (S2), S141–S148.

Fitzpatrick, M.D., Moore, T.J., 2018. The mortality effects of retirement: Evidence from Social Security eligibility at age 62. J. Public Econ. 157, 121–137.

Frandsen, B., Lefgren, L., Leslie, E., 2023. Judging judge fixed effects. Amer. Econ. Rev. 113 (1), 253–277.

French, E., 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., 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., Song, J., 2014. The effect of disability insurance receipt on labor supply. Am. Econ. J.: Econ. Policy 6 (2), 291–337.

French, E., Taber, C., 2011. Identification of models of the labor market. In: Handbook of Labor Economics, Vol. 4. Elsevier, pp. 537–617.

García-Gómez, P., Gielen, A.C., 2018. Mortality effects of containing moral hazard: Evidence from disability insurance reform. Health Econ. 27 (3), 606–621.

Gelber, A., Moore, T., Strand, A., 2023. Disability insurance income saves lives. J. Polit. Econ.

Goldin, J., Lurie, I.Z., McCubbin, J., 2021. Health insurance and mortality: Experimental evidence from taxpayer outreach. Q. J. Econ. 136 (1), 1–49.

Guldi, M., Hawkins, A., Hemmeter, J., Schmidt, L., 2022. Supplemental security income for children, maternal labor supply, and family well-being: Evidence from birth weight eligibility cutoffs. J. Hum. Resour. 0818–9654R2.

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

Heckman, J.J., Urzua, S., Vytlacil, E., 2006. Understanding instrumental variables in models with essential heterogeneity. Rev. Econ. Stat. 88 (3), 389–432.

Hill, M.E., Rosenwaike, I., 2002. The Social Security administration's death master file: The completeness of death reporting at older ages. Soc. Secur. Bull. 64, 45–51.

Hoynes, H.W., Maestas, N., Strand, A., 2022. Legal Representation in Disability Claims. Technical Report, National Bureau of Economic Research.

Hu, J., Lahiri, K., Vaughan, D.R., Wixon, B., 2001. A structural model of social security's disability determination process. Rev. Econ. Stat. 83 (2), 348–361.

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

Kaestner, R., 2021. Mortality and science: A comment on two articles on the effects of health insurance on mortality. Econ. J. Watch 18 (2), 192.

Kolesar, M., 2013. Essays on Instrumental Variables. Harvard University.
Kuhn, A., Staubli, S., Wuellrich, J.-P., Zweimüller, J., 2020. Fatal attraction? Extended unemployment benefits, labor force exits, and mortality. J. Public Econ. 191, 1040–1087.
Lee, D.S., McCrary, J., Moreira, M.J., Porter, J., 2022. Valid t-ratio inference for IV. Amer. Econ. Rev. 112 (10), 3260–3290.
Livermore, G., Stapleton, D.C., 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 Aff. 30 (9), 1664–1672.
Maestas, N., Mullen, K.J., Strand, A., 2013. Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt. Am. Econ. Rev. 103 (5), 1797–1829.
Meyer, B.D., Mok, W.K., 2019. Disability, earnings, income and consumption. J. Public Econ. 171, 51–69.
Miller, S., Johnson, N., Wherry, L.R., 2021. Medicaid and mortality: New evidence from linked survey and administrative data*. Q. J. Econ. qjab004.
Myers, R.J., 1993. Social Security. Univ of Pennsylvania Press.
Parsons, D.O., 1991. Self-screening in targeted public transfer programs. J. Polit. Econ. 99 (4), 859–876.
Rege, M., Telle, K., Votruba, M., 2009. The effect of plant downsizing on disability pension utilization. J. Eur. Econom. Assoc. 7 (4), 754–785.
Rohwedder, S., Willis, R.J., 2010. Mental retirement. J. Econ. Perspect. 24 (1), 119–138.
Rupp, K., Riley, G.F., 2012. Longitudinal patterns of Medicaid and Medicare coverage among disability cash benefit awardees. Soc. Secur. Bull. 72, 19.
Silver, D., Zhang, J., 2022. Impacts of Basic Income on Health and Economic Well-Being: Evidence from the VA's Disability Compensation Program. Technical Report, National Bureau of Economic Research.
Sloczynski, T., 2022. When Should We (Not) Interpret Linear IV Estimands as LATE? Mimeo.
Snyder, S.E., Evans, W.N., 2006. The effect of income on mortality: Evidence from the Social Security notch. Rev. Econ. Stat. 88 (3), 482–495.
Social Security Advisory Board, 2006. Improving the social security administration's hearing process.
Sullivan, D., von Wachter, T., 2009. Job displacement and mortality: An analysis using administrative data. Q. J. Econ. 124 (3), 1265–1306.
Taylor, M.H., 2007. "Refugee Roulette" in an administrative law context: The "Déjà vu" of decisional disparities in agency adjudication. Stanf. Law Rev. 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, 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.
Waddell, M., 2020. Trump plans to tighten Social Security disability criteria. Wall Street J.
Weathers, R.R., Stegman, M., 2012. The effect of expanding access to health insurance on the health and mortality of social security disability insurance beneficiaries. J. Health Econ. 31 (6), 863–875.