The Effect of Disability Insurance Receipt on Labor Supply

Eric French and Jae Song*

June 8, 2012

Abstract

This paper estimates the effect of Disability Insurance receipt on labor supply. Exploiting the effectively random assignment of judges to disability insurance cases, we use instrumental variables to address the fact that those allowed benefits are a selected sample. We find that benefit receipt reduces labor force participation by 26 percentage points three years after a disability determination decision, although the reduction is smaller for those over age 55, college graduates, and those with mental illness. OLS estimates are similar to instrumental variables estimates. We also find that over 60% of those denied benefits by an Administrative Law Judge are subsequently allowed benefits within 10 years, showing that most applicants apply, re-apply, and appeal until they get benefits.

*Comments welcome at jae.song@ssa.gov and efrench@frbchi.org. Affiliations are Social Security Administration and Federal Reserve Bank of Chicago, respectively. French thanks the Social Security Administration for hospitality while much of this paper was written. We thank David Autor, John Bound, Meredith Crowley, Mariacristina DeNardi, John Jones, Wojciech Kopczuk, Fabian Lange, Bruce Meyer, David Pattison, Ted To, Bernie Wixon and seminar participants at the Federal Reserve Bank of Chicago, Federal Reserve Bank of New York, Cambridge University, Lusanne, Northwestern University (Kellogg), University of Chicago (Booth), University of Notre Dame, University of Illinois-Chicago, Syracuse, Washington University, Ohio State University, Yale University, SOLE Meetings, National Tax Association Meetings, Michigan Retirement Research Conference, the Arizona State University Human Capital Conference, and American Economics Association Meetings, for helpful comments, Deanna Becker and especially David Benson, Angshuman Gooptu, An Qi, Shani Schechter, and Zach Seeskin for excellent research assistance. The views in this paper are not necessarily those of the Federal Reserve System or Social Security Administration.

1 Introduction

This paper presents new evidence on the effect of Disability Insurance (DI)/Supplemental Security Income (SSI) receipt on labor supply. We compare the earnings patterns of individuals who applied for and received disability insurance benefits to the earnings patterns of those who applied for benefits but were denied.

Relative to Bound's (1989) classic study on earnings of rejected DI applicants, we make the following key improvement. We address the fact that those who are denied benefits are potentially different than those who are allowed. Using Social Security administrative data, we exploit the assignment of DI cases to Administrative Law Judges (ALJs), an assignment which is essentially random. We document large differences in allowance rates across judges, and show that these differences are unrelated to the health or earnings potential of DI applicants. Using instrumental variables procedures, we use judge specific allowance rates to predict allowance of individual cases. We then use predicted allowance to estimate the effect of allowance on labor supply.

We find that three years after assignment to an ALJ, DI benefit allowance reduces earnings \$4,059 per year and labor force participation 26 percentage points. As it turns out, our estimates are not very sensitive to accounting for the fact that those who are denied benefits are potentially different than those who are allowed: instrumental variables estimates are very close to OLS estimates for those assigned to an ALJ. These estimates imply a high labor supply elasticity with respect to the after-tax wage. The earnings and participation elasticities are 1.8 and 1.5, respectively.

However, many initially-denied DI applicants appeal or re-apply. In fact, we find that 40% of applicants who are denied benefits by an ALJ are eventually allowed benefits within three years. Furthermore, 40% of those not allowed benefits three years after an assignment to an ALJ are allowed benefits within 10 years of assignment. In order to be allowed benefits, the applicant cannot earn above a small amount. As a result, few applicants work during the appeal process, even though they are currently not receiving benefits. This has an important impact on our estimated effects. When we measure earnings and DI benefit allowance five years after assignment to an ALJ, rather than three, we find that DI allowance reduces earnings \$4,915 per year, rather than \$4,059.

Furthermore, we estimate labor supply responses for different subgroups of the population.

We identify many subgroups of the population whose labor supply is not sensitive to benefit receipt, such as those over age 55, college graduates, and those with mental illness. Because we have the population of DI applicants whose case was heard by a judge, we obtain precise estimates of the labor supply responses, even for these narrow subgroups of the population.

Using a Marginal Treatment Effects approach, we find that marginal applicants handled by stricter judges (who allow benefits to relatively few applicants) have slightly smaller labor supply responses than the marginal applicants heard by lenient judges. This is consistent with the view that the marginal applicant handled by a strict judge is slightly less able to work than the marginal case handled by a more lenient judge. The marginal case heard by a stricter judge is, however, slightly more likely to get benefits in the future. This suggests that these strict judges delay benefit receipt rather than deny benefit receipt.

Section 2 gives a literature review, section 3 describes the DI system, section 4 describes our estimation methods, section 5 shows data, section 6 reports basic estimates, and section 7 concludes.

2 Literature Review

Disability Insurance is one of America's largest social insurance programs. In 2005, 4.1% of men ages 25-64 were receiving DI benefits (Autor and Duggan 2006). Furthermore, many disabled individuals with low income receive Supplemental Security Income benefits. Most DI and SSI beneficiaries also receive health insurance benefits through Medicare (for DI beneficiaries) or Medicaid (for SSI beneficiaries). The combined cost of these programs was $428 billion in 2008 (Livermore et al. 2011), making these programs several times more expensive than unemployment insurance. These rapidly rising costs have generated many policy proposals to reform the system (Autor and Duggan 2010, Burkhauser and Daly 2011).

DI is often cited as a major cause of the fall in labor supply of American men aged 55-64. In order to better understand the labor supply effects of DI, Bound (1989) compared earnings patterns of individuals who applied for and received DI benefits to those who applied for benefits but were denied. He found that those who were allowed benefits were less likely to work than those who were denied, but the effect was modest. Even those who were denied benefits had participation rates of less than 50% after denial of benefits. The difference in participation rates of those allowed versus denied was 34 percentage points. Thus, Bound inferred that at most 50% of rejected male applicants during the 1970s would have worked were it not for the availability of disability benefits. These estimates imply that DI is responsible for well under half of the fall in labor supply of American men aged 55-64 over the 1970s and 1980s.

Von Watcher et al. (2011) find that these labor supply responses have if anything grown over time because applicants are now younger and have potentially less severe health impairments. Consistent with Von Watcher et al. (2011), Duggan and Imberman (2008) point out that 13.5 percent of DI awards in 1982-83 were for mental disorders, while in 2002-03 it was 25.7 percent. Nevertheless, Bound's original estimate is still very close to the most recent OLS estimates. For example, Bound's estimate was 0.34. Maestas et al. (2011) use recent administrative data, and find that the estimate is .35. It is worth noting that our OLS estimates are .27, smaller than those of Bound (1989), Von Watcher et al. (2011) and Maestas et al. (2011). The reason for this is that they use estimates from the initial stage, whereas we use estimates from the ALJ stage.

Parsons (1991) and Bound (1989, 1991) discuss three key criticisms of Bound's approach. First, those who are denied benefits are different than those who are allowed. Differences in labor supply between those denied and allowed are partly due to the effect of DI, but also partly due to the two groups having different propensities to work, even when receiving the same DI treatment. People whose applications were denied are likely to be in better health, which, all else equal, should make them more likely to work, which is what Bound (1989) argued. However, those who are denied benefits also tend to have very intermittent work histories (Lahiri et al. 2008), suggesting that their non-health characteristics make them less likely to work. For this reason, OLS might be biased up or down. As a result, it is not clear whether those who are denied are more or less likely to work in the absence of benefits and whether OLS overstates or understates the work disincentive effects of DI.

It is this problem that our study addresses. Our identification approach compares those who are denied benefits to those who are otherwise similar but are allowed benefits. Our approach complements the approach of Chen and Van der Klaauw (2008) who exploit the vocational grid. They use the fact that in many cases, an individual aged 54 applying for benefits would be denied, although the same individual at age 55 would be allowed. Our estimated labor supply effects are similar to Chen and Van der Klaauw (2008). However, we add to their analysis by providing larger sample sizes. This allows for more precise estimates. It also allows us to document how the responsiveness of labor supply varies with demographics, because we can obtain precise estimates for narrow subgroups.

Our estimated effects are also similar to Maestas et al. (2011), who use assignment of disability examiners at the initial stage of the DI application process as a source of variation in allowance rates. This paper makes three contributions relative to that paper. The first is that judges are assigned to cases on a rotational basis, which makes the assignment process random for all practical purposes, whereas examiners at the initial stage may specialize. Thus our source of variation is more clearly exogenous. Second, we obtain more precise estimates, allowing us to document how the responsiveness of labor supply varies with demographics. Third, our data includes earnings and the share of individuals who are allowed or are appealing up to 10 years after the ALJ allowance decision, whereas they have data only on earnings and the share working, and only up to three years after an initial allowance decision. This is important because we find that 40% of those not allowed benefits three years after an assignment to an ALJ are allowed benefits within 10 years of assignment.

Our paper, Van der Klaauw (2008) and Maestas et al. (2011) all obtain identification at different stages of the adjudication process, and thus our estimated effects correspond to different pools of applicants. Thus the three studies are of independent interest. For example, the disparities in allowance rates across ALJs has received a great deal of attention in policy circles (Social Security Advisory Board, 2006), legal studies (Taylor, 2007), and the popular press (Paletta, 2011). Despite the differences between our paper, Chen and Van der Klaauw (2008), and Maestas et al. (2011), all three papers produce similar results and reinforce each other's findings.

The second criticism of Bound's approach is that many individuals who are denied continue to appeal the denial. In order to be deemed eligible for benefits, the individual cannot work while appealing the denial. Thus, many of those who are denied do not work in order to increase the chances of successful appeal. If the option to appeal had not existed, more of these individuals might have returned to the labor force. We partly address this problem by estimating the labor supply response to whether the individual was allowed benefits three years after assignment to a judge, although we show that many re-apply and appeal well after three years. We provide new evidence on the share of denied individuals who appeal and subsequently receive benefits.[1]

Third, in order to apply for benefits, the individual must be out of the labor force for a period of time. For example, the individual can only work a very limited amount in the five months before applying for benefits. During that period, human capital may depreciate (Autor et al. (2012)). Thus the individual may not be able to return to her previous job, even if she is healthy. In other words, the very act of applying for benefits reduces ability to work. Our study does not address this issue.

3 The Disability Insurance System

This section shows that that the DI application process is high stakes: DI benefits are worth about $200,000 to a typical beneficiary if they maintain low earnings. Those allowed benefits face strong work disincentives. Those denied benefits face strong incentives to re-apply and appeal. Judges who make allowance decisions are for all practical purposes randomly assigned to cases. Judicial independence means that judges have a great deal of latitude to determine eligibility (Taylor, 2007), and as a result judges can have very different allowance rates.

3.1 Labor Supply Incentives

Both income effects (through the high replacement rate) and substitution effects (beneficiaries will lose benefits if they earn above the SGA amount) indicate that DI should reduce labor supply. If an applicant is allowed DI benefits, the dollar amount of benefits depends on previous labor earnings. Disabled worker benefits averaged $1,004 per month among DI beneficiaries in 2007 (Social Security Administration, 2008). Because the benefit schedule is progressive, disability benefits replace 60% and 40% of labor income for those at the 10th and 50th percentile of the earnings distribution, respectively (Autor and Duggan 2006). Those receiving benefits can earn up to the Substantial Gainful Activity level (SGA), which was $500 per month (in current dollars) during the 1990s and $900 per month in 2007. Those earning more than this amount for more than a nine month Trial Work Period lose their benefits.

[1]Understanding subsequent allowance and appeal is also an important input into dynamic models of DI application and receipt, such as Bound et al. (2010), Benitez-Silva et al. (2011), Low and Pistaferri (2011).

Furthermore, DI benefits likely reduce labor supply through a third channel – Medicare eligibility. Individuals receiving DI benefits are eligible for Medicare after a two year waiting period. Medicare largely eliminates the value of employer-provided health insurance. For those working at a firms providing health insurance, Medicare eliminates an important work incentive (French and Jones, 2011). Livermore et al. (2011) show that federal and state governments spend more on health care than on cash benefits for the disabled.

Disabled individuals with especially weak earnings histories and low asset levels are eligible for a related program called Supplemental Security Income (SSI). SSI benefits are not a function of previous labor income. The Federal Maximum SSI benefit level was $386 per month in 1990 and $623 in 2007. Some states supplement this benefit. Benefits are reduced by 50 cents for every dollar of labor income. Individuals drawing SSI may also be immediately eligible for Medicaid, the government provided health insurance program for the poor. Many people draw both DI and SSI benefits concurrently.

Relatively few people lose disability benefits for reasons other than death.[2] For example, of 7.1 million individuals (DI worker beneficiaries) drawing DI benefits in 2007, 0.5% had benefits terminated because they earned above the SGA level for an extended period of time in 2007. Another 0.3% had benefits terminated because they were deemed medically able to work after a continuing disability review, which is a periodic review of the health of DI beneficiaries (Social Security Administration, 2007).

The disability allowance decision is high stakes. If the individual is allowed benefits, that individual is typically given disability benefits until the normal retirement age (age 65 during the 1990s and now 66), when these benefits are converted into Social Security benefits. Thus a 52 $\frac{1}{2}$ year old receiving $12,000 in annual disability benefits will likely receive these benefits for 12 $\frac{1}{2}$ years, meaning that she will receive $150,000 in transfers. Furthermore, two years after receiving benefits, she will receive Medicare benefits, which are worth at least $50,000. Thus, being allowed benefits is worth on average $200,000 over a lifetime.

3.2 Determining Eligibility for DI benefits

An individual is deemed eligible for benefits if they have met certain work requirements and if they are deemed medically disabled. Although the exact algorithm is complex (see Hu

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

et al. 2001, Benitez-Silva et al. 1999, for details), one of two conditions must be met for the individual to be deemed disabled.

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

The second condition is inability to perform either past work or other work. This condition involves a combination of medical impairment and vocational factors such as education, work experience, and age. These cases can be especially difficult to evaluate. Myers (1993), a former Social Security Administration Deputy Commissioner, points out that "if a worker has a disability so severe that he or she 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!"

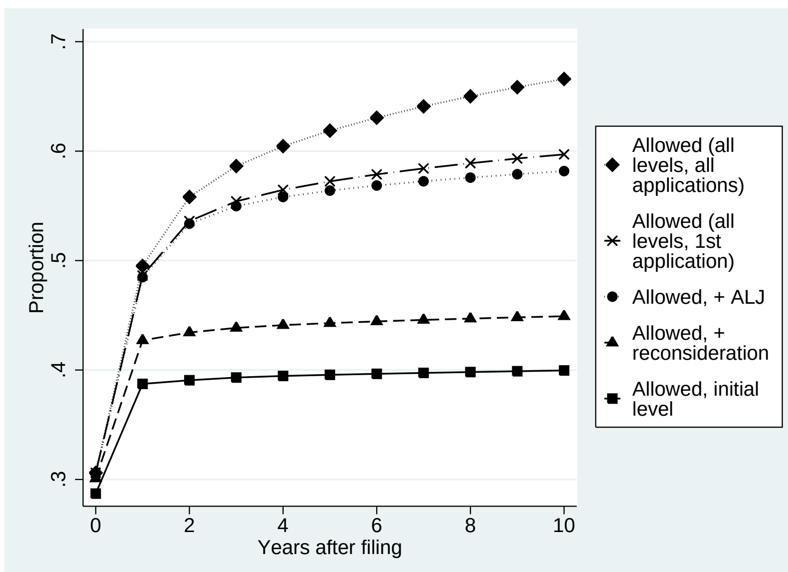

Figure 1: Allowance at different stages of the applications and appeals process.

The disability determination process is a multi-step process. Figure 1 shows the share of applicants who are allowed at different steps during our sample period (described in detail in Section 4 and Appendix A). After an initial waiting period of five months, DI applicants have their case reviewed by a Disability Determination Service review board. Figure 1 shows that 39% of applicants are allowed and 61% are denied at this stage. At this stage the most clear-cut cases are allowed, such as those with a listed impairment. Cases that are more difficult to judge (such as musculoskeletal problems) are usually denied at this stage. About half of all applicants denied for medical reasons appeal at the disability determination service reconsideration stage. About 10% of those that appeal are allowed benefits at this stage (Social Security Administration, 2008). Sixty days after the disability determination service decision, a DI appeal can be requested. DI appeals are reviewed in court by Administrative Law Judges (ALJs) after a delay of about one year.[3] 14% of all initial claims, or 59% of all claims that are appealed, are allowed at the ALJ level.[4] If the case is denied at the ALJ level, the applicant can then appeal to the Appeals Council level. If the applicant is denied at this level, she can then appeal after 60 days at the Federal Court level. However, Figure 1 shows that appeals at the higher levels are rarely successful: less than 2% of all initial claimants receive benefits at the Appeals Council or Federal Court level. Lastly, denied applicants can end their appeal and re-apply for benefits. The last line on Figure 1 includes those who re-apply for benefits. Another 7% of all initial claims are eventually allowed benefits through a re-application. 33% do not get benefits at any stage after 10 years. Figure A1 in the appendix shows that most who do not get benefits after a few years end their appeals. However, 10 years after initially claiming, 6% are still in the process of appealing or re-applying.

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

[3] 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.

[4] The full allowance rate at this stage is slightly higher than 59%. Our 59% allowance rate is for our estimation sample, which drops pre-reviewed cases that have higher allowance rates. See footnote 7.

3.3 Assignment of DI cases to judges

Administrative Law Judges (ALJs) are assigned to appeals cases on a rotational basis, with the oldest cases receiving priority at each hearing office.[5] Thus, the oldest case is given to the judge who most recently finished a case. Therefore, conditional on applying at a given office at a given point in time, the initial assignment of cases to judges is "essentially random" (Social Security Advisory Board, 2006). Judges do not get to pick the cases they handle. Judges are not assigned cases based on the expertise of the judge. Furthermore, an individual cannot choose an alternate judge after being assigned a judge.

The initially assigned judge is not necessarily the judge who decides the case. Paletta (2011) documents a judge who took assigned cases from other judges and made decisions on those cases. Thus the cases were not randomly assigned to the deciding judge.[6] Fortunately, however, we have information on the assigned judge in addition to the deciding judge. Although the deciding judge is not necessarily randomly assigned, the initially assigned judge is. We use the initial assignment to a judge as our source of exogenous variation. As it turns out, the initially assigned judge is the same as the deciding judge in 96% of all cases.

The assigned judge is for all practical purposes randomly assigned conditional on hearing office and day. However, individuals are not randomly assigned to hearing offices. The zip code in which a person lives determines the hearing office to which they are assigned. The characteristics of applicants vary by location (e.g., black lung disease is more common near mining towns) as well as across time (e.g., the share of DI applicants listing mental illness as the main health problem has risen over time). For this reason we condition explicitly on

[5]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." (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. Fortunately we can identify cases that were decided without a hearing and we delete them from our sample. Our analysis focuses on the remaining cases where there was a hearing.

[6]Furthermore, an individual can potentially reject the assigned judge. For example, if an individual misses her court case, she may be reassigned to a different judge. Another possibility is that for some cases in remote areas, cases 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.

hearing office and day in the estimations below. In doing so, we exploit only within hearing office-day variation in judge level leniency.

4 Estimating Equations

In order to estimate the effect of DI allowance on earnings and labor force participation, we use a two-step procedure. In the first step we generate an instrumental variable that is a measure of judge leniency. Conditional on the hearing office and time, this variable is correlated with the probability of allowance, but is independent of health, ability, or preferences for work. In the second step we use instrumental variables procedures to estimate the effect of DI on earnings, participation, appeals, and subsequent allowance.

4.1 Basic Specification

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

$$A_{it} = j_i\gamma_t + X_i\delta_{At} + e_{it}. \tag{1}$$

where A_{it} is a 0-1 indicator equal to 1 if individual i is allowed benefits at time t, j_i is a full set of judge indicator variables equal to 1 if judge j heard individual i's case, and X_i is a full set of hearing office-day indicators (equal 1 if individual i's case is assigned to that hearing office-day pair). The allowance rate and estimated parameters depend on time since many individuals initially denied benefits are subsequently allowed.

For the second stage we adopt the random coefficients model of Bjorklund and Moffitt (1987):

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

where $y_{i\tau}$ is either earnings, participation, appeals or allowance at time τ. We allow for time $\tau \geq t$ so that we can observe the effect of time t allowance on time τ outcomes. We allow for heterogeneity in the parameter $\phi_{i\tau}$ to capture heterogeneity in the effect of benefit receipt on earnings, appeals, and allowance, both across individuals and over time. We allow the variables $u_{i\tau}$ and $\phi_{i\tau}$ to be potentially correlated with A_{it}, and with each other.[7] Ideally we would be able to identify the entire distribution of $\phi_{i\tau}$, although this is not possible. Below we describe what is identified given our data.

4.2 Estimating Equations

When estimating equation (2) we are confronted with three concerns. First, we wish to allow for heterogeneity in the parameter $\phi_{i\tau}$. Second, we have 1,497 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. (2009)). Third, we have over 200,000 hearing office-day interactions in the covariate set X_i.

In order to address these three concerns, our estimation procedure is as follows. First, we de-mean variables by hearing office and day, and construct variables $\widetilde{A}_{it} = A_{it} - \bar{A_{it}}$, $\widetilde{y}_{i\tau} = y_{i\tau} - \bar{y_{i\tau}}$ where $\bar{A_{it}}$ and $\bar{y_{i\tau}}$ are the mean values of $A_{it}, y_{i\tau}$ conditional on the hearing office and on the day that case i was assigned. Second, for every observation i in our sample, we estimate equation (1) in where A_{i1} (the ALJ decision) is the dependent variable. We leave out observation i, as in a jackknife estimator and calculate the mean of the difference between each of judge j_i's allowance decisions and the average allowance rate of all cases heard at the same hearing office and day. We define the estimated value of γ_1 from this procedure as $\hat{\gamma}_{1,-i}$. The instrumental variable is $\widetilde{j}_i\hat{\gamma}_{1,-i}$, which we refer to as the judge allowance differential. Because we remove observation i, the estimated parameter $\hat{\gamma}_{1,-i}$ is independent of e_{it} or $u_{i\tau}$, even in a small sample. Third, we estimate the equations

$$\widetilde{A}_{it} = \lambda_t \widetilde{j}_i \hat{\gamma}_{1,-i} + \epsilon_{it}, \tag{3}$$

$$\widetilde{y}_{i\tau} = \phi_\tau \widehat{\widetilde{A}}_{it} + \widetilde{u}_{i\tau} \tag{4}$$

jointly using two stage least squares.

Given the above assumptions, Heckman, Urzua, and Vytlacil (2006) and French and Taber (2011) point out that this procedure identifies a weighted average of $\phi_{i\tau}$ for the set of indi-

[7]The residual $u_{i\tau}$ is potentially correlated with A_{it} because those allowed benefits potentially have low earnings potential. Furthermore, $\phi_{i\tau}$ is potentially correlated with A_{it} because more disabled people are unlikely to work, even when they get the benefit. Finally, $u_{i\tau}$ and $\phi_{i\tau}$ are potentially correlated with each other since unhealthy individuals have lower earnings, whether or not they are allowed benefits.

viduals affected by the instrument if three conditions are met. First, if judges are randomly assigned to cases, conditional on date and hearing office, then assignment satisfies the "independence assumption". Second, if judges differ only in leniency, then Imbens and Angrist's (1994) "monotonicity assumption" is satisfied. The monotonicity assumption implies that a case allowed by a strict judge will always be allowed by a lenient one. Third, we assume that the instrument causes variation in allowance rates, sometimes known as the rank or existence condition. Sections 6.1 and 6.2 provide evidence on the extent to which the independence, monotonicity, and rank assumptions hold.[8]

4.3 Marginal Treatment Effects

Section 6.6 presents estimated Marginal Treatment Effects (MTEs), which is the participation or earnings response for the individuals whose allowance decision is affected by changing the instrument. We estimate the equations

$$\widetilde{A}_{it} = \widetilde{f(\widetilde{j}_i \hat{\gamma}_{1,-i})} + \eta_{it}, \tag{6}$$

$$\widetilde{y}_{i\tau} = \widetilde{K}(\widehat{\widetilde{A}}_{it}) + \mu_{i\tau} \tag{7}$$

where $\widehat{\widetilde{A}}_{it}$ is the predicted value of $\widetilde{A}_{it}$ from equation (6), and the $\sim$ above the functions $f(.)$ and $K(.)$ means that they are also de-meaned. As shown by Heckman, Urzua, and Vytlacil (2006) and French and Taber (2010), as well as appendix B, the MTE is

$$K'(a_t) = E[\phi_{i\tau} | \text{allowed if } \widehat{\widetilde{A}}_{it} \geq a_t, \text{not allowed if } \widehat{\widetilde{A}}_{it} < a_t,] \tag{8}$$

[8]More formally, we are assuming that allowance follows

$$A_{it} = 1\{g_t(Z_i) - V_i > 0\} \tag{5}$$

where $Z_i = (j_i, X_i)$. The residual V_i can be thought of as the lack of severity of disability observed by the judge (but not by the econometrician). Equation (5) implies that all judges observe the same signal of disability V_i but differ in the level of severity necessary to be allowed benefits $g_t(Z_i)$. We assume V_i is independent of j_i and X_i, sometimes called the independence assumption. The latent variable framework gives rise to the monotonicity assumption. The rank condition is that $plim \widehat{A}_{it} = \Pr(A_{it} = 1 | Z_i)$ is a non-trivial function of Z_i. Equation (5) is not identified because a monotonic transformation of both $g(.)$ and V delivers the same choice probabilities. As a normalization, we assume that V_i is distributed uniformly. Furthermore, as a functional form assumption we assume that $g(.)$ is linear in j_i and X_i so that we can estimate equation (5) using the regression function in equation (1).

where a_t is a particular realization of allowance. Equation (8) shows that the MTE is the mean value of $\phi_{i\tau}$ for those who would be allowed if the value if their assigned judge allowed slightly higher than a share a_t of cases, and would be denied if assigned to a judge allowing slightly lower than a share a_t of cases. This value of a_t can also be interpreted as the (lack of) judge-observed severity of the case. As a_t increases, the instrument affects individuals with lower levels of severity. We estimate $\hat{\gamma}_{1,-i}$ from equation (1) as before, then estimate equations (6) and (7), allowing the functions $f(.)$ and $K(.)$ to be polynomials. Heckman et al. (2006) experiment with different approaches to estimating the MTE. They find that the polynomial approach works about as well as other procedures. Our Monte Carlo simulations suggest there is very little bias when using polynomials. Furthermore, the polynomial procedure is computationally feasible when allowing for large numbers of covariates, such as a full set of hearing office-day interactions. Appendix B provides more details on interpretation and estimation of the MTE.

5 Data

Our initial sample is the universe of individuals who appealed either a DI or SSI benefit denial, and were assigned to an ALJ during the years 1990-1999. Using Social Security Numbers, we match together data from the SSA 831 file, the Office of Hearings and Appeals Case Control System (OHACCS), the Hearing Office Tracking System (HOTS), the Appeals Council Automated Processing System (ACAPS), the Litigation Overview Tracking System (LOTS), the Master Earnings file (MEF), and the Numerical Identification file (NUMIDENT). These data are described in greater detail in the appendix. To the best of our knowledge, neither the OHACCS, HOTS, ACAPS, nor the LOTS datasets have been used for research purposes before. We match in earnings, reapplications and appeals data from 11 years prior to 10 years following assignment to a judge. Thus our earnings and appeals data run from 1979 to 2009.

We drop all observations heard by a judge who heard less than 50 cases during the sample period. We also drop cases with missing education information. Table A1 in Appendix A presents more details on sample selection criteria and table A2 presents mean age, race, earnings histories, and health of individuals in our estimation sample. Our main estimation sample has 1,779,825 DI cases, heard by 1,497 judges, with a mean allowance rate at the ALJ stage of 64.5%. Because many of those denied by an ALJ appeal or re-apply for benefits, the allowance rate three years after assignment is 76.9%. All dollar amounts listed below are in 2006 dollars, deflated by the CPI.

These cases were heard at 227 different hearing offices (including temporary remote sites) over our 10 year sample period. Cases were heard on 217,663 hearing office-day pairs that our procedure must account for. Thus on an average $1,779,825/217,748 = 8.2$ cases were heard at each hearing office-day pair. Although 217,663 hearing office-day fixed-effects is a large number to account for, recall that consistency in fixed effects estimators depends on the number of observations going to infinity, not the number of observations per fixed effect going to infinity. A non-trivial number of cases (15.7%) were heard when there was only a single judge at the hearing office. Given that identification in our instrumental variables estimation comes from across judge variation in allowance rates within hearing office-day pairs, these observations do not contribute any identifying variation. Nevertheless, the other observations contribute useful identifying information, as the results below show.

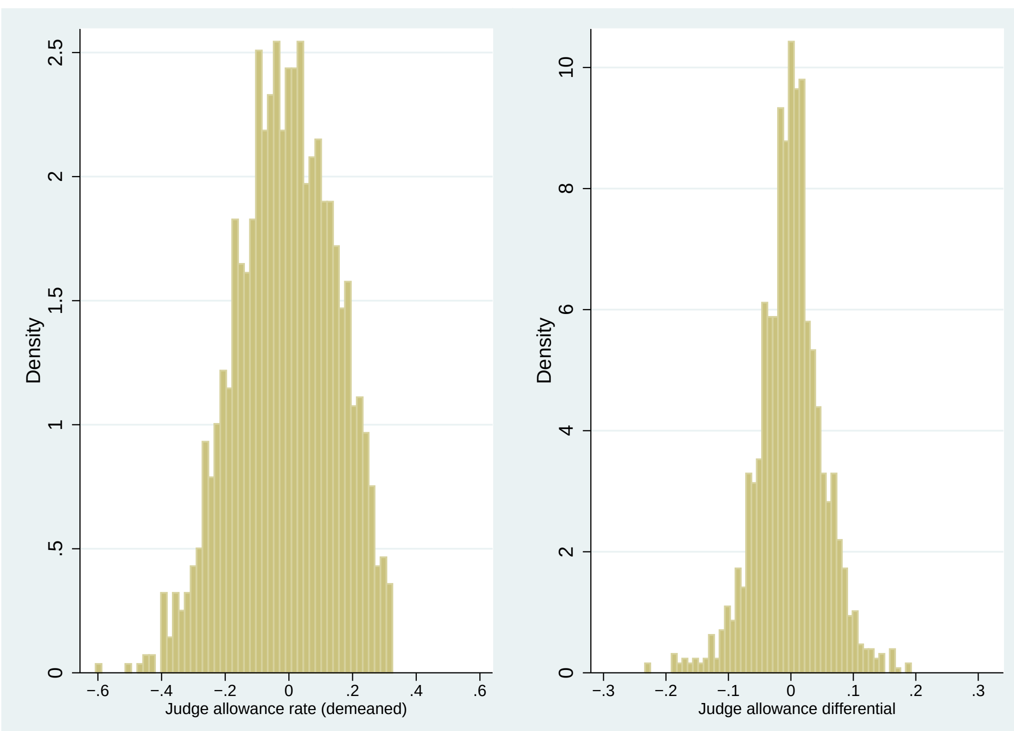

Figure 2: Allowance rate of ALJs, de-meaned, and de-meaned by hearing office and day.

Figure 2 plots the distribution of judge specific allowance rates, both unconditional (left panel) and also conditional on hearing office-day (right panel). Specifically, the left panel plots the distribution of average allowance rates of different judges over the sample period. The right panel plots the judge allowance rate de-meaned by hearing office and day; it is thus the histogram of our instrumental variable. Figure 2 shows that there is less variation in allowance rates after conditioning on hearing office and day; one standard deviation in the unconditional judge allowance rate is .153, whereas conditional on hearing office and day it is .0659 (when weighted by the number of cases handled by the judge). This means that being assigned to a judge one standard deviation more lenient than the average at her office increases the probability of allowance at the ALJ stage by 6.59 percentage points. Thus conditioning on hearing office and day removes a non-trivial share of variation in judge allowance rates, but much of the variation is within hearing office and day.

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 cannot predict the judge using observable characteristics of the judge's caseload. Table 1 presents tests of this hypothesis.

First we consider which variables predict allowance. Column 1 of Table 1 presents estimates from a regression of an allowance indicator (de-meaned by hearing office and day) on the age, race, earnings histories, and health conditions of individuals in our estimation sample. Women, older individuals, whites, those with strong attachment to the labor market, high earners, those represented by a lawyer, and those who did not complete high school are more likely to be allowed benefits. Column 2 presents $t - statistics$ (all standard errors throughout are clustered by judge). It shows that these differences are highly statistically significant. The R^2 shows that the covariates explain 3.9% of the variation in allowance rates.

Our instrumental variable is the judge allowance differential, $j_i\hat{\gamma}_{1,-i}$, de-meaned by hearing office and day. Column 3 presents estimates from a regression of the judge allowance differential on covariates. Column 4 provides $t - statistics$. Of the 22 covariates, two have

TABLE 1: PREDICTORS OF ALLOWANCE AND JUDGE ALLOWANCE DIFFERENTIAL

Covariate	Dependent variable: Allowed Coefficient (1)	 t-statistic (2)	Dependent variable: judge allowance differential Coefficient (3)	 t-statistic (4)
Sex				
Female	0.0290	22.9	0.0002	0.9
Age				
45 to 54	0.0484	37.3	-0.0003	-1.3
55 to 59	0.1379	54.5	-0.0005	-1.0
60 or older	0.1476	49.7	-0.0004	-0.6
Race				
Black	-0.0497	-23.1	0.0001	0.1
Other (non-black, non-white) or unknown	-0.0215	-7.0	-0.0001	0.0
Labor force participation and income				
Average participation rate, years -11 to -2	0.0082	24.9	0.0000	0.1
Average earnings/1,000,000, years -11 to -2 ($2006)	0.9480	10.2	-0.0002	0.0
Represented by lawyer				
Represented by lawyer	0.0743	41.8	0.0008	1.0
Application type				
SSDI	-0.0027	-1.7	-0.0004	-0.6
Education				
High school graduate, no college	-0.0092	-8.8	0.0000	0.0
Some college	-0.0292	-17.3	-0.0010	-1.4
College graduate	-0.0127	-5.6	-0.0004	-0.5
Health conditions (by diagnosis group)				
Neoplasms (e.g., cancer)	-0.0124	-4.4	-0.0016	-3.1
Mental disorders	-0.0153	-7.7	-0.0016	-2.6
Mental retardation	-0.0063	-1.9	-0.0008	-0.8
Nervous system	0.0158	8.6	0.0001	0.2
Circulatory system (e.g., heart disease)	0.0040	2.3	-0.0006	-1.2
Musculoskeletal disorders (e.g., back pain)	0.0036	2.4	0.0000	0.0
Respiratory system	-0.0218	-10.3	-0.0006	-1.0
Injuries	0.0098	5.3	0.0009	1.9
Endocrine system (e.g., diabetes)	0.0215	10.3	-0.0003	-0.5
Standard deviation of dependent variable	0.4293		0.0659	
R^2	0.0389		0.0002	
Number of applicants = 1,779,825, number of judges = 1,497				

Notes: variables allowed and judge allowance differential are demeaned. Standard errors are clustered by judge. Omitted category is male, younger than 45, white, not represented by a lawyer, applying for SSI or SSI and DI concurrently, not a high school graduate, with a health condition other than the those listed above.

TABLE 2: ALLOWANCE RATES, BY DEMOGRAPHICS

	Observations	Allowance rate ALJ stage	Allowance rate 3 years later	Allowance 3 years late Coeff on judge allowance rate	Std. Error	T-ratio	Relative likelihood*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
All groups							
All groups	1,779,825	0.645	0.769	0.764	0.008	101	1.000
Sex							
Male	894,927	0.638	0.763	0.738	0.010	74	0.966
Female	884,898	0.652	0.774	0.791	0.009	84	1.035
Age							
44 or younger	647,528	0.580	0.698	0.898	0.015	60	1.175
45 to 54	754,191	0.644	0.783	0.752	0.010	74	0.983
55 to 59	245,948	0.755	0.866	0.550	0.016	34	0.720
60 or older	132,158	0.762	0.848	0.612	0.023	26	0.801
Race							
White	416,177	0.673	0.791	0.742	0.008	89	0.971
Black	1,154,269	0.586	0.725	0.793	0.015	54	1.037
Other (non-black, non-white) or unknown	209,379	0.608	0.733	0.835	0.019	44	1.092
Labor force participation and income							
Average participation rate, years -11 to -2<70%	688,194	0.581	0.696	0.914	0.013	73	1.197
Average participation rate, years -11 to -2≥70%	1,091,631	0.685	0.814	0.668	0.009	72	0.874
Average earnings, years -11 to -2 ($2006)<$10000	919,519	0.587	0.709	0.886	0.011	78	1.159
Average earnings, years -11 to -2 ($2006)≥$10000	860,306	0.707	0.833	0.635	0.011	60	0.831
Represented by lawyer							
Represented by lawyer	1,136,584	0.684	0.802	0.738	0.009	79	0.965
Not represented by lawyer	643,241	0.576	0.710	0.802	0.013	62	1.049
Application type							
SSDI	673,444	0.696	0.814	0.680	0.012	57	0.890
SSI or Concurrent (both SSDI and SSI)	1,106,381	0.614	0.741	0.817	0.010	80	1.069
Education							
Less than high school	726,027	0.649	0.776	0.741	0.010	75	0.969
High school graduate, no college	771,339	0.647	0.767	0.778	0.010	76	1.018
Some college	197,533	0.615	0.738	0.812	0.016	51	1.062
College graduate	84,926	0.673	0.786	0.715	0.021	34	0.936
Health conditions (by diagnosis group)							
Neoplasms (e.g., cancer)	34,436	0.644	0.762	0.698	0.036	19	0.914
Mental disorders	272,508	0.591	0.759	0.749	0.018	42	0.980
Mental retardation	31,336	0.602	0.813	0.578	0.034	17	0.756
Nervous system	99,666	0.658	0.776	0.711	0.021	34	0.931
Circulatory system (e.g., heart disease)	191,883	0.670	0.787	0.681	0.015	45	0.891
Musculoskeletal disorders (e.g., back pain)	640,712	0.664	0.776	0.785	0.012	68	1.028
Respiratory system	75,079	0.632	0.760	0.757	0.025	31	0.991
Injuries	119,617	0.655	0.748	0.840	0.020	43	1.100
Endocrine system (e.g., diabetes)	86,024	0.661	0.790	0.741	0.022	34	0.970
All other	228,564	0.630	0.740	0.825	0.014	58	1.079
Year assigned to judge							
1990	125,293	0.682	0.830	0.549	0.020	28	0.718
1991	145,136	0.717	0.842	0.564	0.016	36	0.739
1992	170,759	0.719	0.829	0.620	0.015	40	0.812
1993	162,315	0.687	0.792	0.736	0.018	40	0.963
1994	179,567	0.659	0.758	0.802	0.018	44	1.050
1995	197,684	0.629	0.738	0.850	0.016	54	1.113
1996	209,342	0.588	0.715	0.872	0.020	44	1.142
1997	197,951	0.589	0.723	0.852	0.017	49	1.115
1998	202,123	0.608	0.745	0.872	0.015	60	1.142
1999	184,045	0.626	0.768	0.775	0.018	43	1.014

Notes: variables allowed and judge allowance differential are demeaned. Standard errors are clustered by judge.
*Relative likelihood is the ratio of the group specific coefficient on judge allowance rate (what is in column 4) to the full sample coefficient (0.764).

coefficients that are statistically different than 0 at the 95% level. Sex, age, race, previous earnings, past labor market participation, an indicator equal to 1 if the individual is a DI (but not SSI) applicant, an indicator for whether the case is represented by a lawyer, and education all have little explanatory power for whether or not the case was assigned to a lenient judge. All the estimated coefficients are small in comparison to the coefficients on the same variables in the allowance equation. The only statistically significant differences are for mental disorders and neoplasms. Those with mental disorders and neoplasms are assigned to judges who have 0.16% lower allowance rates than average. These coefficients are small, especially in comparison to the coefficients on the same variables in the allowance equation. The R^2 shows that the covariates explain .02% of the variation in judge specific allowance rates. Thus there is little evidence against the hypothesis of random assignment. Random assignment satisfies the independence assumption described in section 4.1. The next section provides some evidence on whether the rank and monotonicity conditions hold.

6.2 First Stage Estimates

Column 1 of table 2 shows the number of observations for different groups of DI cases heard by an ALJ. Column 2 shows the allowance rate at the ALJ stage for that group. Column 3 shows the allowance rate of the group three years after assignment to an ALJ. Columns 2 and 3 show that older individuals, high earners, and those represented by lawyers have relatively high allowance rates.[9] Nevertheless, differences in allowance rates across subgroups are small.

Column 4 shows the estimated first stage regression coefficient $\hat{\lambda_3}$ on the judge allowance differential from equation (3). Column 5 shows the standard error and column 6 the t-statistic. Column 4 shows that the probability of allowance is increasing in the judge allowance differential and column 5 shows that the increase is highly statistically significant for all the subgroups we consider. The estimated value of $\hat{\lambda_3}$ for the full sample is .764, meaning that the probability that case i is allowed 3 years after assignment rises .764% for every 1% increase in the judge allowance differential (which measures the allowance rate on all cases other than case i). The main reason $\hat{\lambda_3}$ is less than 1 is because we use allowance by the ALJ as the measure of the judge allowance differential in table 1, whereas we use allowance three years

[9]This 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.

after assignment as our key measure of allowance in table 2. Many cases denied by an ALJ are later allowed.

Column 4 shows that the estimated coefficient $\hat{\lambda_3}$ 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 an injury. Abadie (2003) shows that the ratio of the group specific estimate of $\hat{\lambda_3}$ relative to full sample estimate of $\hat{\lambda_3}$ is informative for understanding the characteristics of those allowed by a small increase in the ALJ allowance rate. He shows that this ratio yields the relative likelihood that someone with a given characteristic is allowed given a small increase in the allowance rate. Thus, an increase in the allowance threshold of all judges would increase the allowance rate of those with low earnings and injuries more than for other groups, holding the applicant pool and the rest of the re-applications and appeals process constant.

An important implication of the monotonicity assumption described in section 4.1 is that the probability of allowance is non-decreasing in the judge allowance differential for all subgroups of the population. If the allowance rate was rising in the judge allowance differential for some subgroups of the population, but was declining for others, it would show that lenient judges were less likely to allow benefits than strict judges for some types of cases. We do not observe this and thus cannot reject an important implication of the monotonicity assumption. Furthermore, estimates are highly significant, so the rank conditions hold.

6.3 Second Stage: the Effect of Disability Recipiency on Labor Supply

Table 3 presents estimates of the effect of disability recipiency on earnings and labor force participation using both OLS and IV estimators. The first two rows show mean earnings and labor force participation (measured as earnings $>$ \$100) for those allowed and denied benefits, three years after assignment to an ALJ. Row 3 shows the allowance coefficient from a regression of earnings or participation on allowance. Note that the coefficient on allowance is just the difference in earnings or participation between those allowed and those denied. The next row shows the associated standard error. The next rows show OLS and IV estimates of de-meaned (by hearing office and day) earnings or participation on similarly de-meaned allowance. The OLS estimates show that de-meaning does not much affect the results. The IV estimate is the estimate from equation (4). The next row includes the covariates listed in

TABLE 3: ESTIMATED EFFECT OF DI RECIPIENCY ON LABOR SUPPLY

	Dependent Variable: Earnings		*Dependent Variable: Participation*	
	OLS	IV	OLS	IV
Without Covariates:				
Allowed	1442		0.130	
Denied	5345		0.395	
Coef on allowance	-3903		-0.265	
(Std. Error)	(37)		(0.002)	
Coef on demeaned allowance*	-3857	-4059	-0.262	-0.256
(Std. Error)	(34)	(140)	(0.002)	(0.006)
With Covariates:				
Coef on demeaned allowance*	-4247	-4023	-0.271	-0.255
(Std. Error)	(65)	(127)	(0.002)	(0.005)
Lagged labor supply covariates only				
Coef on allowance	-4688		-0.295	
(Std. Error)	(76)		(0.002)	
Non-labor-supply covariates only				
Coef on allowance	-3773		-0.253	
(Std. Error)	(34)		(0.002)	

Notes: N=1,779,825. Standard errors are clustered by judge. Instrument is judge allowance differential.
Earnings, participation, and allowance are measured 3 years after assignment to a judge.
Earnings in 2006 dollars. Participation is and indicator for earnings over $100 in a year.
Covariates are those in Table 1; 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 de-meaned allowance, all variables are de-meaned from the hearing office-day average.

table 1. Parameter estimates are remarkably similar whether using IV or OLS, and whether using additional covariates or not.

Our preferred results are the IV estimates with no covariates. These estimates suggest that those who are allowed benefits earn on average $4,059 per year less than their denied counterparts. IV estimated participation rates for allowed individuals are 25.6% lower than for their denied counterparts. Adding all the covariates listed in table 1 to this specification has only a tiny effect on the estimates. For example, adding covariates to the IV participation equation changes the estimated participation response from 25.6% to 25.5%. Recall that our IV estimation procedure should deliver consistent estimates, with or without covariates. Thus it is reassuring to see that adding covariates barely changes the estimates.

Perhaps the most surprising fact in table 3 is that OLS and IV estimates are so similar. In contrast, Chen and van der Klauww (2008) and Maestas, Mullin and Strand (2011) find the OLS estimates are larger than IV. Our IV estimates are larger than those of both Chen and van der Klauww (2008) and Maestas, Mullin and Strand (2011), although our OLS estimates are smaller. Our OLS estimates are likely smaller because our initial sample is the set of individuals who appealed an ALJ decision. These individuals potentially have weaker attachment to the labor force than the pool of all initial applicants, which is the sample used in those other two papers. However, for all three papers we are estimating labor supply responses for the "marginal applicant", whose condition is severe enough that they have a good chance of allowance, but are not sufficiently disabled that they are guaranteed allowance at the initial stage. Thus it should not be particularly surprising that the our IV estimates are similar to those of Chen and van der Klauww (2008) and Maestas, Mullin and Strand (2011).

Bound (1989) suggests that OLS should overstate the true work disincentive effect of DI, because those who are allowed are on average less healthy and thus less likely to work than those who are not allowed. Differences in labor supply across the two groups is partly due to the effect of DI, but also partly due to the fact that those denied benefits would be more likely to work, even if they were allowed. Consistent with this view, table 2 shows that older individuals have high allowance rates. Tables 4 and 5 show that these individuals are unlikely to work. Moreover, only 16.2% of those allowed benefits in our sample die within 10 years, whereas 12.6% of those denied benefits die within 10 years. However, as pointed out by Bound (1989, 1991), Parsons (1991), and more recent research, those allowed benefits have stronger attachment to the labor market prior to applying for benefits. It is possible that this attachment extends to after when they apply for benefits. Thus it is possible that those allowed benefits are *more* likely to work in the absence of benefit receipt. This would imply that OLS understates the work disincentive effect of DI. Consistent with this view, table 2 shows that those allowed benefits have higher earnings and participation prior to applying. Thus it is an empirical question whether OLS overstates or understates the effect of DI receipt on participation.

The bottom rows of table 3 present OLS earnings and participation estimates with different sets of additional covariates. The table reveals two offsetting biases in the OLS estimates. Recall the the coefficient on allowed when including no covariates is -.265, but is potentially biased up or down. OLS potentially understates the effect (i.e., OLS is biased towards 0) because those allowed benefits have stronger prior attachment to the labor market. Thus, accounting for prior attachment to the labor market should increase the magnitude of the estimated effect. Consistent with this view, accounting for earnings and participation prior to appeal, but nothing else, increases the estimated effect from -.265 to -.295. OLS potentially overstates the effect (i.e., OLS is biased towards -1) because those allowed benefits are older and less healthy. Thus accounting for age and health condition should reduce the magnitude of the effect. Consistent with this view, when we omit labor supply variables, but include all the other variables listed in table 1, the estimated effect declines from -.265 to -.253. Thus there is evidence for the two offsetting effects.

The results in this section are robust to a number of other modifications to sample selection and functional form. We discuss these results in table A3 in the appendix. These robustness checks include: dropping people who die within 3 years after assignment, including people with missing education information, and conditioning on hearing office-years interactions rather than hearing office day interactions.

Table 4 disaggregates the participation responses by demographics, earnings, and health conditions. Column 1 reports mean earnings for allowed individuals, column 2 for denied individuals, column 3 the difference, and column 4 the standard error. Column 5 reports the IV estimate of allowance on earnings and column 6 the standard error. Table 4 shows that the effect of DI allowance on participation is relatively small for college graduates and those with mental disorders, but is larger for high school graduates and those with musculoskeletal problems and injuries. Participation responses are larger in the late 1990s than the early 1990s and early 2000s (recall that participation is measured three years after assignment, so assignment in 1999 refers to participation in 2002), potentially giving evidence that the work disincentive from DI is larger when it is easier to get a job. For most groups, the OLS estimates are very close to the IV estimates. One interesting exception is those with neoplasms. OLS estimates suggest decline in participation of 30.2% in response to allowance, whereas IV suggests a decline of only 19.4%. The low responsiveness of labor supply of those with mental illness is particularly surprising. Mental health is more difficult to monitor than many other health conditions. As a result, some analysts believe that many who claim mental illness are those who are healthy and would have worked in the absence of benefit allowance (Bound and Burkhauser, 1999). This turns out not to be the case.

Table 5 disaggregates the earnings responses by demographics, earnings, and health conditions. Results from this table are consistent with the results in table 4. For all groups, allowance reduces earnings. Earnings estimates tend to be less precise than estimates for

TABLE 4: ESTIMATED EFFECT OF DI RECIPIENCY ON PARTICIPATION, DISAGGREGATED

	OLS				IV	
	Allowed	Denied	Difference	Std. Error	Difference	Std. Error
All groups						
All groups	0.130	0.395	-0.265	0.002	-0.256	0.006
Sex						
Male	0.133	0.403	-0.270	0.002	-0.263	0.009
Female	0.127	0.386	-0.260	0.002	-0.250	0.008
Age						
45 or younger	0.174	0.467	-0.293	0.002	-0.290	0.009
45 to 54	0.116	0.359	-0.244	0.002	-0.254	0.009
55 to 59	0.094	0.282	-0.189	0.003	-0.248	0.019
60 to 64	0.099	0.179	-0.080	0.003	-0.069	0.023
Race						
Black	0.138	0.425	-0.287	0.003	-0.252	0.014
White	0.133	0.393	-0.260	0.002	-0.265	0.008
Other (non-black, non-white) or unknown	0.097	0.343	-0.246	0.004	-0.221	0.016
Labor force participation and income						
Average participation rate, years -11 to -2<70%	0.065	0.264	-0.199	0.002	-0.176	0.009
Average participation rate, years -11 to -2≥70%	0.165	0.531	-0.365	0.002	-0.327	0.012
Average earnings, years -11 to -2 ($2006)<$10000	0.087	0.325	-0.239	0.002	-0.202	0.008
Average earnings, years -11 to -2 ($2006)≥$10000	0.169	0.525	-0.356	0.002	-0.335	0.014
Represented by lawyer						
Represented by lawyer	0.130	0.400	-0.270	0.002	-0.274	0.008
Not represented by lawyer	0.129	0.389	-0.260	0.002	-0.226	0.010
Application type						
SSDI	0.175	0.429	-0.254	0.002	-0.277	0.016
SSI or SSI/SSDI concurrent	0.100	0.380	-0.280	0.002	-0.244	0.008
Education						
Less than high school	0.076	0.327	-0.251	0.002	-0.230	0.009
High school graduate, no college	0.148	0.425	-0.277	0.002	-0.279	0.009
Some college	0.210	0.479	-0.269	0.003	-0.261	0.019
College graduate	0.254	0.472	-0.219	0.004	-0.179	0.031
Health conditions (by diagnosis group)						
Neoplasms (e.g., cancer)	0.155	0.457	-0.302	0.006	-0.194	0.043
Mental disorders	0.146	0.383	-0.237	0.003	-0.202	0.016
Mental retardation	0.094	0.322	-0.227	0.007	-0.282	0.048
Nervous system	0.140	0.392	-0.251	0.004	-0.237	0.027
Circulatory system (e.g., heart disease)	0.111	0.367	-0.256	0.003	-0.250	0.018
Musculoskeletal disorders (e.g., back pain)	0.136	0.419	-0.283	0.002	-0.285	0.009
Respiratory system	0.089	0.363	-0.274	0.004	-0.254	0.023
Injuries	0.147	0.468	-0.320	0.003	-0.367	0.022
Endocrine system (e.g., diabetes)	0.089	0.324	-0.235	0.004	-0.224	0.024
All other	0.128	0.365	-0.237	0.003	-0.211	0.015
Year assigned to judge						
1990	0.100	0.323	-0.223	0.004	-0.234	0.023
1991	0.108	0.332	-0.224	0.004	-0.186	0.021
1992	0.115	0.362	-0.247	0.004	-0.277	0.020
1993	0.123	0.370	-0.246	0.004	-0.231	0.018
1994	0.137	0.395	-0.259	0.004	-0.293	0.015
1995	0.142	0.410	-0.268	0.003	-0.276	0.015
1996	0.141	0.431	-0.289	0.003	-0.273	0.014
1997	0.147	0.424	-0.277	0.003	-0.252	0.013
1998	0.140	0.410	-0.270	0.003	-0.265	0.014
1999	0.134	0.386	-0.252	0.003	-0.222	0.017

Notes: OLS estimates are in levels with no covariates.
IV estimates use demeaned variables and the judge allowance differential as the instrument
Allowance and participation measured 3 years after assignment to an ALJ. Standard errors clustered by judge.

TABLE 5: ESTIMATED EFFECT OF DI RECIPIENCY ON EARNINGS, DISAGGREGATED

	OLS				IV	
	Allowed	Denied	Difference	Std. Error	Difference	Std. Error
All groups						
All groups	1442	5345	-3903	37	-4059	140
Sex						
Male	1731	6231	-4500	48	-4695	234
Female	1153	4405	-3252	36	-3438	174
Age						
45 or younger	2085	6251	-4166	46	-4698	228
45 to 54	1286	5026	-3740	45	-4038	205
55 to 59	872	3728	-2855	69	-3218	427
60 to 64	747	1773	-1026	59	-1496	460
Race						
Black	1193	5175	-3982	48	-3675	249
White	1581	5637	-4056	44	-4383	197
Other (non-black, non-white) or unknown	1100	4431	-3331	67	-3143	381
Labor force participation and income						
Average participation rate, years -11 to -2<70%	521	2654	-2132	24	-2025	171
Average participation rate, years -11 to -2≥70%	1937	8124	-6186	51	-5847	287
Average earnings, years -11 to -2 ($2006)<$10000	578	3025	-2448	23	-2134	165
Average earnings, years -11 to -2 ($2006)≥$10000	2227	9661	-7434	66	-6888	370
Represented by lawyer						
Represented by lawyer	1461	5474	-4013	41	-4431	190
Not represented by lawyer	1402	5189	-3787	47	-3459	239
Application type						
SSDI	2341	7649	-5307	70	-5787	418
SSI or SSI/SSDI concurrent	840	4337	-3497	34	-3138	168
Education						
Less than high school	638	3798	-3160	37	-3086	202
High school graduate, no college	1584	5889	-4305	44	-4750	207
Some college	2577	6953	-4375	74	-4077	479
College graduate	4478	9245	-4767	187	-4368	1272
Health conditions (by diagnosis group)						
Neoplasms (e.g., cancer)	2332	6751	-4420	179	-2038	1323
Mental disorders	1350	4607	-3257	57	-2844	318
Mental retardation	545	3120	-2575	107	-2920	1079
Nervous system	1501	5425	-3924	95	-3926	723
Circulatory system (e.g., heart disease)	1178	4823	-3645	67	-3294	385
Musculoskeletal disorders (e.g., back pain)	1619	5974	-4355	50	-4942	245
Respiratory system	774	4377	-3603	94	-3177	477
Injuries	2070	7178	-5108	94	-6606	578
Endocrine system (e.g., diabetes)	741	3727	-2986	77	-2589	437
All other	1411	4850	-3439	59	-3634	344
Year assigned to judge						
1990	851	4208	-3357	93	-2848	516
1991	1078	4374	-3296	99	-3360	650
1992	1154	4692	-3538	88	-4205	418
1993	1213	4460	-3247	76	-4017	318
1994	1444	4803	-3359	67	-3748	350
1995	1661	5415	-3754	70	-4317	357
1996	1716	5976	-4260	68	-4366	348
1997	1773	6016	-4243	71	-3766	316
1998	1704	5991	-4287	71	-4745	326
1999	1566	5555	-3989	71	-4078	367

Notes: OLS estimates are in levels with no covariates
IV estimates use demeaned variables and the judge allowance differential as the instrument
Allowance and earnings measured 3 years after assignment to an ALJ. Standard errors clustered by judge.
Earnings in 2006 dollars.

participation, however.

6.4 Dynamics of the Response

This section shows the dynamics of the response of both earnings and participation. Figure 3 shows the earnings and participation responses to benefit allowance. The top left panel shows annual earnings for those who are allowed and those who are denied DI benefits by an ALJ both before and after the date of assignment to a judge. Prior to assignment, those who are allowed benefits have higher earnings than their denied counterparts. By the year of assignment, earnings for allowed and denied individuals are similar. Three years after assignment, earnings of those allowed benefits average \$1,490 while earnings of those denied average \$3,842, a difference of \$2,352. Differences in earnings between those allowed and those denied emerge rapidly, are very stable 2-5 years after assignment, and decline slowly thereafter.[10]

Consistent with the evidence on earnings, the bottom-left panel of figure 3 shows that 10 years prior to assignment, those who are subsequently allowed benefits have participation rates that are seven percentage points higher than those subsequently denied benefits. Three years after the date of assignment, those who are allowed benefits have participation rates that are 17 percentage points lower than those who are denied. Afterwards, the differences between the two groups narrow slightly.

The right-hand panels show IV estimates of earnings and labor force participation of allowed and denied individuals both before and after assignment to a judge. We estimate the effect of allowance for each year relative to the assignment year, as predicted by the judge allowance differential. Using the estimation procedure described in section 4.2 we can estimate the effect of DI receipt on earnings or participation at any point in time (at least for those affected by the instrument). The vertical difference between the allowed and denied lines is this estimated effect. In order to make the figures more concrete, we also present the level of earnings and participation. To identify the level, we make the additional assumption that $E[\phi_{i\tau}]$ for those affected by the instrument is the same as $E[\phi_{i\tau}]$ for those not affected by the instrument: see appendix C for details. This assumption is untestable, although section

[10]Some care must be taken in interpreting the decline in earnings of denied individuals 5 years after assignment because after 5 years, 7% of all sample members are at least 65 and after 10 years 21% are at least 65. These people are eligible for full Social Security benefits, even if they were initially denied.

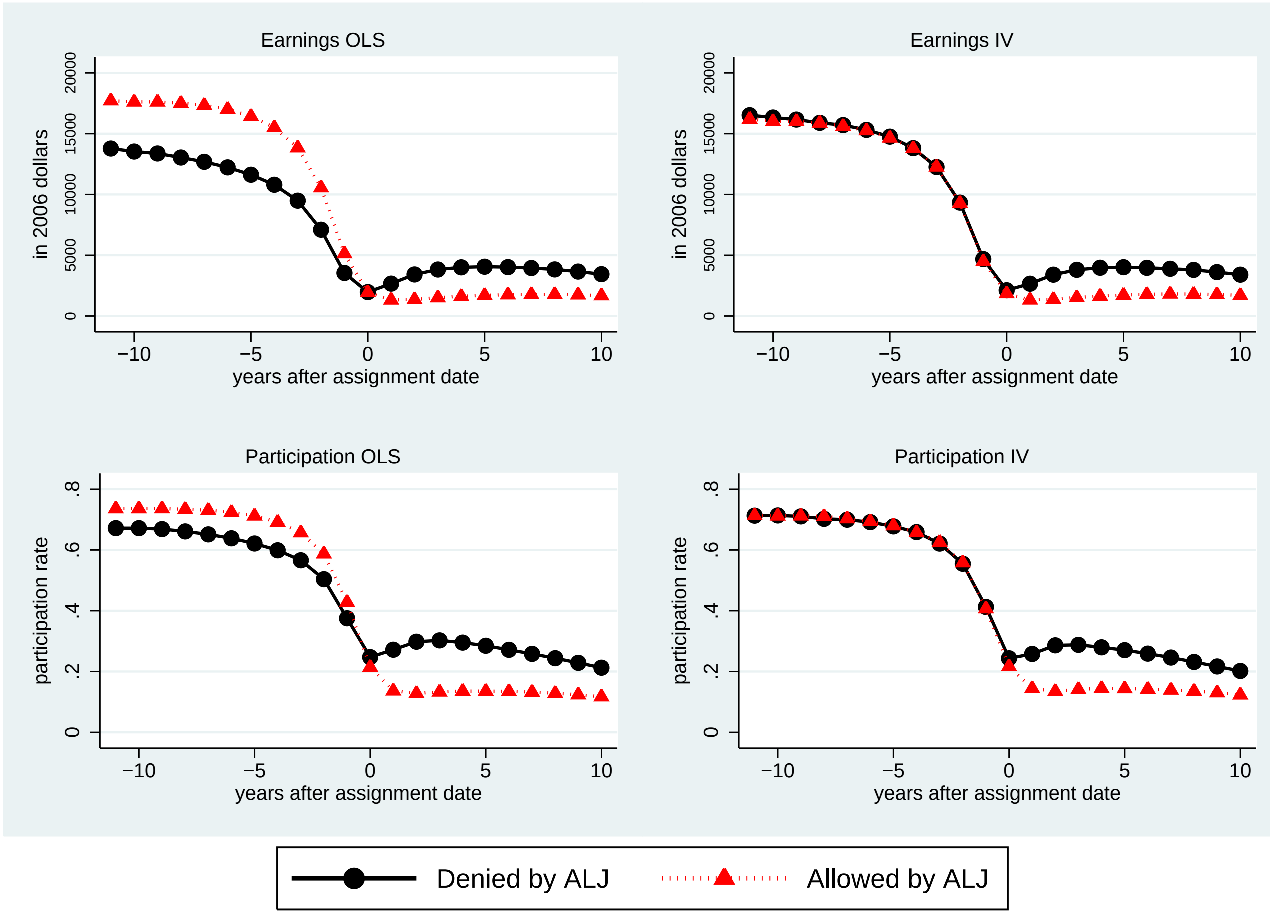

Figure 3: Dynamics of earnings and participation, allowed versus denied by ALJ.

6.6 gives evidence that $E[\phi_i]$ does not vary much over the support of our data.[11]

IV estimates for those allowed versus denied are virtually identical prior to assignment. Recall that the difference in participation between the two groups is that predicted by the instrument of the judge allowance differential. A difference of 0 prior to assignment is a reassuring result, as it shows that we are unable to predict labor supply prior to assignment using our instrument. This is an important testable implication of the independence assumption.

However, after assignment, earnings and participation of allowed individuals are lower. The top right panel shows that three years after the time of assignment, the difference in earnings between the two groups is $2,314 (virtually identical to the OLS estimate) and remains very stable thereafter. Similarly, the bottom right panel shows that three years after

[11] In contrast to our findings, Maestas et al. (2011) do find variability in $E[\phi_{i\tau}]$ across the support of their data.

assignment the difference in participation between the two groups is 14.8%, and does not change much thereafter. The standard errors are tiny and thus omitted. For example, the standard error on the effect of allowance on participation averages less than 1% when using either OLS or IV.

Note that the IV estimate of the effect of allowance on earnings 3 years after allowance is smaller in figure 3 ($2,314) than in table 3 ($4,059). The difference arises because figure 3 uses allowance by the ALJ, whereas table 3 uses allowance 3 years after assignment to the ALJ. Section 6.5 discusses the difference between allowance by an ALJ and allowance at any point in time.

6.5 Appeals, Re-applications, and Subsequent Allowance

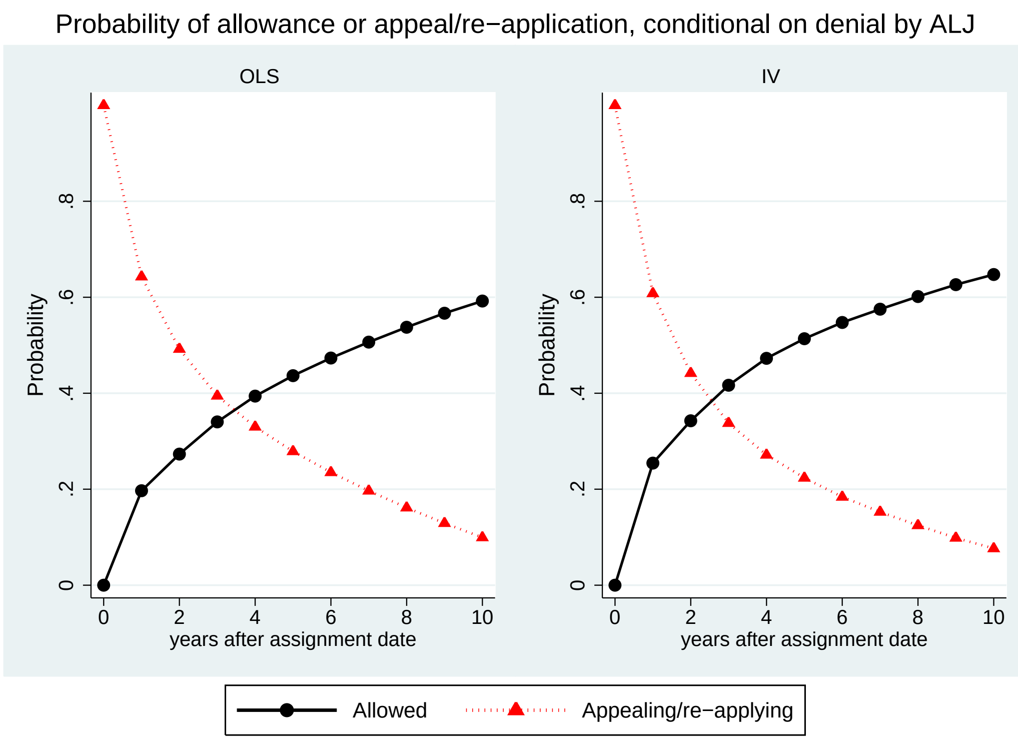

Figure 4: Allowance and Appeals/Re-applications following denial by ALJ.

The left panel of figure 4 shows the share of denied (at the ALJ stage) individuals who are reapplying/appealing and allowed relative to when they are assigned to a judge.[12] It

[12]We use data from ACAPS and LOTS to identify denied applicants who successfully appealed at either the Appeals Council or the Federal Court level. We use data from SSA 831 files, MBR (Master Beneficiary

shows that 35% of all applicants denied by an ALJ were allowed benefits within three years. Furthermore, many initially denied individuals continue to reapply or appeal for many years after their initial denial. Three years after assignment to an ALJ, 40% of all individuals denied benefits are still in the process of appealing or reapplying for benefits. Combined, fully 75% of those denied by an ALJ are either allowed or in the process of appealing 3 years after assignment to an ALJ.

The right panel of figure 4 presents the share of initially denied individuals who are allowed benefits or are still in the process of reapplying/appealing relative to when they are assigned to a judge, where the shares are instrumented using the judge allowance differential. To do this we estimate the effect of predicted ALJ allowance on allowance and appeals at future points in time, as well as the procedure in appendix to C to infer the effect of ALJ denial on future allowance.[13] Thus the left panel uses OLS and the right panel uses IV, where initial denial is instrumented using the judge allowance differential. Those affected by the instrument are likely the marginal cases who have a better chance of final allowance than others denied benefits. For this reason we might think that subsequent allowance rates of those initially denied would be higher when instrumented. In fact, this is the case, although the OLS estimates and the IV estimates are similar. For example, the right panel figure 4 shows that for those initially denied benefits, the IV estimate of allowance is 42% three years after assignment, versus 35% from the OLS estimates.

Sections 6.4 and 6.5 show that most denied applicants do not work, but engage in re-applications and appeals until they get DI benefits. This has an important effect on our main estimated effects. Table 3 shows that DI benefit allowance reduces earnings \$4,059 per year when measuring earnings and allowance three years after assignment to an ALJ. However, DI benefit allowance reduces earnings \$4,915 per year when measuring earnings and allowance five years after assignment to an ALJ.

Record), and SSR (Supplemental Security Record) to identify denied applicants who reapplied for benefits and were allowed at either the DDS, Reconsideration, ALJ, Appeals, or Federal Court level stage.

[13]Using the full sample, we regress de-meaned allowance on a set of wave dummies and predicted de-meaned ALJ allowance × wave dummies (where allowance is predicted using the judge allowance differential). The estimated coefficient on allowance×wave measures increased probability of allowance at a given wave conditional on initial denial. Next, we regress de-meaned appeal on a set of wave dummies and predicted de-meaned ALJ allowance interacted with wave dummies (where allowance is predicted using the judge allowance differential). The estimated coefficient on allowance×wave measures increased probability of allowance at a given wave conditional on initial denial. The right panel of figure 4 plots the coefficient on predicted allowance×wave for both the allowance and appeal equations.

6.6 Estimates of the Distribution of Labor Supply, Allowance, and Appeal Responses: Marginal Treatment Effects

Using the the Marginal Treatment Effects approach described in section 4.3 and appendix B, this section shows how DI benefit allowance affects the distribution of labor supply, subsequent allowance, and appeals.

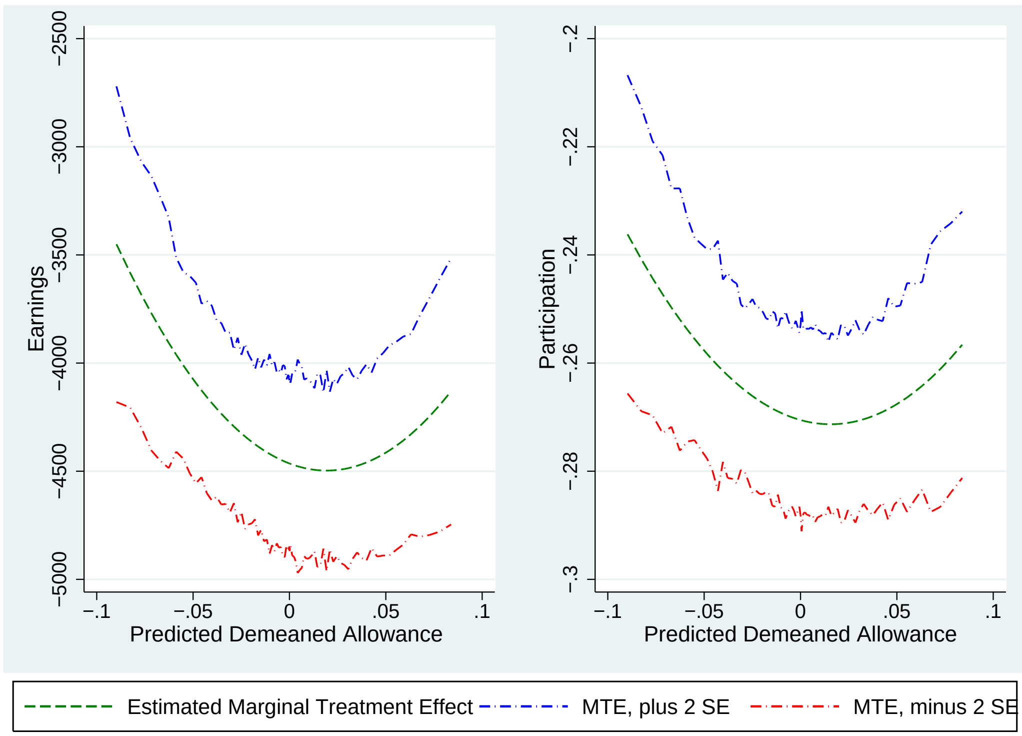

Figure 5: Earnings and participation decline when allowed for marginal applicant.

The left panel of figure 5 shows the earnings decline and the right panel shows the participation decline of the marginal case when allowed (i.e., the Marginal Treatment Effect). We use third order polynomials for both the instrument and the endogenous variable (de-meaned allowance) when estimating equations (6) and (7). Both Akaike's information criterion and the Bayesian information criterion reject quadratic and quartic specifications in favor of the cubic. Furthermore, results from the quartic specification are very similar to the cubic specification. Since polynomial smoothers have poor endpoint properties, we show estimated MTEs over the middle 90% of the distribution of the judge allowance differential. Based upon Monte Carlo experiments, we found our procedure produced little bias over the middle 90% of the distribution. Figure 5 also shows bootstrapped 95% confidence intervals.

On average, annual earnings and participation decline \$4,300 and 26% in response to benefit allowance, similar to the main estimates reported in table 3. However, there is heterogeneity in the declines. The earnings decline is \$3,451 for the marginal applicant heard by an ALJ who is stricter than 95% of all judges, whose decisions lead to allowance rates that are nine percentage points below the average three years after assignment. The earnings decline is \$4,131 for the marginal applicant heard by an ALJ who is more lenient than 95% of all judges, whose decisions lead to allowance rates that are eight percentage points above the average three years after assignment. When judge specific allowance rates rise, the labor supply response of the marginal case also rises. This result is consistent with the notion that as allowance rates rise, more healthy individuals are allowed benefits. These healthier individuals are more likely to work when not receiving DI benefits and thus their labor supply response to DI receipt is greater. Nevertheless, the differences in the earnings response are not statistically significant and is modest in size.

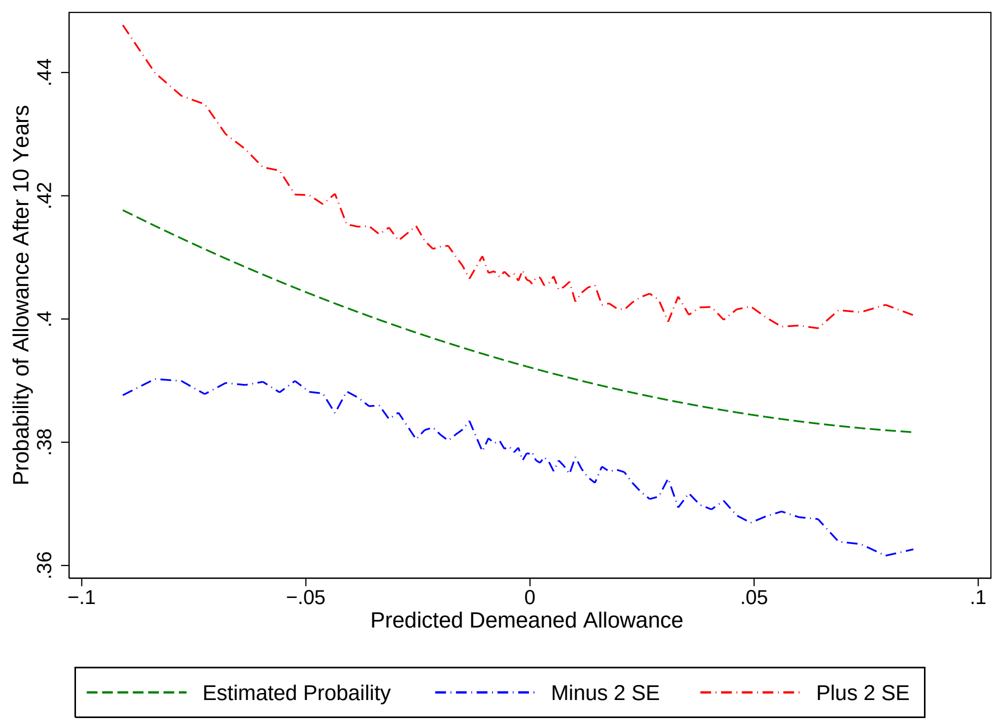

Figure 6: Marginal applicant's allowance probability 10 years after assignment conditional on not allowed 3 years after assignment to an ALJ.

Figure 6 shows how allowance three years after assignment to an ALJ affects allowance 10 years afterwards. It shows that 40% of those not allowed three years after assignment were allowed benefits 10 years after assignment. For marginal applicants assigned to lenient judges and are not allowed three years after assignment, the probability of allowance 10 years after assignment is .38. For those assigned to strict ones it is .42. Recall that marginal applicants assigned to lenient judges and not allowed benefits are healthier than those assigned to strict judges. Thus it is unsurprising that they are less likely to be allowed benefits in the future. What is remarkable, however, is that conditional on being denied 3 years after assignment, 40% have been allowed benefits 10 years after assignment.

6.7 Elasticity of Labor Supply with Respect to the After-Tax Wage

In this section we present estimates of the effect of DI on the after-tax (and after DI benefit) wage, as well as the earnings and participation elasticity with respect to the after-tax wage. Table 6 shows participation and earnings elasticities with respect to the after-tax wage, which we calculate as follows:

$$\varepsilon_{y,w} = \frac{(E[y_i|A_i = 0] - E[y_i|A_i = 1])/(E[y_i|A_i = 0] + E[y_i|A_i = 1])}{(E[w_i|A_i = 0] - E[w_i|A_i = 1])/(E[w_i|A_i = 0] + E[w_i|A_i = 1])} \tag{9}$$

where $E[y_i|A_i = 0]$ is the average outcome variable (either mean earnings or participation) of denied individuals and $E[y_i|A_i = 1]$ is the average outcome variable for allowed individuals. $E[w_i|A_i = 0]$ is the average after-tax wage for denied individuals and $E[w_i|A_i = 1]$ is the average after-tax wage for allowed individuals. The after-tax wage is defined as the income gain from wage earnings plus DI benefits (net of federal, state and payroll taxes) when working. Appendix B presents the details of how we estimate after-tax wages.

We first predict the distribution of pre-tax wages for everyone in the sample. The first row of table 6 shows that the average predicted pre-tax wage of workers in our sample is \$11,047. Next, we use Social Security earnings histories, the year, and state of residence to calculate DI/SSI benefits for everyone in the sample. The second row shows that the average DI/SSI benefit is \$9,023. The third row shows the DI/SSI benefit reduction resulting from high earnings. People who are allowed benefits will lose most of their benefits if they work. The fourth column shows that the average Federal, State, and payroll tax paid by those working is \$2,081. The fifth row is after-tax income, which is labor income plus the DI/SSI

TABLE 6: EARNINGS AND PARTICIPATION ELASTICITIES

	Means				Allowed versus Denied Percent Change/100	Elasticity
	Allowed		Denied			
	working	not working	working	not working		
Pre Tax Wage Income	11,047	0	11,047	0		
DI/SSI benefit if Allowed	9,525	9,525	0	0		
DI/SSI benefit reduction	4,572	0	0	0		
Taxes	2,081	0	2,081			
After Tax Income*	13,915	9,525	8,966	0		
After Tax Wage**	4,390		8,966		0.64	
Earnings	1,412		5,471		1.19	1.86
Participation	0.135		0.391		0.98	1.53

Notes: Earnings and Participation estimates are from Table 3
Elasticity is an arc elasticity: see equation (12)
*After Tax Income is sum of pre-tax wage income and DI/SSI benefit, less DI/SSI benefit reduction and taxes
**After tax wage = after tax income if working - after tax income if not working
2006 dollars.

benefit, less DI/SSI reductions and taxes. The sixth row shows the average after-tax wage, defined as the difference between the after-tax income if working and the after-tax income if not working. The after-tax wage is \$8,966 on average for those who are denied benefits and is \$4,599 for those allowed benefits. Because most DI beneficiaries who are working earn above the SGA level, most people who are allowed benefits will lose their DI benefit if they work. Thus, most of the gain from working is lost when the individual has been allowed DI benefits. We take estimates of earnings and participation declines when allowed (i.e., $E[y_i|A_i = 0] - E[y_i|A_i = 1]$) from table 4 and use the procedure in section C to infer $E[y_i|A_i = 1]$ and $E[y_i|A_i = 0]$. Table 6 shows that the implied earnings elasticity is 1.9 and participation elasticity is 1.5. While our estimates suggest that most DI/SSI applicants would not work even if denied benefits, labor supply is elastic for this group of individuals.

In order to infer a labor supply elasticity with respect to the after-tax wage from the labor supply response to DI allowance, we make two strong assumptions. First, we assume that individuals are only responding to current work incentives and not future incentives. However, individuals must keep their earnings below the SGA level in order to appeal or reapply for benefits. Therefore, the low earnings level of denied applicants may be caused by the incentives to keep earnings low in order to appeal or to reapply for benefits. Thus we are overstating the percent difference in the present value of future after-tax wages and understating the labor supply elasticity. To better assess this issue, we measure the labor supply response to allowance five years after allowance. Figures 1 and 3 show that after five years most DI/SSI applicants have either received benefits or have given up on the application process. Five years after assignment to an ALJ, the participation elasticity is 1.6, slightly higher than the elasticity three years after assignment.

Second, we omit the value of health insurance benefits from both work and from DI/SSI receipt. When individuals lose their DI and SSI benefits due to high earnings, they also typically lose their Medicare and Medicaid health insurance benefits. Thus the percent change in the after-tax wage is likely larger and the true labor supply elasticity is smaller than what we report in table 6. As such, our two strong assumptions lead to two potentially important, but offsetting, biases. Interestingly, our estimates are similar to those of Kostøl and Mogstad (2012).

7 Conclusion

This paper estimates the effect of Disability Insurance receipt on labor supply. Using instrumental variables procedures, we address the fact that those allowed benefits are a selected sample. We find that benefit receipt reduces labor force participation by 26 percentage points three years after a disability determination decision, although the reduction is smaller for those over age 55, college graduates, and those with mental illness. OLS estimates are similar to instrumental variables estimates. The participation elasticity with respect to the after-tax wage is 1.5. Over 60% of those denied benefits are allowed benefits within 10 years, showing that most applicants apply, re-apply, and appeal until they get benefits.

References

[1] Abadie, A. (2003), "Semiparametric instrumental variable estimation of treatment response models", *Journal of Econometrics* 113, 231-263.

[2] Autor, D. and M. Duggan (2006), "Supporting Work: A Proposal for Modernizing the U.S. Disability Insurance System", *The Hamilton Project*.

[3] Autor, D. and M. Duggan (2010), "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding", *Journal of Economic Perspectives* 20(3), 71-96.

[4] Autor, D., Maestas, N., K. Mullen, and A. Strand (2011), "Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation and Earnings of Disability Applicants", *manuscript*.

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

[6] Benitez-Silva, H., M. Buchinsky, and J. Rust (2011), "Induced Entry Effects of a $1 for $2 Offset in SSDI Benefits", *manuscript*.

[7] Bjorklund, A. and R. Moffitt (1987), "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models", *Review of Economics and Statistics* 69(1), 42-49.

[8] Burkhauser, R. and M. Daly (2011) *The Declining Work and Welfare of People with Disabilities: What Went Wrong and a Strategy for Change* American Enterprise Institute Press.

[9] Bound, J. (1989), "The Health and Earnings of Rejected Disability Insurance Applicants", *The American Economic Review* 79(3), 482-503.

[10] Bound, J. (1991), "The Health and Earnings of Rejected Disability Insurance Applicants: Reply", *The American Economic Review* 81(5), 1427-1434.

[11] Bound, J. and R.V. Burkhauser (1999), "Economic analysis of transfer programs targeted on people with disabilities", In: Ashenfelter, O. and D. Card (Eds.), *Handbook of Labor Economics*, Vol. 3, Part 3, 3417-3528.

[12] Bound, J., T. Stinebrickner, and T. Waidmann (2010), "Health, Economic Resources and the Work Decisions of Older Men", *Journal of Econometrics* 156(1), 106-129.

[13] Chen, S. and W. van der Klaauw (2008), "The Effect of Disability Insurance on Labor Supply of Older Individuals in the 1990s", *Journal of Econometrics* 142(2), 757-784.

[14] David, M., R. Little, M. Samuhel, and R. Triest (1986), "Alternative Methods for CPS Income Imputation", *Journal of the American Statistical Association* 81(393), 29-41.

[15] De Jong, P., M. Lindeboom, B. van der Klaauw (2011), "Screening Disability Insurance Applications", *Journal of the European Economic Association* 9(1), 106-129.

[16] Duggan, M., and S. Imberman (2008), "Why Are the Disability Rolls Skyrocketing? The Contribution of Population Characteristics, Economic Conditions, and Program Generosity", In: Ashenfelter, O. and D. Card (Eds.), *Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly*, , 337-379.

[17] French, E. and J. Jones (2011), "The Effects of Health Insurance and Self-Insurance on Retirement Behavior", *Econometrica* 79(3), 693-732.

[18] French, E. and C. Taber (2011), "Identification of Models of the Labor Market", In: Ashenfelter, O. and D. Card (Eds.), *Handbook of Labor Economics*, Vol. 4, Part 1, 537-617.

[19] Hausman, J., W. Newey, T. Woutersen, J. Chao, and N. Swanson (2007) "Instrumental Variable Estimation with Heteroskedasticity and Many Instruments", *manuscript.*

[20] Heckman, J. and E. Vytlacil (2005), "Structural Equations, Treatment Effects, and Econometric Policy Evaluation", *Econometrica* 73(3), 669-738.

[21] Heckman, J., S. Urzua, and E. Vytlacil (2006), "Understanding Instrumental Variables in Models with Essential Heterogeneity", *Review of Economics and Statistics* 88(3), 389-432.

[22] Hu, J., K. Lahiri, D. Vaughan, and B. Wixon (2011), "A Structural Model of Social Security's Disability Determination Process", *The Review of Economics and Statistics* 83(2), 348-361.

[23] Imbens, G. and J. Angrist (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica* 62(2), 467-475.

[24] Kostl, A. and M. Mogstad (2012), "How Financial Incentives Induce Disability Insurance Recipients to Return to Work", *manuscript.*

[25] Lahiri, K., J. Song, and B. Wixon (2008), "A model of Social Security Disability Insurance using matched SIPP/Administrative data", *Journal of Econometrics* 145(1-2), 4-20.

[26] Livermore, G., D. Stapleton, and M. O'Toole (2011), "Health Care Costs Are a Key Driver of Growth in Federal and State Assistance to Working-Age People with Disabilities, *Health Affairs* 30(9), 1664-1672.

[27] Low, H. and L. Pistaferri (2011), "Disability Risk, Disability Insurance and Life Cycle Behaviour", *NBER Working Paper 15962.*

[28] Maestas, N., K. Mullen, and A. Strand (2011), "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt", *manuscript.*

[29] Meyers, R., *Social Security,* 1993.

[30] Paletta, D., "Disability-Claim Judge has Trouble Saying 'No'; Near Perfect Approval Record; Social Security Program Strained", *Wall Street Journal*, May 19, 2011

[31] Parsons, D.O. (1991), "The Health and Earnings of Rejected Disability Insurance Applicants: Comment", *American Economic Review* 81(5), 1419-1426.

[32] Social Security Advisory Board, *Improvind the Social Security Administration's Hearing Process*, Sept. 2006. Available at `http://www.ssab.gov/documents/HearingProcess.pdf.`

[33] Taylor, M. (2007), "Refugee Roulette in an Administrative Law Context: the Deja Vu of Decisional Disparities in Agency Adjudication", *Stanford Law Review* 60(2), 475-502.

[34] U.S. Social Security Administration, *Social Security Bulletin Annual Statistical Supplement*, various years.

[35] U.S. Social Security Administration, *Annual Statistical Report on the Social Security Disability Insurance Program*, 2007.

[36] U.S. Social Security Administration, *Hearings, Appeals and Litigation Law Manual (HALLEX)*, 2009.

[37] von Wachter, T., J. Song, and J. Manchester (2011), "Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program", *The American Economic Review*, forthcoming.

Appendix A: Data Appendix

We use the universe of all DI appeals heard by ALJs, 1990-1999. We use data from the Office of Hearings and Appeals Case Control System (OHACCS), the Hearing Office Tracking System (HOTS), the Appeals Council Automated Processing System (ACAPS), the Litigation Overview Tracking System (LOTS), the SSA 831 file, SSA Master Earnings file (MEF), the Master Beneficiary Record (MBR), the Supplemental Security Record (SSR), and the SSA Numerical Identification (NUMIDENT) file.

The OHACCS data contain details of Social Security DI and SSI cases adjudicated at the ALJ level (and also contain limited information on cases heard at the Appeals Council, Federal or Supreme Court). In addition to SSI and DI, they include cases involving Retirement and Survivors Insurance as well as Medicare Hospital insurance. We keep only the SSI and DI cases. The OHACCS data are used for administering DI and SSI cases, and are thus very accurate. The OHACCS data include information on the judge assigned to the case, the hearing office, the date of assignment, and the outcome of the case (such as allowed or denied). It also has data on the claimant's Social Security number, and type of claim (DI versus SSI). The data include all cases filed in 1982 to present. Because our earnings data go back to 1980, and we use earnings data 10 years prior to assignment, we use OHACCS data 1990-2009.

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

OHACCS also contains Appeals Council records. However, data on Appeals Council decisions are sometimes missing from OHACCS. Thus we use the Appeals Council Automated Processing System (ACAPS) data to track actions on cases heard at the Appeals Council level. ACAPS is the Appeals Council's data for administration of cases.

The Litigation Overview Tracking System (LOTS) data are used for administration of cases that are heard at the Federal or Supreme Court level. These data provide information on which cases that were denied at the Appeals Council level were appealed at the Federal

Court level. We combine the LOTS data with information provided by the Federal Court to determine whether the cases was eventually allowed or denied.

The SSA 831 data have information on the details of the DI application received at the Disability Determination Service. The data include information on the type of application (whether DI or SSI or concurrent) and whether the claim is on one's own earnings history or on the history of a spouse or parent. It also has all the information relevant for determining whether the application should be allowed, either through a medical listing or the vocational grid. Thus we have detailed medical information, such as the health condition of the individual. Because of the vocational grid, we have information on age, education, industry and occupation. We also have some other demographic information such as sex. Since a new 831 record is established whenever a new application is filed and adjudicated, we use information in the 831 file to identify those who reapplied for benefits.

The Master Earning File (MEF) includes annual longitudinal earnings data for the US population. It includes not only individuals' annual Social Security covered earnings from 1951 to the present (which we use to calculate the Primary Insurance Amount for DI benefits), but also individuals' annual wages directly taken from the W-2 starting from 1978. We use data back to 1981. Wage earnings are not top-coded, but self-employment earnings are top coded until 1992. Our earnings measure is the sum of wage earnings and self employment earnings, which we topcode at $200,000 per year.

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

Lastly, we use the SSA NUMIDENT for information on date of death. The NUMIDENT file includes information from the Social Security Number application form such as name, date of birth and Social Security number. Once the individual dies, the date of death is placed on the file. We treat individuals who die as missing, although we found that this assumption does not affect our results.

For Figure 1 and A1 we use all cases filed 1989-1999. We include all primary disability – auxiliary benefit claimants (i.e., child and spouse) are excluded. We make no other sample restrictions for these cases. For all other figures and tables, we begin with the universe of all cases adjudicated by an ALJ and make the following sample restrictions, described in Table A1:

1. We drop all Medicare cases. These Medicare cases are typically disputes over whether Medicare will pay for certain medical treatments.

2. We drop all remand cases (cases sent to Appeals Council, then sent back to the hearing office). We drop these because this would lead to double counting of cases, as a remand is a case that was already heard by an ALJ.

3. We drop cases with a missing Social Security number. This leaves us with 3,525,787 cases for 1990-1999.

4. We drop all cases younger than 35 or older than 64.

5. We drop cases with missing judge or hearing office information.

6. We drop cases that were previewed prior to being assigned to a judge. These cases are extremely likely to be critical cases that are reviewed by a senior attorney.

7. We drop cases where the claim is against the earnings record of a spouse or parent.

8. We drop cases with missing education data. This leaves us with 1,779,825 cases.

Table A2 presents sample means.

TABLE A1: SAMPLE SELECTION

	Sample size
Original sample	3,525,787
Number of drops	
(1): Age at assignment <35 or >64	792,939
(2): Missing judge or hearing office information	174
(3): case is pre-viewed	794,470
(4): DI Child case	30,221
(5): Survivor case	3,564
(6): Missing education data	123,911
(7): Judge handled fewer than 50 cases	683
total number of sample dropped (sum of drops 1-7)	1,745,962
Remaining sample	1,779,825

Robustness Checks

Table A3 provides robustness checks. The first row shows estimates from our benchmark model, which are also presented in table 3. The benchmark model estimates the effect of

TABLE A2: MEANS

Female	0.497
Age	
45 or younger	0.364
45 to 54	0.424
55 to 59	0.138
60 to 64	0.074
Race	
Black	0.234
Other (non-black, non-white) or unknown	0.118
Labor force participation and income	
Average participation rate, years -11 to -2≥70%	0.922
Average earnings, years -11 to -2 ($2006)≥$10000	0.483
Not represented by lawyer	0.639
SSDI (not SSI or SSI/SSDI concurrent)	0.378
Education	
Less than high school	0.408
High school graduate, no college	0.433
Some college	0.111
College graduate	0.048
Health conditions (by diagnosis group)	
Neoplasms (e.g., cancer)	0.128
Mental disorders	0.019
Mental retardation	0.153
Nervous system	0.018
Circulatory system (e.g., heart disease)	0.056
Musculoskeletal disorders (e.g., back pain)	0.108
Respiratory system	0.360
Injuries	0.042
Endocrine system (e.g., diabetes)	0.067
All other	0.048
Year assigned to judge	
1990	0.070
1991	0.082
1992	0.096
1993	0.091
1994	0.101
1995	0.111
1996	0.118
1997	0.112
1998	0.114
1999	0.104
Allowance by ALJ	0.645
Allowance 3 years after assignment to an ALJ	0.769
Participation 3 years after assignment to an ALJ	0.191
Earnings 3 years after assignment to an ALJ	2345

N=1,779,825

allowance 3 years after assignment to a judge on participation 3 years after assignment to a judge. It conditions on a full set of hearing office-day interactions, drops observations that are missing education information, and includes those who died in the 3 years after assignment (and uses allowance status at time of death for allowance and sets participation to 0 for these individuals). In this appendix, we re-estimate the model, making different assumptions in the specifications below.

In the second row we include the 123,911 individuals with missing education. When we do this the estimate for participation rises in magnitude from -0.256 to -0.257. The third row drops both those with missing education (as in the baseline case) as well as the 49,017 individuals who died within 3 years following assignment. When we do this the estimate for participation rises in magnitude to -0.260. The fourth row uses the baseline sample and conditions on a full set of hearing office-year interactions, rather than a full set of hearing office-day interactions. This leaves the point estimate unchanged; more precisely, it changes the estimate from -0.2561 to -0.2565.

TABLE A3: ROBUSTNESS CHECKS, IV ESTIMATES

Dependent Variable: Participation	Estimate	Std. Error	N
Benchmark specification	-0.256	0.006	1,779,825
Include those with missing education	-0.257	0.006	1,903,736
Drop those who died within 3 years after assingment	-0.260	0.006	1,730,808
Condition on hearing office-year interactions (rather than hearing office-day interactions)	-0.256	0.006	1,779,825

Re-applications and appeals

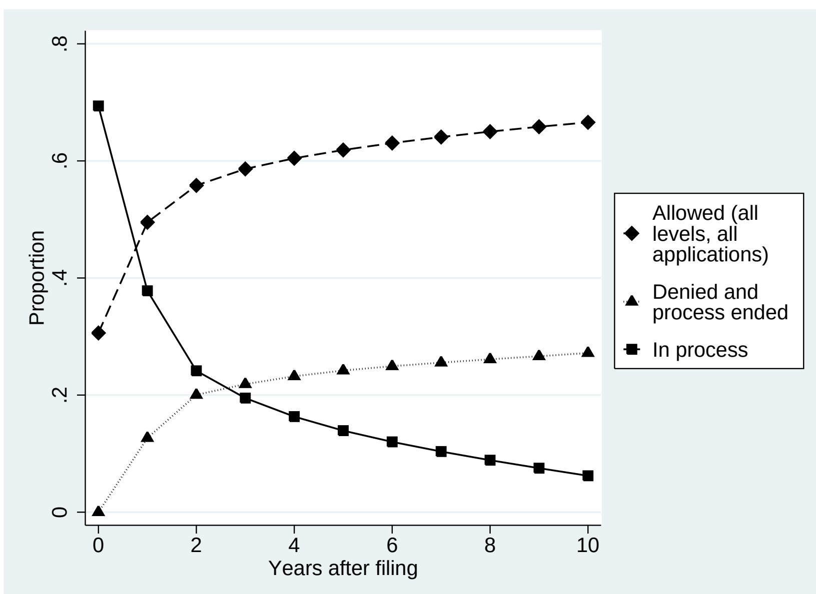

FIGURE A1: SHARE OF ALL DI/SSI APPLICANTS WHO ARE ALLOWED BENEFITS, ARE APPLYING/APPEALING, AND SHARE WHO ARE DENIED, NO LONGER RE-APPLYING OR APPEALING

Figure A1 uses the same data as in figure 1 shows the total share of initial claims allowed at any level. It also disaggregates those cases not allowed into those where the application process ended versus those who were re-applying or appealing a denial. 10 years after the initial filing, 67% of all claimants were allowed benefits, 27% were denied and the process ended, and 6% were still in the process of applying for benefits. Together, figures 1 and A1 emphasize the fact that re-applications and appeals are important for understanding the DI system.

Appendix B: Derivations

Marginal Treatment Effects

All derivations in this are purely for completeness – they are straightforward adaptations of that discussed in Heckman et al. (2006) or French and Taber (2011). Define A_i as a 0-1 indicator =1 if individual i is allowed benefits, y_i is earnings, participation, appeals, or future allowance. We drop t subscripts for simplicity. Individual i's earnings are characterized by

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

where

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

Combining equations (10) and (11) yields:

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

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

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

where $1\{.\}$ is the indicator function, $Z_i = (j_i, X_i)$, and j_i represents a full set of judge dummy variables. By assumption, u_i and ϕ_i are potentially correlated with each other but

V_i is independent of j_i and X_i. The Marginal Treatment Effect is

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

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

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

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

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

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

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

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

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

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

Demeaning the data

We have over 500,000 hearing office-day interactions as covariates, so directly estimating equations (1) and (2) is not computationally feasible. To simplify the problem we de-mean the data. Specifically, we take the difference between $f(j_i\hat{\gamma}_{1,-i})$, A_{it}, $K(\widehat{A}_{it})$, and y_i and the means of the same variables heard at the same hearing office and same day.[14] We then estimate:

$$\begin{aligned} \widetilde{A}_{it} &= f(\widetilde{\widetilde{j_i}\hat{\gamma}_{1,-i}}) + \eta_{it}, & (19) \\ \widetilde{y}_{i\tau} &= K(\widetilde{\widehat{\widetilde{A}}_{it}}) + \mu_{i\tau} & (20) \end{aligned}$$

where "$\sim$" represents a de-meaned variable, e.g., $\widetilde{A}_{it} = A_{it} - \bar{A}_{it}$ and $\bar{A}_{it}$ is the mean allowance rate at the hearing office and on the day that case i was assigned and $\widetilde{j}_i = j_i - \bar{j}_i$ and $\bar{j}_i$ is the mean value of j at the hearing office and on the day that case i was assigned. For the functions $f(.)$ and $K(.)$ we choose polynomials, so $f(\hat{j}_i) = \sum_{k=1}^{K} \omega_k \hat{j}_i^{\,k}$ and $K(\hat{A}_{it}) = \sum_{k=1}^{K} \lambda_{kt} \hat{A}_{it}^{\,k}$. Polynomials are straightforward to demean, so $\widetilde{f(\hat{j}_i)} = \sum_{k=1}^{K} \omega_k \widetilde{\hat{j}_i^{\,k}}$, where $\widetilde{\hat{j}_i^{\,k}} = j_i^k - \overline{j_i^k}$ (where $\overline{j_i^k}$ is demeaned value of the k^{th} power of the judge-specific allowance rate of all judges at the office where case i was heard) and $\widetilde{K(\hat{A}_{it})} = \sum_{k=1}^{K} \lambda_{kt} \widetilde{\hat{A}_{it}^{\,k}}$, where $\widetilde{\hat{A}_{it}^{\,k}} = \hat{A}_{it}^{\,k} - \overline{\hat{A}_{it}^{\,k}}$. We choose the order of polynomial P that minimizes Akaike's information criterion, $\ln \hat{\sigma}^2 + 2P/N$ and the Bayesian information criterion, $\ln(\hat{\sigma^2}) + P/N \cdot \ln(N)$. Because of the well known endpoint problems with polynomials, we experimented with the order of the polynomial. We found that the results were largely unchanged when we increased or decreased the order of the polynomial by 1.

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

$$A_{i1} = j_i\hat{\gamma}_1 + X_i\delta_{A1} + e_{i1} \tag{21}$$

implies

$$E[A_{s1}|X_s] = E[j_s\hat{\gamma}_1|X_s] + X_s\delta_{A1} \tag{22}$$

[14]This is equivalent to taking residuals from first stage regressions of $f(j_i\hat{\gamma}_{1,-i})$, A_{it}, $K(\widehat{A}_{it})$, and y_{it} on X_i.

for any given s and so

$$E[j_s\hat{\gamma}_1 - E[j_s\hat{\gamma}_1|X_s]] = E[A_{s1} - E[A_{s1}|X_s]] \tag{23}$$

where the left-hand side object is $E[j_s\hat{\gamma}_1 - E[j_s\hat{\gamma}_1|X_s]]$, the de-meaned instrumental variable. We approximate the right-hand side object, but using the sample analog and leaving observation i out, as in a jackknife estimator, so the constructed instrument is:

$$\widetilde{j_i}\hat{\gamma}_{1,-i} = \frac{1}{N_j - 1} \sum_{s \in \{J\}, s \neq i} A_{s1} - \overline{A_{s1}} \tag{24}$$

where N_j is the number of cases heard by judge j_i over the sample period, $\{J\}$ is the set of cases heard by judge j_i, $\overline{A_{s1}}$ is the mean allowance rate by ALJs at case s's hearing office on the day case s was heard. Doyle (2008) uses a similar approach. Because we remove case i from $\widetilde{j_i}\hat{\gamma}_{1,-i}$, as in a jackknife estimator, it should be independent of η_i and μ_i, even in a small sample.

Based on Monte Carlo experiments with what seemed reasonable parameters, the procedure produced accurate approximations in the linear models, as well as for the true MTE from the 10th to 90th percentiles of the distribution of the estimated judge allowance differentials, so we present estimates of the MTE over the middle 80 percent of the data.

Appendix C: Using IV estimates to identify the effect of ALJ allowance on the level of labor supply, future allowance, and appeals

Level of labor supply

The plim of the IV estimator is $E[y_{i\tau}|A_{iti} = 1] - E[y_{i\tau}|A_{it} = 0]$ where $y_{i\tau}$ is an outcome measure (participation, earnings, allowance or appeals) at time τ and A_{it} is an indicator equal to 1 if the individual was allowed at time t.

First we describe identification of the effect of ALJ allowance on the level of labor supply. The estimation procedure described in section 4.2 identifies the change in earnings or participation caused by DI receipt. To obtain the level, note that the law of total probability gives

$$E[y_{i\tau}] = E[y_{i\tau}|A_{it} = 1] \Pr[A_{it} = 1] + E[y_{i\tau}|A_{it} = 0] \Pr[A_{it} = 0]. \tag{25}$$

Furthermore, equation (2) shows that

$$E[\phi_{i\tau}] = E[y_{i\tau}|A_{it} = 1] - E[y_{i\tau}|A_{it} = 0]. \tag{26}$$

Using equations (25) and (26) we can solve for the two unknowns:

$$E[y_{i\tau}|A_{it} = 1] \;=\; E[y_{i\tau}] + E[\phi_{i\tau}]\,\Pr[A_{it} = 1] \tag{27}$$

$$E[y_{i\tau}|A_{it} = 0] \;=\; E[y_{i\tau}] - E[\phi_{i\tau}]\,\Pr[A_{it} = 0]. \tag{28}$$

We can identify $E[y_{i\tau}]$, $\Pr[A_{it} = 1], \Pr[A_{it} = 0]$ directly from the data. Our estimation procedure delivers $E[\phi_{i\tau}]$ for cases who are affected by our instrument. Assuming that $E[\phi_{it}]$ for those affected by the instrument is the same as $E[\phi_{it}]$ for those not affected by the instrument yields estimates of $E[y_{i\tau}|A_{it} = 1]$ and $E[y_{i\tau}|A_{it} = 0]$ for the full sample. This assumption is untestable, although section 6.6 gives evidence that $E[\phi_{i\tau}]$ does not vary much over the support of our data.

Future Allowance and Appeals

Next we describe identification of time t allowance on the level of future allowance and appeals. To do this we estimate equation (2), or in de-meaned form, equation (4), where the left hand side variable is time τ allowance $A_{i\tau}$ or appeals $a_{i\tau}$ and the coefficient on time t allowance converges to $E[\phi_{i\tau}]$ for the set of individuals affected by the instrument. The regression coefficient identifies $E[\phi_{i\tau}] = E[A_{i\tau}|A_{it} = 1] - E[A_{i\tau}|A_{it} = 0]$. Because allowance is a binary variable, and because allowance is an absorbing state, $E[A_{i\tau}|A_{it} = 1] = prob[A_{i\tau} = 1|A_{it} = 1] = 1$. Thus the regression coefficient identifies

$$E[A_{i\tau}|A_{it} = 1] - E[A_{i\tau}|A_{it} = 0] \;=\; 1 - prob[A_{i\tau} = 1|A_{it} = 0] \tag{29}$$

and so $prob[A_{i\tau} = 1|A_{it} = 0] = 1 - E[\phi_{i\tau}]$.

When considering appeals define $a_{i\tau}$ as an indicator equal to 1 if the individual was appealing at time τ. Then

$$\begin{aligned} E[a_{i\tau}|A_{it} = 1] - E[a_{i\tau}|A_{it} = 0] &= 0 - E[a_{i\tau}|A_{it} = 0] \\ &= -prob[a_{i\tau} = 1|A_{it} = 0] \end{aligned} \tag{30}$$

and so $prob[A_{i\tau} = 1|A_{it} = 0] = -E[\phi_{i\tau}]$ where $E[\phi_{i\tau}]$ is the plim of the regression coefficient on the appeals equation.

Appendix D: Calculation of the After-Tax Wage

We estimate after-tax wages as follows. We impute pre-tax wage income of non-working DI applicants using a predictive mean matching regression approach, described in David et al. (1986). We first regress income y on the vector of observable variables w described in table 1, yielding $y = wb + \vartheta$. Second, for each sample member i we calculate the predicted value $\hat{y}_i = w_i\hat{b}$, and for each member with an observed value of y_i we calculate the residual $\hat{\vartheta}_i = y_i - \hat{y}_i$. Third, we sort the predicted value $\hat{y}_i$ into deciles. Fourth, for non-working individuals, we impute ϑ_i by finding a random individual j with a value of $\hat{y}_j$ in the same decile as $\hat{y}_i$, and setting $\vartheta_i = \hat{\vartheta}_j$. The imputed value of y_i is $\hat{y}_i + \hat{\vartheta}_j$. We estimate models for DI and SSI beneficiaries separately because the two groups face different labor supply incentives.

Once we impute pre-tax wage income for every member of the sample, we calculate the after-tax wage. First, we use year, state, and the Social Security earnings data to calculate the DI/SSI benefit for everyone in the sample. We impute SSI benefits using state and year for those drawing SSI benefits. Second, we predict the distribution of post-tax wages plus DI benefits (i.e., the difference between income if working and income if not working) for everyone in our data using the federal, state, and local tax schedule shown in French and Jones (2011). Those who are allowed benefits will have DI benefits if predicted income from working is below the SGA limit ($6,000 in 1993 to $9,360 in 2002). If income is above the SGA limit, then the individual will lose benefits. If the individual is denied benefits, then there are no DI benefits to be lost when working. We assume that SSI benefits above the disregard level are reduced 50 cents for each dollar of earnings, until all SSI benefits are lost. Third, we take the sample average after-tax wage if denied and allowed, which is our measure of $E[w_i|A_i = 0]$ and $E[w_i|A_i = 1]$. Our main limitation on these measurements is that ideally we should know family structure and all sources of income to calculate taxes. Family structure is important because the DI/SSI benefit depends on marital status and the number of dependants. Unfortunately, we do not have this information, so we assume that the individual can claim no dependants for the DI/SSI benefit and is not pushed into a higher marginal tax bracket from spousal or other non-labor income.