Federal Reserve Bank of Chicago

The Effect of Disability Insurance Receipt on Labor Supply: A Dynamic Analysis

Eric French and Jae Song

WP 2012-12

The Effect of Disability Insurance Receipt on Labor Supply: a Dynamic Analysis

Eric French and Jae Song*

December 6, 2012

Abstract

This paper estimates the effect of Disability Insurance receipt on labor supply, accounting for the dynamic nature of the application process. 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 when not considering the dynamic nature of the applications process. 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. Next, we estimate a dynamic programming model of optimal labor supply and appeals choices. Consistent with the law, we assume that people cannot work and appeal at the same time. We match labor supply, appeals, and subsequent allowance decisions predicted by the model to the decisions observed in the data. We use the model to predict labor supply responses to benefit denial when there is no option to appeal. We find that if there was no appeals option, those denied benefits are 35 percentage points more likely to work. However, there is considerable heterogeneity in responses. Most individuals in their 40s would return to work if denied benefits, for example. Our results suggest that many of those denied benefits not because they are

*Comments welcome at jae.song@ssa.gov and efrench@frbchi.org. Affiliations are Social Security Administration and Federal Reserve Bank of Chicago, respectively. We thank seminar participants at the Conference in Honor of John Kennan, the Cowles Foundation Structural Microeconomics Conference at Yale, IZA, and Tilburg University for helpful comments and David Benson An Qi for excellent research assistance. The views in this paper are not necessarily those of the Federal Reserve System or Social Security Administration.

unable to work, but because they remain out of the labor force in order to appeal their benefit denial.

1 Introduction

This paper presents new evidence on the effect of Disability Insurance (DI) 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 improvements. First, 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. 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 50% of applicants who are denied benefits by an ALJ are eventually allowed benefits within five years. During the appeal process, these applicants tend not to work, even though they are currently not receiving benefits. This has an important effect 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 similar labor supply responses to those handled by lenient judges. This is consistent with the view that the marginal applicant handled by a strict judge is as physically unable to work as 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.

Next, we estimate a dynamic programming model of optimal labor supply and appeals choices. Consistent with the law, we assume that people cannot work and appeal at the same time. We match labor supply, appeals, and subsequent allowance decisions predicted by the model to the decisions observed in the data. We use the model to predict labor supply responses to benefit denial when there is no option to appeal. We find that if there was no appeals option, those denied benefits are 35 percentage points more likely to work. Our results suggest that many of those denied benefits not because they are unable to work, but because they remain out of the labor force in order to appeal their benefit denial.

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, section 7 reports results from the dynamic programming model, and section 8 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 disability insurance benefits. The total cost of the program was $85.4 billion, making it more costly than unemployment insurance. Furthermore, after two years on the disability rolls, individuals become eligible for Medicare benefits. The total cost of Medicare payments to DI beneficiaries was $49 billion in 2005 (Autor and Duggan 2006).

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. 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. Von Watcher et al. (2011) find that the patterns documented by Bound have changed little over time.

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. Bound's claim was that this should lead to an overstatement of the effect of disability on labor supply, because those who are denied are on average healthier and thus more likely to work than those who are 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. However, Lahiri et al. (2008) found that those who are denied benefits tend to have very intermittent work histories. Those who are allowed benefits are more likely to work and have higher earnings before applying for benefits. Thus it is not clear whether those who are denied are more or less likely to work in the absence of benefits.

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 compliments 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. The advantage of our study relative to theirs 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. Furthermore, 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. 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.

[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).

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 reapply 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.

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 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

[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.

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!”

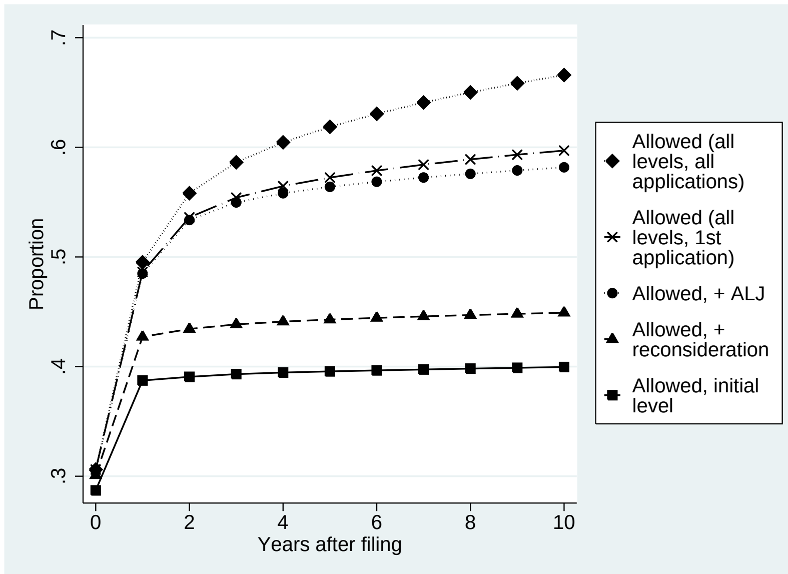

Figure 1: ALLOWANCE AT DIFFERENT STAGES OF THE APPLICATIONS AND APPEALS PROCESS.

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 mental and 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

[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

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, are allowed at the Disability Determination Service stage. The healthiest individuals will be denied by every judge and will be denied 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.

3.3 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"

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.

[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.

(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 the same as the deciding judge in 96% of all cases. Although the deciding judge is not necessarily randomly assigned, the initially assigned judge is.[6] We use the initial assignment to a judge as our source of exogenous variation.

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 and participation.

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

[6]The initially assigned judge is not necessarily the judge who handles the case. This fact can potentially be exploited by DI claimants. 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. In this way, an individual can potentially reject a judge.

(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 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 500,000 hearing office-day interactions as 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

[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.

independent of e_{it} or $u_{i\tau}$, even in a small sample. Third, we estimate the equations

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

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

jointly using two stage least squares.

Given the above assumptions, Heckman, Urzua, and Vytlacil (2006) and French and Taber (2010) point out that this procedure identifies a weighted average of $\phi_{i\tau}$ for the set of individuals 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

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

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

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

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

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 (5).

where $\widehat{\widetilde{A}}_{it}$ is the predicted value of $\widetilde{A}_{it}$ from equation (6). As shown by Heckman, Urzua, and Vytlacil (2006) and French and Taber (2010), as well as appendix B, the MTE is

$$K'(a_t) = E[\phi_{i\tau}|\widetilde{A}_{it} = a_t] \tag{8}$$

where a_t is a particular value of A_{it}. This value of a_t can also be interpreted as the (lack of) judge-observed severity of the case. As $\widetilde{A}_{it}$ 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.

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.

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 $t-statistics$. 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 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 mental retardation. Those with mental disorders and mental retardation are assigned to judges who have 0.16% lower allowance rates than average. These coefficients

TABLE 1: PREDICTORS OF ALLOWANCE AND JUDGE ALLOWANCE DIFFERENTIAL

	Dependent variable: Allowed		Dependent variable: judge allowance differential	
	Coefficient	t-statistic	Coefficient	t-statistic
Covariate	(1)	(2)	(3)	(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).

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 and high earners have relatively high allowance rates. 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 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 after assignment as our key measure of allowance in table 2. Many cases denied by an ALJ are later allowed.

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.

TABLE 3: ESTIMATED EFFECT OF DI RECIPIENCY ON LABOR SUPPLY

	Dependent Variable: Earnings		*Dependent Variable:*
	OLS	IV	OLS
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
(Std. Error)	(34)	(140)	(0.002)
With Covariates:			
Coef on demeaned allowance*	-4247	-4023	-0.271
(Std. Error)	(65)	(127)	(0.002)
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 ear
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.

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 columns show mean earnings and labor force participation (measured as earnings > $100) for those allowed and denied benefits, three years after assignment to an ALJ. Column 3 shows the difference and column 4 the associated standard error. Columns 5 and 6 show OLS and IV estimates of de-meaned (by hearing office and day) earnings on similarly de-meaned allowance. The IV estimate is the estimate from equation (2). The next column includes the covariates listed in 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. Recall that our estimation procedure should deliver consistent estimates, with or without covariates. Thus the fact that adding covariates does not change the estimates is reassuring.

6.4 Dynamics of the Response

This section shows the dynamics of the response of both earnings and participation. Using the estimation procedure described in section 4.2 we can identify the change in earnings or participation caused by DI receipt at any point in time. 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 $E[\phi_i]$ for those affected by the instrument is the same as $E[\phi_i]$ for those not affected by the instrument: see appendix C for details. This assumption is untestable, although section 6.6 gives evidence that $E[\phi_i]$ does not vary much over the support of our data.[9]

Figure 2 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 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 2 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. We then infer the level of labor supply using the approach described

[9]In contrast to our findings, Maestas et al. (2011) do find variability in $E[\phi_i]$ across the support of their data.

[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.

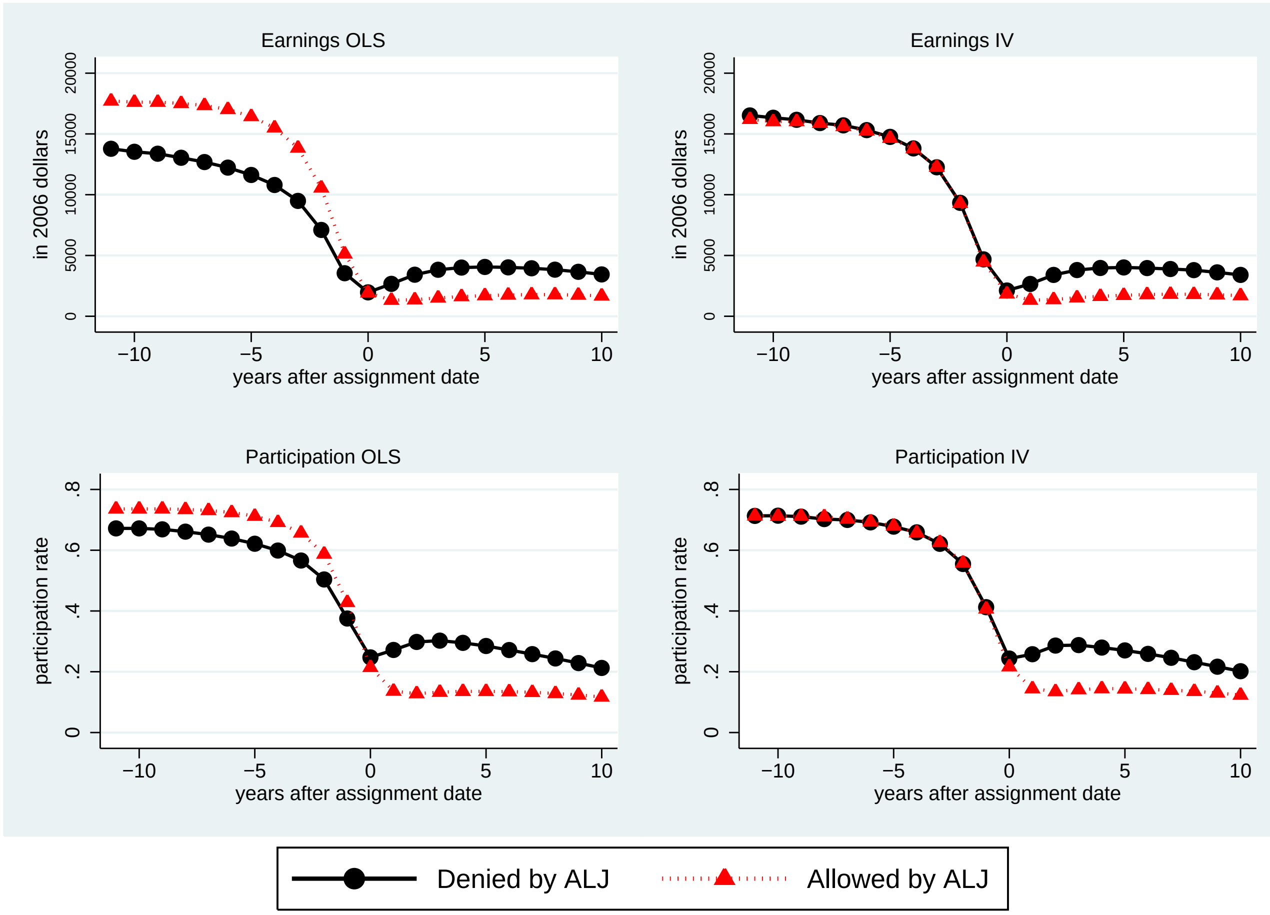

Figure 2: Dynamics of earnings and participation, allowed versus denied by ALJ.

in section ??. Earnings and participation rates of the two groups are virtually identical before assignment to a judge, which is unsurprising given that our instrument is uncorrelated with earnings prior to assignment. 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 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 2 (\$2,314) than in table 3 (\$4,059). The difference arises because figure

2 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

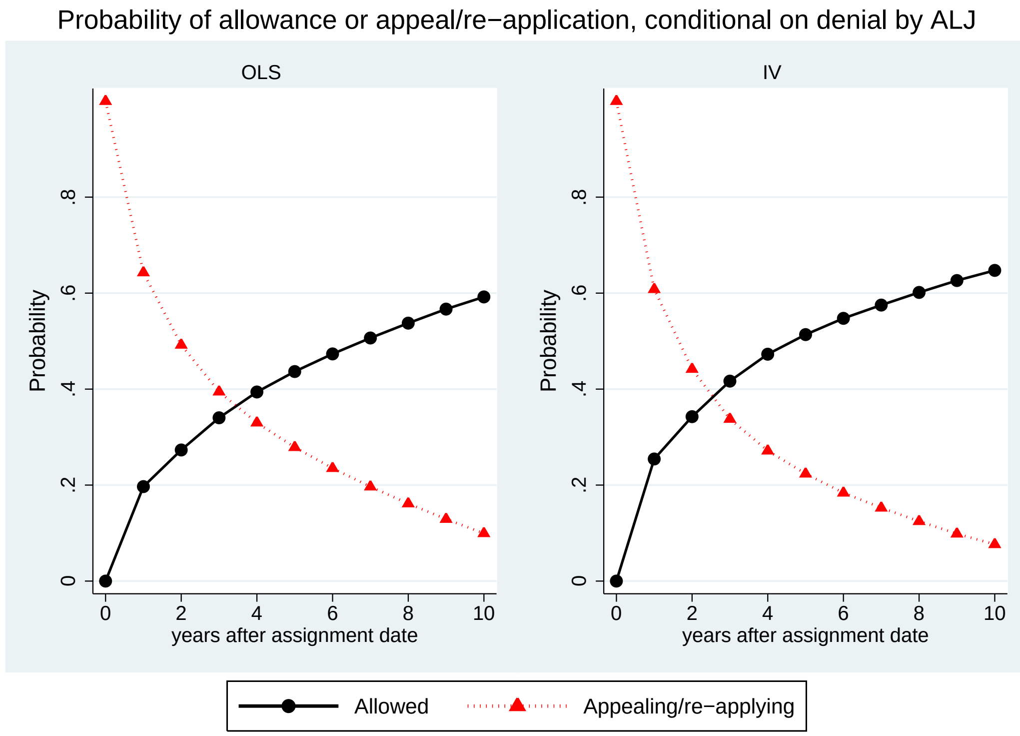

Figure 3: Allowance and Appeals/Re-applications following denial by ALJ.

The left panel of figure 3 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.[11] It 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. Because most denied applicants have either been allowed benefits or have given up applying for benefits by

[11]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 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.

this point, we focus on allowance rates and labor supply decisions three years after assignment to a judge in this paper.

The right panel of figure 3 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.[12] 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 3 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 reapplications 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.

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.

The left panel of figure 4 shows the earnings decline and the right panel shows the partic-

[12] Using the set of individuals who were denied by an ALJ, we regress de-meaned allowance on a set of wave dummies and predicted de-meaned ALJ allowance $\times$ wave dummies (where allowance is predicted using the judge allowance differential). The estimated coefficient on allowance$\times$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$\times$wave measures increased probability of allowance at a given wave conditional on initial denial. The right panel of figure 3 plots the coefficient on predicted allowance$\times$wave for both the allowance and appeal equations.

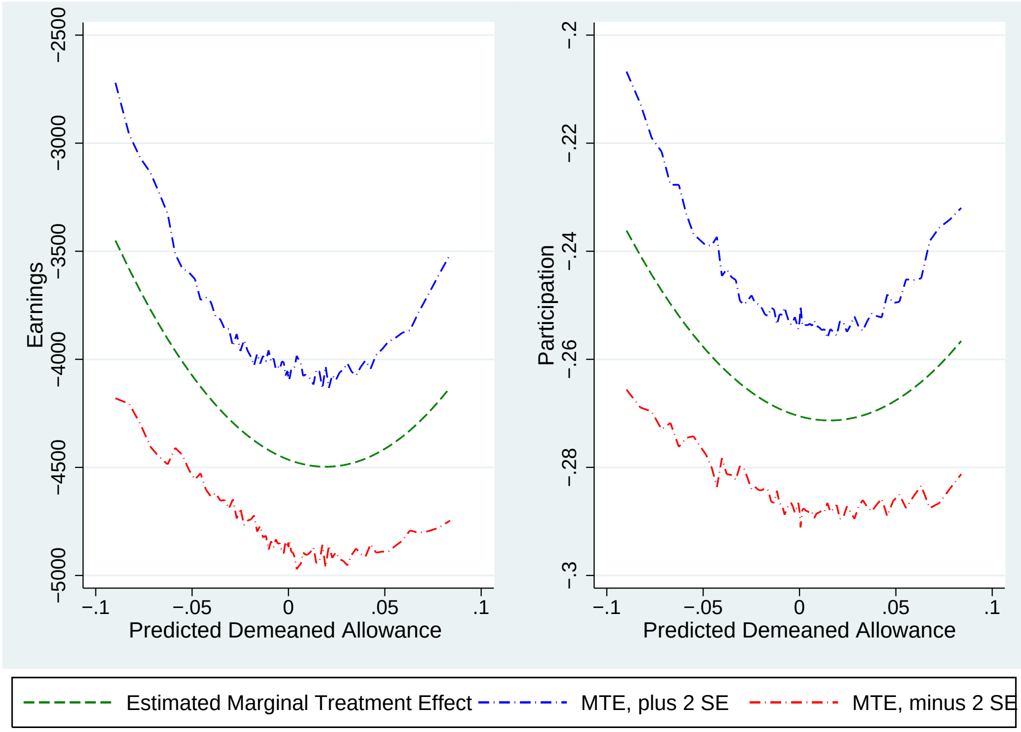

Figure 4: Earnings and participation decline when allowed for marginal applicant.

ipation 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 (**??**) and (**??**). 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 4 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 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.

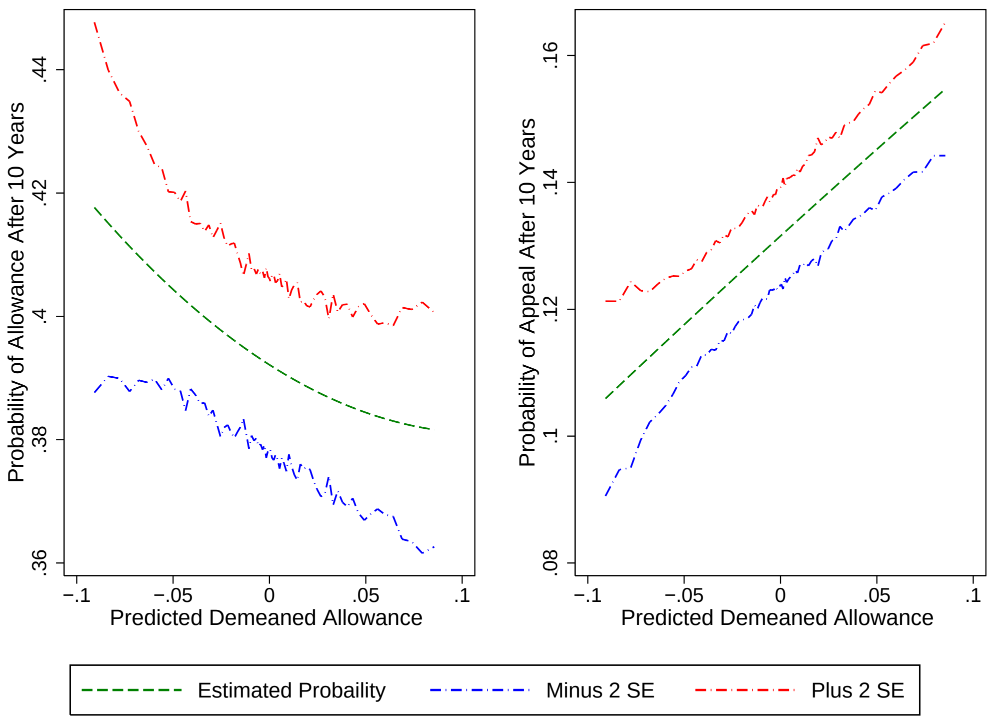

Figure 5: Marginal applicant's allowance and appeal probability 10 years after assignment conditional on not allowed 3 years after assignment to an ALJ .

Figure 5 shows how allowance three years after assignment to an ALJ affects allowance and appeal 10 years afterwards. The left panel shows that allowance three years after assignment to an ALJ increases the probability of allowance 10 years after assignment by .60 on average. Put differently, 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 (1-.62)=.38. For those assigned to strict ones it is (1-.58)=.42. The right panel of figure 5 shows that allowance three years after assignment reduces the average probability of appealing 10 years after assignment by .13, so 13% of those not allowed three years after assignment are still appealing 10 years after assignment. For marginal applicants assigned to lenient judges it is 15% and for those assigned to strict judges it is 11%. Figure 5 shows that for a marginal applicant not allowed three years after assignment to a lenient judge, the probability that she is either allowed benefits or appeals 10 years after assignment is .38 and .15, respectively. Thus conditional on not being allowed, 1-(.38+.15)=47% of those who do not get benefits and do not appeal or re-apply. For those assigned to the stricter judges, the numbers are 42% for allowance, 11% for appealing, and 1-(.42+.11)=47% for not being allowed and not appealing.

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. Nevertheless, the right panel of figure 5 shows that these people continue trying to get the benefit. Remarkably, conditional on being denied 3 years after assignment, over half of all cases are allowed, appealing, or re-applying for benefits 10 years after assignment.

7 Re-interpretation of results using a model

7.1 Model set up

Previous sections showed that many people who are denied benefits do not work and are later allowed upon appeal. It is not clear if they would have worked if there was no appeals option. This section uses a dynamic model to predict the effect of DI receipt on labor supply when there is no option for appeals. We estimate the model's parameters that best match the previously estimated profiles of labor force participation, appeals, and allowances. Upon doing so, we use the model to infer whether many denied individuals would work if they lost the option to appeal.

Consider a single person seeking to maximize his or her expected lifetime utility at age $t = \tau, \tau + 1..., 100$, where 100 is the age of certain death. At each time t she must decide to make a decision $d_t = \{p, a, n\}$, where p is participate in the labor force, a is appeal or re-apply for benefits, and n is neither work nor appeal. The individual makes these decisions depending on her age and whether she is allowed DI benefits $A_t \in \{0, 1\}$. If $A_t = 0$ and she applies for benefits, the probability of being allowed benefits next year is $\Pr(A_{t+1} = 1|a, A_t = 0, h) = \Pr_{t+1}$, where h represents her observed health. Her flow utility is

$$v(d_t) = u(c_t) - \upsilon_a 1\{d_t = a\} - \upsilon_{pt} 1\{d_t = p\} \tag{9}$$

where the parameter $\upsilon_p t$ represents the distulity of work and υ_a represents the the disutility of applying for benefits. We allow the disutility of work can change with age so that:

$$\upsilon_{pt} = \upsilon_{p0} + (t - 42)\upsilon_{p1}$$

The Before age 65 consumption $c_t = w$ if $d_t = p$, $c_t = b$ if $d_t = n$ and $A_t = 1$, and $c_t = \underline{c}$ if $d_t = n$ or a and $A_t = 0$, where $u(w) > u(b) > u(\underline{c})$. After age 65 consumption is $c_t = b$ since everyone is eligible for full benefits at the Normal Retirement Age (which was 65 during the sample period) and few DI applicants work after that age. Individuals discount the future at rate β. The probability of surviving to time t conditional on being alive at time $t - 1$ is S_t.

The parameters υ_{p0}, υ_a and υ_h vary across members of the population, but do not vary across time ($\upsilon_p 1$ is a constant). Furthermore, the distribution of these variables is joint normal:

$$\begin{pmatrix} \upsilon_{p0} \\ \upsilon_a \\ \upsilon_h \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_p \\ \mu_a \\ \mu_h \end{pmatrix}, \begin{pmatrix} V[\upsilon_p] & C[\upsilon_p, \upsilon_a] & C[\upsilon_p, \upsilon_h] \\ C[\upsilon_p, \upsilon_a] & V[\upsilon_a] & C[\upsilon_a, \upsilon_h] \\ C[\upsilon_p, \upsilon_h] & C[\upsilon_a, \upsilon_h] & V[\upsilon_h] \end{pmatrix} \right) \tag{10}$$

The probability of allowance varies across members of the population according to both their age and health status. The Social Security Administration can perfectly observe health status. However, there is heterogeneity in the threshold rule used by different administrators (such as judges). Individuals are uncertain of which administrator they will be assigned to, and this means that allowance is stochastic on the part of individuals.

Specifically, decision to allow benefits to an individual is

$$A_{t+1} = 1\{\upsilon_h > \chi_{ht}, d_t = a\} \tag{11}$$

where

$$\chi_{ht} = \chi_t + \chi_{jt}, \quad \chi_{jt} \sim N(0, V[\chi_j]), \tag{12}$$

$$\chi_t = \begin{cases} \alpha_0 & \text{if} \quad t = 0 \\ \alpha_1 \exp(\alpha_2 t) & \text{if} \quad t > 0 \end{cases}$$

where χ_{jt} represents the judge specific threshold of the judge who handled the case at time t. Every period the individual appeals, she receives a new independent draw of χ_{jt}. To parameterize the time varying but deterministic component of the threshold χ_t, note that the ALJ allowance rate is 65%, much higher than in later stages of the adjudication process. Therefore, we let the time 0 threshold have mean α_0. We let the deterministic component of the threshold decline exponentially thereafter, which is consistent with the estimated profile shown in figure XX (somewhere we should have the figure on the probability of allowance, conditional on appealing). Equations (11) and (12) imply that the probability of allowance for an individual is

$$\begin{aligned} \Pr(A_{t+1} = 1 | a, A_t = 0, \upsilon_h) &= \Pr(\upsilon_h > \chi_{ht} | a, A_t = 0, \upsilon_h) \\ &= \Pr(\upsilon_h > \chi_t + \chi_{jt} | \upsilon_h) \\ &= \Pr(\upsilon_h - \chi_t > \chi_{jt} | \upsilon_h) \\ &= \upsilon\left(\frac{\upsilon_h - \chi_t}{\sqrt{V[\chi_j]}}\right) \end{aligned} \tag{13}$$

7.2 Value Functions

The value function is

$$V_t(A_t) = \max_{d_t} \left\{ v(d_t) + \beta S_{t+1} E_t V_{t+1}(A_{t+1}) \right\}. \tag{14}$$

As we will show below, if $A_t = 1$ the decision problem is a simple one: consume the DI benefit and do not work and so the value function will be

$$V_t(A_t = 1) = \sum_{\tau=t}^{100} \beta^{\tau-t} \left(\prod_{k=t}^{\tau} S_k \right) u(b) \tag{15}$$

and $S_t = 1$. At the Normal Retirement Age DI benefits are converted into Social Security benefits. Furthermore, everyone is eligible for full Social Security benefits at age 65, so $V_{65}(A_{65} = 1) = V_{65}(A_{65} = 0) = \sum_{t=64}^{100} \beta S_{t+1} u(b)$.[13] Thus the key decisions in the model are for those younger than 65 without DI benefits. Expected discounted lifetime utility from choosing $d_t = p$ and making optimal decisions thereafter is

$$u(w) - \upsilon_p + \beta S_{t+1} V_{t+1}(A_{t+1} = 0, \upsilon_h). \tag{16}$$

Expected discounted lifetime utility from choosing $d_t = n$ and making optimal decisions thereafter is

$$u(\underline{c}) + \beta S_{t+1} V_{t+1}(A_{t+1} = 0, \upsilon_h). \tag{17}$$

Expected discounted lifetime utility from choosing $d_t = a$ and making optimal decisions thereafter is

$$u(\underline{c}) - \upsilon_a + \beta S_{t+1} \left[[1 - \Pr_{t+1}] V_{t+1}(A_{t+1} = 0, \upsilon_h) + \Pr_{t+1} V_{t+1}(A_{t+1} = 1) \right]. \tag{18}$$

Comparing equations (16)-(18) shows that the individual's optimal decision rule is

$$d_t = \begin{cases} p & \text{if } \; u(w) - u(\underline{c}) > \upsilon_p, \;\; u(w) - \upsilon_p > u(\underline{c}) - \upsilon_a + \beta S_{t+1} \Pr_{t+1}[V_{t+1}(A_{t+1} = 1) - V_{t+1}(A_{t+1} = 0, \upsilon_h)] \\ n & \text{if } \; u(w) - u(\underline{c}) \leq \upsilon_p, \;\; \upsilon_a > \beta S_{t+1} \Pr_{t+1}[V_{t+1}(A_{t+1} = 1) - V_{t+1}(A_{t+1} = 0, \upsilon_h)] \\ a & \text{if } \; u(w) - \upsilon_p > u(\underline{c}) - \upsilon_a + \beta S_{t+1} \Pr_{t+1}[V_{t+1}(A_{t+1} = 1) - V_{t+1}(A_{t+1} = 0, \upsilon_h)], \\ & \;\;\;\; \upsilon_a < \beta S_{t+1} \Pr_{t+1}[V_{t+1}(A_{t+1} = 1) - V_{t+1}(A_{t+1} = 0, \upsilon_h)] \end{cases} . \tag{19}$$

Optimal decision rules (for a given (υ_h, t) in this model are shown in figure 6. Those with high disutility of work and appealing never appeal or work.

[13] Thus the optimal decision rules for the dynamic programming problem is the same as for the decision rules coming from the problem where $V_{65}(A_{65} = 1) = V_{65}(A_{65} = 0) = 0$. Thus solve the model backwards from 64.

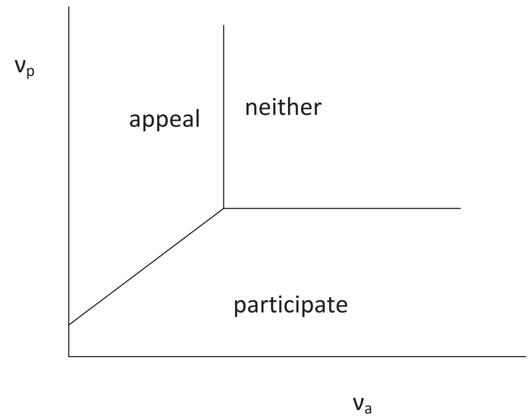

Figure 6: OPTIMAL DECISION RULE FOR $A_t = 0$.

7.3 Estimation

Inspection of equation (19) shows that the model is under-identified, so we make some additional assumptions. First, neither the scale nor location of discrete choice models are identified. Thus, as normalizations, we set $u(\underline{c}) = 0$, $u(b) = 1$, $\mu_h = 0$, and $V[\upsilon_h] = 1$. Next, mean of $u(w)$ is not separately identified from the mean of υ_p. Thus we estimate $u(w) - \mu_p$, which should be interpreted as mean utility when working, including the mean disutility of work.

We estimate the utility parameters $(u(w) - \mu_p, \mu_a)$ five of the six unique elements of Σ (where we normalize the sixth $V[\upsilon_h] = 1$), as well as the probability of allowance parameters $(\alpha_0, \alpha_1, \alpha_2, V[\chi_j])$. [ebf to self: do we estimate β? Is it identified?] This gives us 11 parameters to estimate. We match the model to the profiles for participation (10 years), allowance (10 years), and appeals (10 years), using both OLS and IV estimates, both for the cohort aged 40-44 at the time of assignment and the cohort aged 50-54 at the time of assignment. This gives 120 moment conditions.

Matching OLS estimates in the model to those in the data is fairly straightforward. For participation, we regress simulated participation, allowance, and appeals at different ages on time 0 simulated allowance. We match OLS coefficients on the data simulated by the model to OLS coefficients estimated in the data.

For the IV estimates, our procedure is as follows. First we generate the instrumental variable, which is the mean allowance rate given χ_{j_0}. We calculate mean allowance for each centile of the χ_{j0} distribution. Next, using OLS, we regress participation, allowance, and appeals at different ages on the time 0 predicted allowance rate. We then match these coefficients to the IV parameters estimated in the data. See appendix E for details.

[where to put this para – it looks like maybe move to parameter estimates section.] Inspection of equation (19) and figure 6 shows that 2 parameters are identified from the the participation and allowance decisions at a point in time. Equation (19) also highlights the non-stationarity in the gains from applying for benefits, which also gives valuable identification. The gains from appealing shrink to 0 as the individual nears age 65.

We estimate the model via indirect inference. The solution method is as follows. First, pick parameter values $(u(w) - \mu_p, \mu_a, \Sigma, \alpha_0 \alpha_1, \alpha_2, V[\chi_j])$. Second, given these parameters we solve the model using value function iteration for two cohorts of individuals (with average ages of 42 and 52 when first observed and ages 52 and 62 when last observed), taking 5,000 draws from the joint distribution of $\upsilon_p, \upsilon_a, \upsilon_h$ for each cohort. For each simulated individual, we calculate optimal decision rules using equation (19) where the probability of allowance conditional on application is given in equation (13). Third, we simulate the model using optimal decision rules, and random draws of χ_{ht}. Fourth, we estimate the OLS and IV profiles for appeals, allowance and participation on the simulated data using the same methods we used to estimate the profiles on the real data. Fifth, we compare the model predictions to the profiles observed in the data and calculate sum of squared deviations between model predicted profiles and the profiles observed in the data. Then we draw a new set of parameter values and repeat until we minimize weighted sum of squared errors.

7.4 Parameter Estimates

Given estimated parameters, we can predict labor force participation in the absence of the option to appeal. Optimal decision rules in the absence of the ability to appeal is shown in figure 9.

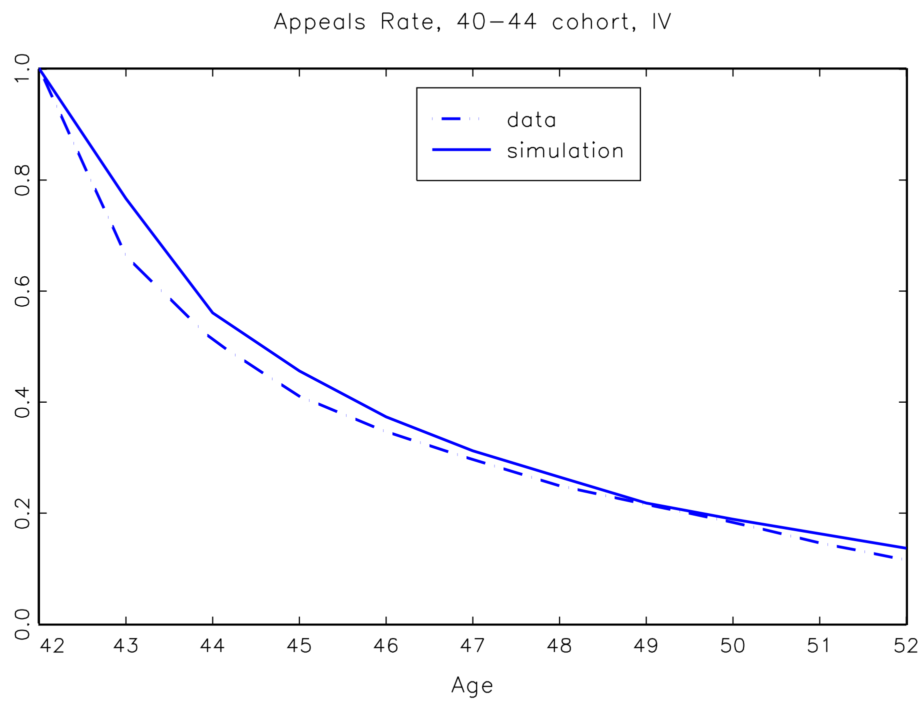

Figure 7: Share appealing/re-applying for benefits, conditional on denial by ALJ, model versus data.

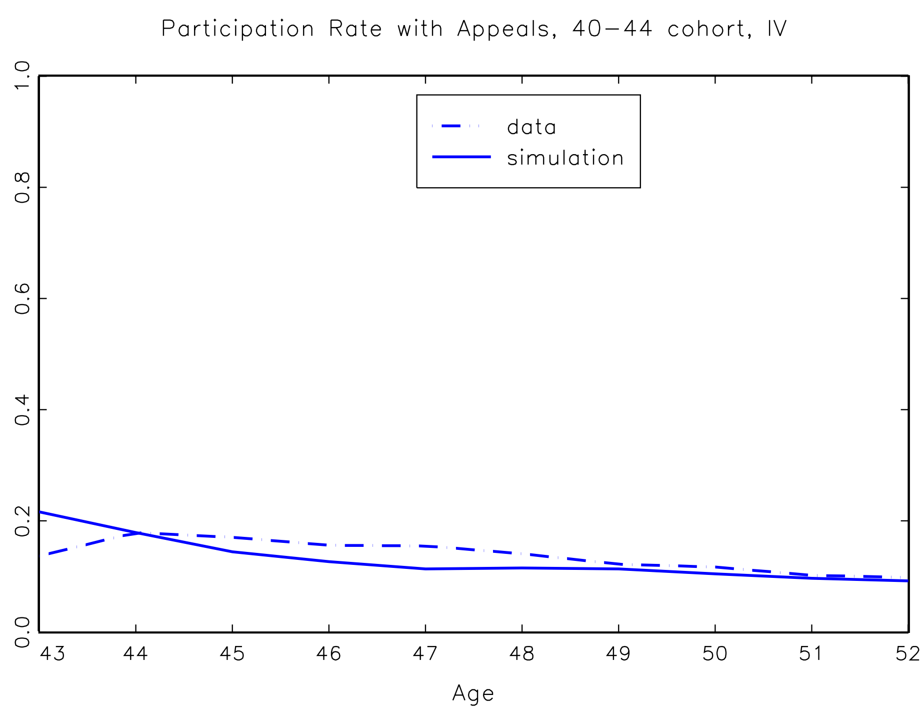

Figure 8: Share working, conditional on denial by ALJ, model versus data.

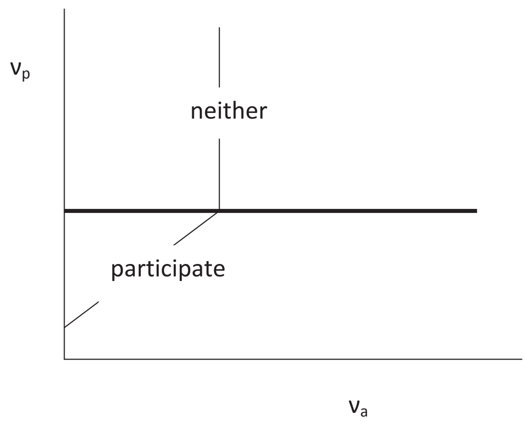

Figure 9: Optimal decision rule for $A_t = 0$, no appeals option.

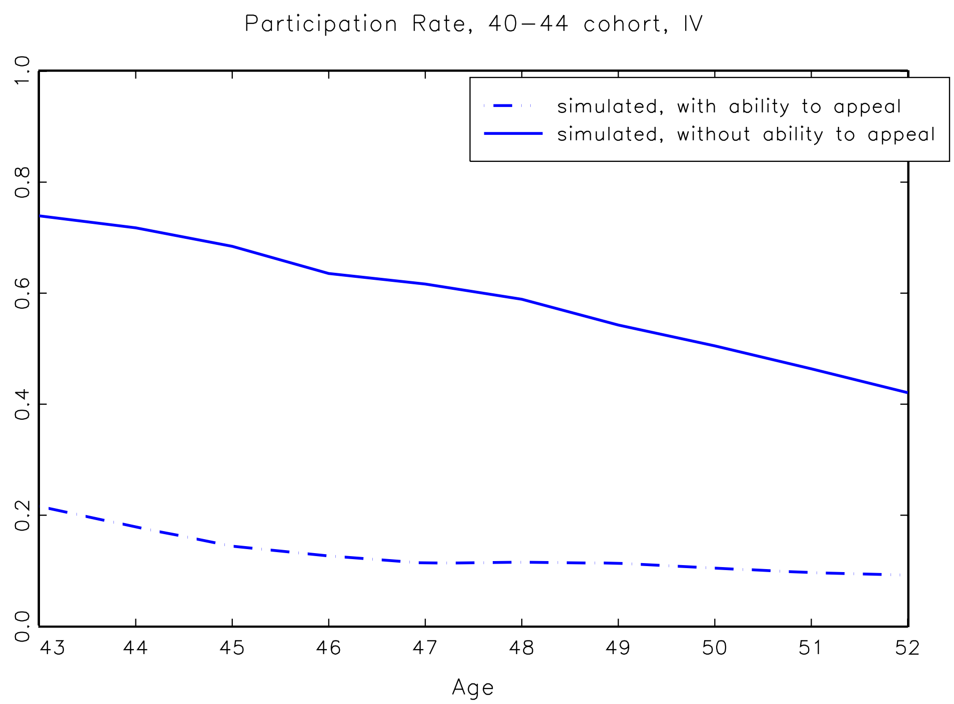

Figure 10: Share working, conditional on denial by ALJ, both with and without the appeals option.

Table 6: Structural Parameter Estimates

Utility function parameters: means	
$u(w) - \mu_p$	0.148
μ_a	0.179
μ_h	0.407
Utility function parameters: variances	
$V[\upsilon_p]$	0.55343000
$V[\upsilon_a]$	0.57897000
Utility function parameters: covariances	
$C[\upsilon_p, \upsilon_h]$	0.10070538
$C[\upsilon_a, \upsilon_h]$	0.28910927
$C[\upsilon_p, \upsilon_a]$	0.42286080
Judge threshold parameters	
α_0	0.38709000
α_1	0.59628000
α_2	0.22642000
$V[\chi_j]$	0.91787000

Next, we summarize the effect of disability insurance receipt with no option to appeal. The next two tables show this. The final table shows that when taking a weighted average over all ages and cohorts, the model predicted participation rate when there is no appeals option is 35%.

Participation, No Appeal, 40-44

Age	OLS	IV
43	0.725	0.731
44	0.693	0.705
45	0.658	0.676
46	0.626	0.644
47	0.590	0.620
48	0.547	0.584
49	0.501	0.545
50	0.451	0.510
51	0.413	0.490
52	0.379	0.450
53	0.337	0.409
54	0.293	0.356
55	0.244	0.308
56	0.199	0.251
57	0.161	0.206
58	0.117	0.162
59	0.071	0.105
60	0.043	0.067
61	0.024	0.043
62	0.008	0.018
63	0.001	0.002
64	0.000	0.000

Participation, No Appeal, 50-54

Age	OLS	IV
53	0.482	0.511
54	0.436	0.466
55	0.397	0.427
56	0.352	0.389
57	0.308	0.347
58	0.275	0.321
59	0.230	0.280
60	0.191	0.243
61	0.149	0.198
62	0.123	0.166
63	0.094	0.128
64	0.065	0.090

Participation, Weighted By Cohort Size

With Appeal	0.077
No Appeal	0.352

8 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.

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. Next, we estimate a dynamic programming model of optimal labor supply and appeals choices. Consistent with the law, we assume that people cannot work and appeal at the same time. We match labor supply, appeals, and subsequent allowance decisions predicted by the model to the decisions observed in the data. We use the model to predict labor supply responses to benefit denial when there is no option to appeal. We find that if there was no appeals option, those denied benefits are 35 percentage points more likely to work.

References

[1] Autor, D. and M. Duggan (2006), "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding", *Journal of Economic Perspectives* 20(3), 71-96.

[2] 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.

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

[4] 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.

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

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

[7] 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.

[8] 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.

[9] 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.

[10] 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.

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

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

[13] 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.

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

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

[16] 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.

[17] 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.

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

[19] Krieder, B. (1998), "Workers' Applications to Social Insurance Programs When Earnings and Eligibility Are Uncertain", *Journal of Labor Economics* 16(4), 848-877.

[20] 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.

[21] 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.

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

[23] 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.*

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

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

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

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

[28] 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.

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

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

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

[32] 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

Reapplications and appeals

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

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

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

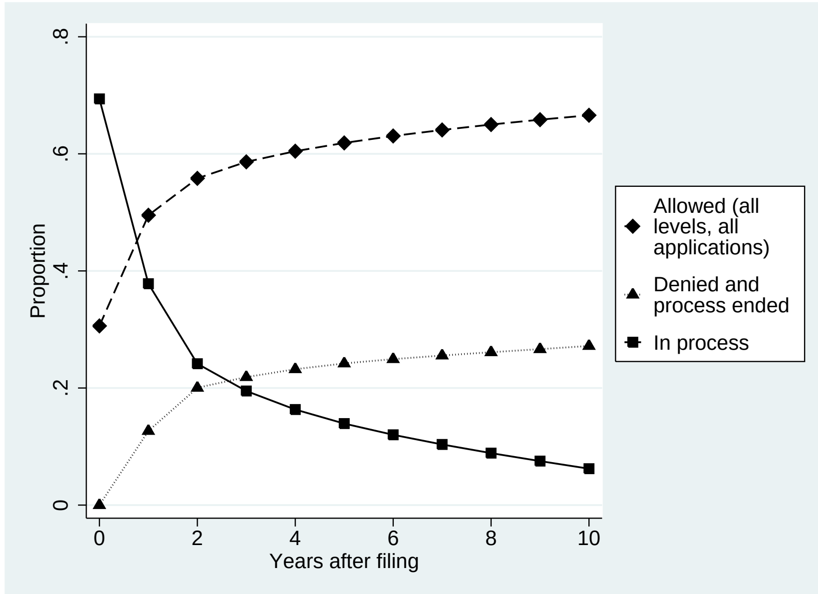

Figure A1: Share of all DI/SSI applicants who are allowed benefits, are applying/appealing, and share who are denied, no longer re-applying or appealing

of that discussed in Heckman et al. (2006) or French and Taber (2010). 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

$$y_i = \begin{cases} y_{1i} & \text{if} \quad A_i = 1 \\ y_{0i} & \text{if} \quad A_i = 0 \end{cases} \tag{20}$$

where

$$\begin{aligned} y_{1i} &= \phi + X_i \delta_y + u_{1i} \\ y_{0i} &= X_i \delta_y + u_i \end{aligned} \tag{21}$$

Combining equations (20) and (21) yields:

$$y_i = A_i \phi_i + X_i \delta_y + u_i. \tag{22}$$

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

$$A_i = 1\{g(Z_i) - V_i > 0\} \tag{23}$$

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

$$MTE(X_i = x, V_i = p) \equiv E[y_{1i} - y_{0i}|X_i = x, V_i = p] \tag{24}$$

where $P(Z_i) \equiv \Pr(A_i = 1|Z_i)$. Given equation (21), $MTE(X_i = x, V_i = p) = \phi + u_{1i} - u_{0i} = \phi_i$. Using equation (22), we estimate the conditional expectation function

$$\begin{aligned} E[y_i|X_i = x, P(Z_i) = p] &= E[A_i\phi_i + X_i\delta_y + u_i|X_i = x, P(Z_i) = p] \\ &= E[A_i(\phi + u_{1i} - u_i)|X_i = x, P(Z_i) = p] + X_i\delta_y + E[u_i|X_i = x, P(Z_i) = p] \\ &= E[A_i\phi|X_i = x, P(Z_i) = p] + E[(u_{1i} - u_i)|A_i = 1, X_i = x, P(Z_i) = p]p + X_i\delta_A \\ &\quad +E[u_i|X_i = x, P(Z_i) = p] \end{aligned} \tag{25}$$

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:

$$\begin{aligned} E[y_i|X_i = x, P(Z_i) = p] &= \phi p + E[(u_{1i} - u_i)|A_i = 1, P(Z_i) = p] + X_i\delta_A + E[u_i|P(Z_i) = p] \\ &= X_i\delta_A + \phi p + E[(u_{1i} - u_i)|A_i = 1, P(Z_i) = p]p + E[u_i|P(Z_i) = p] \\ &= X_i\delta_A + K(p) \end{aligned} \tag{26}$$

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 (26) with respect to p yields

$$\frac{\partial E[y_i|X_i = x, P(Z_i) = p]}{\partial p} = K'(p) \tag{27}$$

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

$$\begin{aligned}
\frac{\partial E[y_i|X_i=x,P(Z_i)=p]}{\partial p} &= \frac{\partial\left[\int_0^p E[y_{1i}|X_i=x,V_i=p]+\int_p^1 E[y_{0i}|X_i=x,V_i=p]\right]}{\partial p} \\
&= E[y_{1i}|X_i=x,V_i=p]-E[y_{0i}|X_i=x,V_i=p] \\
&\equiv MTE(X_i=x,V_i=p). \qquad (28)
\end{aligned}$$

Thus estimation of equation (26) 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:

$$\begin{aligned}
\widetilde{A}_{it} &= f(\widetilde{\widetilde{j}_i\hat{\gamma}_{1,-i}})+\eta_{it}, \qquad (29) \\
\widetilde{y}_{i\tau} &= K(\widetilde{\widehat{\widetilde{A}}_{it}})+\mu_{i\tau} \qquad (30)
\end{aligned}$$

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.

[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.

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

$$A_{i1} = j_i \hat{\gamma}_1 + X_i \delta_{A1} + e_{i1} \tag{31}$$

implies

$$E[A_{s1}|X_s] = E[j_s \hat{\gamma}_1 | X_s] + X_s \delta_{A1} \tag{32}$$

for any given s and so

$$E[j_s \hat{\gamma}_1 - E[j_s \hat{\gamma}_1 | X_s]] = E[A_{s1} - E[A_{s1}|X_s]] \tag{33}$$

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:

$$\widetilde{j_i} \hat{\gamma}_{1,-i} = \frac{1}{N_j - 1} \sum_{s \in \{J\}, s \neq i} A_{s1} - \overline{A_{s1}} \tag{34}$$

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

$$E[y_{i\tau}] = E[y_{i\tau}|A_{it} = 1]\Pr[A_{it} = 1] + E[y_{i\tau}|A_{it} = 0]\Pr[A_{it} = 0]. \tag{35}$$

Furthermore, equation (2) shows that

$$E[\phi_{i\tau}] = E[y_{i\tau}|A_{it} = 1] - E[y_{i\tau}|A_{it} = 0]. \tag{36}$$

Using equations (35) and (36) we can solve for the two unknowns:

$$E[y_{i\tau}|A_{it} = 1] \quad = \quad E[y_{i\tau}] + E[\phi_{i\tau}]\Pr[A_{it} = 1] \tag{37}$$

$$E[y_{i\tau}|A_{it} = 0] \quad = \quad E[y_{i\tau}] - E[\phi_{i\tau}]\Pr[A_{it} = 0]. \tag{38}$$

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

$$E[A_{i\tau}|A_{it} = 1] - E[A_{i\tau}|A_{it} = 0] \quad = \quad 1 - prob[A_{i\tau} = 1|A_{it} = 0] \tag{39}$$

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

$$\begin{aligned} E[a_{i\tau}|A_{it} = 1] - E[a_{i\tau}|A_{it} = 0] &= 0 - E[a_{i\tau}|A_{it} = 0] \\ &= -prob[a_{i\tau} = 1|A_{it} = 0] \end{aligned} \tag{40}$$

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.

Inferring allowance rates given appeals

We must recover $\Pr[A_{it+1} = 1|A_{it} = 0, d_{it} = a]$ given the profiles for allowance and appeals. We do this by first noting that the Law of Total Probability gives us:

$$\begin{aligned} \Pr[A_{it+1} = 1|A_{it} = 0] &= \Pr[A_{it+1} = 1|A_{it} = 0, d_{it} = a]\Pr[d_{it} = a|A_{it} = 0] \\ &+ \Pr[A_{it+1} = 1|A_{it} = 0, d_{it} \neq a]\Pr[d_{it} \neq a|A_{it} = 0]. \end{aligned} \tag{41}$$

Since $\Pr[A_{it+1} = 1|A_{it} = 0, d_{it} \neq a] = 0$, equation (41) becomes

$$\Pr[A_{it+1} = 1|A_{it} = 0, d_{it} = a] = \tfrac{\Pr[A_{it+1}=1|A_{it}=0]}{\Pr[d_{it}=a|A_{it}=0]} \tag{42}$$

Again, using the Law of Total Probability,

$$\begin{aligned} \Pr[A_{it+1} = 1] &= \Pr[A_{it+1} = 1|A_{it} = 1]\Pr[A_{it} = 1] \\ &+ \Pr[A_{it+1} = 1|A_{it} = 0]\Pr[A_{it} = 0]. \end{aligned} \tag{43}$$

Since $\Pr[A_{it+1} = 1|A_{it} = 1] = 1$ we get

$$\Pr[A_{it+1} = 1|A_{it} = 0] = \tfrac{\Pr[A_{it+1}=1]-\Pr[A_{it}=1]}{1-\Pr[A_{it}=1]} \tag{44}$$

Similarly, we can calculate $\Pr[d_{it} = a|A_{it} = 0]$ also using the Law of Total probability

$$\Pr[d_{it} = a|A_{it} = 0] = \tfrac{\Pr[d_{it}=a]}{\Pr[A_{it}=0]} \tag{45}$$

Combining equations (42)-(45) yields:

$$\Pr[A_{it+1} = 1 | A_{it} = 0, d_{it} = a] \quad = \frac{\Pr[A_{it+1}=1] - \Pr[A_{it}=1]}{\Pr[d_{it}=a]} \tag{46}$$

All of the above can be conditioned on denial by an ALJ, all the right hand side objects are those presented in the right hand panel of figure 3.

Appendix D: Standard errors of the indirect inference estimator

This appendix derives standard error formulae for the indirect inference estimator. Relative to the usual GMM standard error formulae, we must confront three challenges. First, "our moment conditions" are estimated parameters. Second, we have panel data, so it is likely that residuals are correlated across equations. Third, because our estimates are for two cohorts, our data are unbalanced: if an individual is observed in one cohort, she is not observed in the other. This appendix describes the procedure for overcoming these obstacles.

Our procedure is to make the dynamic programming model match the OLS and IV estimated parameters. We match both IV and OLS estimates for participation, appeals, and allowance for both the cohorts (with average ages of 42 and 52 when first observed) over the 10 periods for which we have data. This gives us a total of 2 (OLS and IV) $\times$ 3 (participation, appeals, and allowance) $\times$ 2 (cohorts) $\times$ 10 (years of data) = 120 moments. In addition, we also match mean allowance of each cohort, and also the standard deviation of the judge allowance differential. This gives us 123 moment conditions in all. For the OLS and IV moments, the l^{th} moment condition (where $l \in \{1, ..., 120\}$ is an estimated equation) for the c^{th} cohort can be written according to the form

$$\hat{m}^{lc}(\theta) = \widehat{\phi^{lc}} - \phi^{lc}(\theta) \tag{47}$$

where $\widehat{\phi^{lc}}$ is a regression coefficient from either an OLS or IV regression and $\phi^{lc}(\theta)$ is the model-generated value. The de-meaned OLS regression is of the form

$$\dot{y}_i^{lc} = \phi^{lc} \dot{A}_{i1}^c + \dot{u}_i^{lc} \tag{48}$$

where $\dot{y}_i^{lc} = y_i^{lc} - \sum_{j=1}^{N_c} y_j^{lc}$ and $\dot{A}_{i1}^c = A_{i1}^c - \sum_{j=1}^{N_c} A_{j1}^c$ are de-meaned outcomes and de-meaned time 1 (i.e., ALJ) allowance decision, where means are constructed over all members of their cohort. N_c is the number of individuals in cohort c, where the two cohorts are those on average 42 and those on average 52 in they appeal for benefits. Borrowing the notation from equation (6), de-meaned IV regression

$$\widetilde{y}_i^{lc} = \phi^{lc}\widehat{\widetilde{A}}_{i1}^c + \widetilde{u}_i^{lc} \tag{49}$$

where $\widehat{\widetilde{A}}_{i1}^c = \widehat{\lambda_1^c \widetilde{j}_i}\hat{\gamma}_{1,-i}$ is estimated using the regression

$$\widetilde{A}_{i1}^c = \lambda_1^c \widetilde{j}_i \hat{\gamma}_{1,-i} + \epsilon_i \tag{50}$$

using all observations for cohort c, where $\widetilde{A}_{i1}$ is the time 1 (i.e., ALJ) decision.

Denote the vector of parameters θ, the number of observations N, and the 123×1 vector of estimated moment conditions as $\hat{m}(\theta)$. We minimize

$$\frac{N}{1+\varsigma}\hat{m}(\theta)' W \hat{m}(\theta), \tag{51}$$

where ς is the ratio of the number of individuals in the data to the number of simulated agents and W is a weighting matrix. We tried using both the identity weighting matrix (i.e., equal weighting) and the the inverse of the empirical variance-covariance matrix of moment conditions (i.e., optimal weighting). Both produced similar estimates [need to check this]. We describe the distribution of the standard errors and the overidentification statistics when using optimal weighting below. In order to estimate the optimal weighting matrix, we assume that $\lim_{N\to\infty}(N_c/N) = k^c$, where k^c is a constant. In other words, N_c and N converge to infinity at the same rate. Denoting by $\widehat{\theta}$ the estimated vector of coefficients and by θ_0 the true vector, the estimator has a sampling distribution given by

$$\sqrt{N}(\widehat{\theta} - \theta_0) \rightsquigarrow_D \mathcal{N}(0, (1+\varsigma)(D'WD)^{-1}), \tag{52}$$

$$D = \frac{\partial m(\theta_0)}{\partial \theta}, \tag{53}$$

To understand the variance covariance matrix of moment conditions, note that the moments are a collection of OLS coefficients and (exactly identified) IV coefficients. Thus for the case of OLS, the difference between true parameter and its estimated value, equation (47) is

$$m^{lc}(\theta_0)\widehat{\varepsilon}^{lc} = [(\dot{A}_1^c)'(\dot{A}_1^c)]^{-1}(\dot{A}_1^c)'\dot{u}^{lc} \tag{54}$$

where A_1^c and $\dot{u}^{lc}$ are the $N_c \times 1$ vectors of allowance at time 1 and residuals, where the i^{th} element is $\dot{u}_i^{lc} = \dot{y}_i^{lc} - \phi^{lc}\dot{A}_{i1}^c$ are from the OLS regression of equation (48). For the IV estimates it is

$$m^{kc}(\theta_0) = [(\widetilde{j}\hat{\gamma}_{1,-i})'(\widetilde{A}_1^c)]^{-1}(\widetilde{j\hat{\gamma}_{1,-i}})'\widetilde{u}^{kc} \tag{55}$$

where $\widetilde{A}_1^c$, $\widetilde{j}\hat{\gamma}_{1,-i}$ and $\widehat{\widetilde{u}}^{kc}$ are the $N_c \times 1$ vectors of de-meaned allowance, the instrumental variables, and IV residuals. For the mean allowance equations equation (47) is

$$m^{Ac}(\theta_0) = \frac{1}{N_c}\sum_{i=1}^{N_c}(A_{i1}^c - E[A_{i1}^c]) \tag{56}$$

where $E[A_{i1}^c]$ is model predicted mean allowance. Lastly, we match the variance of the judge allowance differential. Equation (34) shows how we calculate the judge allowance differential in the data: $\widetilde{j}_i\hat{\gamma}_{1,-i} = \frac{1}{N_j-1}\sum_{s\in\{J\},s\neq i}(A_{s1} - \overline{A_{s1}})$. We use the full sample of N observations to calculate this object. The estimated variance of this object is:

$$\frac{1}{N}\sum_{i=1}^{N}\left(\widetilde{j}_i\hat{\gamma}_{1,-i} - \overline{\widetilde{j}_i\hat{\gamma}_{1,-i}}\right)^2 \tag{57}$$

Appendix E describes how we calculate the same variance in the model. The difference between the object in equation (57) and the asymptotic variance is:

$$m^{\gamma}(\theta_0) = \frac{1}{N}\sum_{i=1}^{N}\left(\left(\widetilde{j}_i\hat{\gamma}_{1,-i} - \overline{\widetilde{j}_i\hat{\gamma}_{1,-i}}\right)^2 - E\left(\widetilde{j}_i\hat{\gamma}_{1,-i} - \overline{\widetilde{j}_i\hat{\gamma}_{1,-i}}\right)^2\right). \tag{58}$$

Thus equations (56) (for both cohorts) and (58) show the final three moment conditions that we match. The optimal weighting matrix is the inverse variance-covariance matrix of the moment conditions: $W^{-1} = E[m(\theta_0)m(\theta_0)']$. In order to estimate this object we replace all expectations with sample means (e.g., we assume $E[A_{i1}^c] \approx \frac{1}{N_c}\sum_{i=1}^{N_c} A_{i1}^c$ in equation (56)) and all residuals given true parameters with residuals given estimated parameters (e.g., we assume $\dot{u}_i^{lc} = \dot{y}_i^{lc} - \phi^{lc}\dot{A}_{i1}^c \approx \hat{\dot{u}}_i^{lc} = \dot{y}_i^{lc} - \hat{\phi}^{lc}\dot{A}_{i1}^c$, where $\hat{\phi}^{lc}$ the estimated value of ϕ^{lc}). Given the estimated moment conditions, the sample analogs of equations (54), (55), (56), and (58) give rise to the following variance-covariance matrix $\hat{W}^{-1}$, where $\hat{W}_{l,k}^{-1}$ is given in table 7.

Although calculation of most of the elements of $\hat{W}^{-1}$ is straightforward, calculation of the sample analog of $E[m^{Ac}(\theta_0)m^{\gamma}(\theta_0)'] = \left(\frac{N}{N_c}\right)\frac{1}{N_c}\sum_{i=1}^{N}\sum_{n=1}^{N_c}(A_{n1}^c - E[A_{n1}^c]) \times \left(\left(\widetilde{j}_i\hat{\gamma}_{1,-i} - \overline{\widetilde{j}_i\hat{\gamma}_{1,-i}}\right)^2 - E\left(\widetilde{j}_i\hat{\gamma}_{1,-i} - \overline{\widetilde{j}_i\hat{\gamma}_{1,-i}}\right)^2\right)$ merits explicit derivation. Direct calculation of this object is computationally infeasible since it is the sum of $N \times N_c$ objects. But using the definition of $\widetilde{j}_i\hat{\gamma}_{1,-i}$ from equation (34), note that $\overline{\widetilde{j}_i\hat{\gamma}_{1,-i}} \approx 0$ and thus

$$\begin{aligned}
& E[m^{Ac}(\theta_0)m^{\gamma}(\theta_0)'] \\
= \; & E\left[\left(\frac{N}{N_c}\right)\frac{1}{N_c}\sum_{i=1}^{N}\sum_{n=1}^{N_c}(A_{n1}^c - E[A_{n1}^c]) \times \left(\left(\widetilde{j}_i\hat{\gamma}_{1,-i} - \overline{\widetilde{j}_i\hat{\gamma}_{1,-i}}\right)^2 - E\left(\widetilde{j}_i\hat{\gamma}_{1,-i} - \overline{\widetilde{j}_i\hat{\gamma}_{1,-i}}\right)^2\right)\right] \\
= \; & E\left[\left(\frac{N}{N_c}\right)\frac{1}{N_c}\sum_{i=1}^{N}\sum_{n=1}^{N_c}(A_{n1}^c - E[A_{n1}^c]) \times \left(\left(\widetilde{j}_i\hat{\gamma}_{1,-i}\right)^2 - E\left(\widetilde{j}_i\hat{\gamma}_{1,-i}\right)^2\right)\right] \\
= \; & E\left[\left(\frac{N}{N_c}\right)\frac{1}{N_c}\sum_{i=1}^{N}\sum_{n=1}^{N_c}1\{n,\, i \text{ heard by same judge}\}(A_{n1}^c - E[A_{n1}^c]) \times \left(\left(\widetilde{j}_i\hat{\gamma}_{1,-i}\right)^2 - E\left(\widetilde{j}_i\hat{\gamma}_{1,-i}\right)^2\right)\right]
\end{aligned} \tag{59}$$

Equation (59) is not 0 because the same judge who heard individual i's case also potentially heard individual n's case, affecting the probability of allowance of both cases. Assuming that judges handle similar number of cases, the probability both i and n were heard by the same judge is equal to $\frac{1}{\text{number of judges}}$. If the number of cases heard by each judge is large, then $E(A_{n1}^c - E[A_{n1}^c])1\{n,\, i \text{ heard by same judge}\}\left(\widetilde{j}_i\hat{\gamma}_{1,-i}\right)^2 \approx \frac{1}{\text{number of judges}}E(A_{i1}^c - E[A_{i1}^c])\left(\widetilde{j}_i\hat{\gamma}_{1,-i}\right)^2$. Furthermore, as done above, assume $(A_{n1}^c - E[A_{n1}^c]) \approx \dot{A}_{i1}^c$. Then equation (59) equals

$$\begin{aligned}
& E\left(\frac{N}{N_c}\right)\frac{1}{N_c}\sum_{i=1}^{N_c}\sum_{n=1}^{N}\frac{1}{\text{number of judges}}(A_{i1}^c - E[A_{i1}^c])\left(\widetilde{j}_i\hat{\gamma}_{1,-i}\right)^2 \\
\approx \; & \left(\frac{N}{N_c}\right)^2\frac{1}{\text{number of judges}}\sum_{i=1}^{N_c}\dot{A}_{i1}^c\left(\widetilde{j}_i\hat{\gamma}_{1,-i}\right)^2
\end{aligned} \tag{60}$$

Assuming that the number of judges grows at the same rate as the number of observations, this will converge to a non-stochastic non-degenerate object.

The first three lines table 7 are derived using the assumption $E[u_i^{kc}u_j^{lc}] = 0$ for $i \neq j$:

Table 7: Elements of Covariance Matrix

$\hat{W}^{-1}_{l,k} =$	if
$\left(\frac{N}{N_c}\right) \sum_{i=1}^{N_c} \widehat{\dot{u}}_i^{kc} \widehat{\dot{u}}_i^{lc} [(\dot{A}_1^c)'(\dot{A}_1^c)]^{-1}$	l, k are OLS moments, same cohort
$\left(\frac{N}{N_c}\right) \sum_{i=1}^{N_c} \widehat{\dot{u}}_i^{kc} \widehat{\widetilde{u}}_i^{lc} [(\dot{A}_1^c)'(\dot{A}_1^c)]^{-1} [(\dot{A}_1^c)'(\widetilde{j\hat{\gamma}_{1,-i}})][(\widetilde{j\hat{\gamma}_{1,-i}})'(\widetilde{A}_1^c)]^{-1}$	l an IV, k an OLS moment, same cohort
$\left(\frac{N}{N_c}\right) \sum_{i=1}^{N_c} \widehat{\widetilde{u}}_i^{kc} \widehat{\widetilde{u}}_i^{lc} \times$ $((\widetilde{j\hat{\gamma}_{1,-i}})'(\widetilde{A}_1^c))^{-1} [(\widetilde{j\hat{\gamma}_{1,-i}})'(\widetilde{j\hat{\gamma}_{1,-i}})][(\widetilde{j\hat{\gamma}_{1,-i}})'(\widetilde{A}_1^c)]^{-1}$	l, k are IV moments, same cohort
$\left(\frac{N}{N_c}\right) \frac{1}{N_c} \sum_{i=1}^{N_c} (\dot{A}_{i1}^c)^2$	if $l = k$ is allowance for cohort c
$\frac{1}{N} \sum_{i=1}^{N} \left(\left(\widetilde{j}_i \hat{\gamma}_{1,-i} - \overline{\widetilde{j}_i \hat{\gamma}_{1,-i}}\right)^2 - \frac{1}{N} \sum_{i=1}^{N} \left(\widetilde{j}_i \hat{\gamma}_{1,-i} - \overline{\widetilde{j}_i \hat{\gamma}_{1,-i}}\right)^2 \right)^2$	$l = k$ is variance of judge allowance differenti
$\left(\frac{N}{N_c}\right)^2 \frac{1}{\text{number of judges}} \sum_{i=1}^{N_c} \dot{A}_{i1}^c \left(\widetilde{j}_i \hat{\gamma}_{1,-i}\right)^2$	l is variance of judge allowance differential, k allowance for cohort c
0	l is allowance for cohort c, $l \neq k$
0	l is variance of judge allowance rate, $l \neq k$
0	l, k are for different cohorts

i.e., there is no correlation across individuals. The fourth line is derived by noting that by construction outcome residuals are uncorrelated with the allowance residuals. The fifth line also utilizes no correlation across individuals: thus cross cohort correlations should be 0 because they include different individuals. Thus the matrix $\hat{W}^{-1}$ has block diagonal form: all elements of $\hat{W}^{-1}$ referring to members of different cohorts are equal to 0.

Assuming that the model is properly specified, the objective function in equation (51) is distributed χ^2_{123-12}.

Appendix E: Generating moment conditions in the model

We generate moment conditions as follows.

- Draw two cohorts of simulated agents: one age 42 at time $t = 0$, one 52 at time $t = 0$.
- For each cohort, draw a total of $s = 1, ..., S$ (we use $S = 5,000$) agents, where each agent is a $\upsilon_p, \upsilon_a, \upsilon_h$ triple

- Solve for optimal decision rules for each simulated individual, including whether each agent should appeal at the ALJ stage. If it is not optimal for them to appeal at the ALJ stage, they are dropped from the simulated sample.

- Assign these individuals to a time period 0 judge, with allowance theshold $\chi_{h0} = \chi_0 + \chi_{j_0}$

- Calculate the mean allowance rate for each "judge", which is a centile of the χ_{h0} distribution. We want the values of α_0, $V[\chi_h]$ so that the mean time 0 allowance probability for those who apply is .65, and 1 standard deviation of the judge allowance probability is 0.0659 (where 0.0659 is the standard deviation of the difference of the average allowance rate of all judges, and the average allowance rate of each judge).

  - Note that the average allowance rate is (ignoring the fact that not all simulated individuals appeal at time 0)

$$\begin{aligned}
\Pr(A_1 = 1|a, A_0 = 0) &= \Pr(\upsilon_h > \chi_{h0}|a, A_t = 0) \\
&= \Pr(\mu_h + \epsilon_h > \alpha_0 + \chi_{j_0}) \\
&= \Pr(\mu_h - \alpha_0 > -\epsilon_h + \chi_{j_0}) \\
&= \Phi\left(\frac{\mu_h - \alpha_0}{\sqrt{V[\upsilon_h] + V[\chi_j]}}\right) \qquad (61)
\end{aligned}$$

    where $\upsilon_h = \mu_h + \epsilon_h$.

  - Note average allowance probability of a judge with threshold $\alpha_0 + \chi_{j_0}$ is (again ignoring the fact that not all simulated individuals appeal at time 0)

$$\begin{aligned}
\Pr(A_1 = 1|a, A_0 = 0, \chi_{j_0}) &= \Pr(\upsilon_h > \chi_{h0}|a, A_t = 0, \chi_{j_0}) \\
&= \Pr(\mu_h - (\alpha_0 + \chi_{j_0}) > -\epsilon_h) \\
&= \Phi\left(\frac{\mu_h - (\alpha_0 + \chi_{j_0})}{\sqrt{V[\upsilon_h]}}\right) \qquad (62)
\end{aligned}$$

  We calculate the probabilities $\Pr(A_1 = 1|a, A_0 = 0)$ and $\Pr(A_1 = 1|a, A_0 = 0, \chi_{j_0})$ numerically, but equations (61) and (61) are useful checks on the accuracy of the numerical results.

- 1 standard deviation of the judge allowance differential is 0.0659. We calculate this as follows:

1. Calculate the mean allowance rate in the simulated sample (which should be close to the probability in equation (61).
2. Sort the simulated data by χ_{j_0}. Give every observation a centile rank.
3. Calculate the mean allowance probability $\bar{A}_j$ at every centile of χ_{j_0} distribution. Treat each centile as a "judge". There will be 100 centiles. There are 5,000 simulated individuals in each cohort, for a total of 10,000 simulated individuals. Thus there will be 100 simulated individuals observations per judge.
4. Next we need the judge allowance differential for each simulated individual s. This is the judge allowance for each simulated individual (calculated in the above step), taking out simulated individual s (so it will be the allowance rate over the 99 other simulated individuals in simulated individual s's centile). So it is the value of $j_s\widehat{\gamma}_{1,-s} = \bar{A}_j(100/99) - (1/99)A_{0s} - \Pr(A_1 = 1|a, A_0 = 0)$.
5. We calculate the allowance rate differential for each simulated individual s. We then take the standard deviation of this object. We find parameters so that the simulated allowance differential matches the 0.0659 found in the data.

- For each simulated individual who found it optimal to appeal, simulate whether that individual is allowed benefits, where from equation (13) we know that $\Pr(A_1 = 1|a, A_0 = 0, \upsilon_h) = \Phi\left(\frac{\upsilon_h - \alpha_0}{\sqrt{V[\chi_j]}}\right)$ [note: that is right if $V[\upsilon_h]$ is variance of judge fixed effects].
- At this point we can go back and recalculate the mean and sd of judge allowance rates. The de-meaned judge allowance rate (the time 1 allowance rate, conditional on the centile of the υ_{j0} distribution for those who appeal at time 0) is the instrument.
- Next we simulate the model. Each individual now has a value of time 1 allowance. We solve and simulate the model for all time periods. For those allowed at time 1, they are allowed benefits in all subsequent periods and never work or appeal.
- Regress participation, allowance, appeals, on time dummies and time dummies × time 1 allowance. The coefficients should be identical to mean participation , 1-(mean allowance), -(mean appeals), conditional on time 1 denial.

- demean participation, allowance, appeals at each age.

- Estimate parameters using IV. The two stages are:

  1. Regress de-meaned allowance $A_{s1} - \bar{A}_1$ (where $\bar{A}_1$ is the mean allowance probability for all simulated individuals who appealed at time 0) on $j_s \widehat{\gamma}_{1,-s}$, ie $A_{s1} - \bar{A}_1 = \theta j_s \widehat{\gamma}_{1,-s} + \varepsilon_{s1}$. Predicted allowance is then $\hat{\theta} j_s \widehat{\gamma}_{1,-s}$
  2. Regress de-meaned (by age) participation, allowance, appeals at each age on predicted allowance $\hat{\theta} j_s \widehat{\gamma}_{1,-s}$.

Still to do

- probably need an age trend in preferences to work to get the model to match falling participation rates.

- address issue that many of those allowed work. simple fix: make "participation" earnings>SGA.

- need to do this for fixed age groups, to that we can better handle the timing and age issues. do it for ages 50-54 at time of assignment (so 60-64 in final time period), as well as 40-44 at time of assignment to a judge. In model assume everyone is of the average age of that group (ie, age 42 and 52 at time of time of assignment)

- Policy experiments

  – welfare gains of eliminating appeals option (need to think about health shocks if so)

  – welfare gains of standardizing the appeals process (ie eliminating uncertainty)

  – clarify the issues associated with using estimates 3 (or other years after assignment). on the one hand, many people continue to appeal, even 3 years after assignment. on the other hand, many of those who never would have worked (and thus appealed) are now counted in the allowed group. in principle, this means that using allowance 3 years after assignment could either lead to an overstatement of understatement of the true causal effect. What is the "optimal" number of years after assignmnent to use?

Working Paper Series

A series of research studies on regional economic issues relating to the Seventh Federal Reserve District, and on financial and economic topics.

Why Has Home Ownership Fallen Among the Young? **WP-09-01**
Jonas D.M. Fisher and Martin Gervais

Why do the Elderly Save? The Role of Medical Expenses **WP-09-02**
Mariacristina De Nardi, Eric French, and John Bailey Jones

Using Stock Returns to Identify Government Spending Shocks **WP-09-03**
Jonas D.M. Fisher and Ryan Peters

Stochastic Volatility **WP-09-04**
Torben G. Andersen and Luca Benzoni

The Effect of Disability Insurance Receipt on Labor Supply **WP-09-05**
Eric French and Jae Song

CEO Overconfidence and Dividend Policy **WP-09-06**
Sanjay Deshmukh, Anand M. Goel, and Keith M. Howe

Do Financial Counseling Mandates Improve Mortgage Choice and Performance? **WP-09-07**
Evidence from a Legislative Experiment
Sumit Agarwal,Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff

Perverse Incentives at the Banks? Evidence from a Natural Experiment **WP-09-08**
Sumit Agarwal and Faye H. Wang

Pay for Percentile **WP-09-09**
Gadi Barlevy and Derek Neal

The Life and Times of Nicolas Dutot **WP-09-10**
François R. Velde

Regulating Two-Sided Markets: An Empirical Investigation **WP-09-11**
Santiago Carbó Valverde, Sujit Chakravorti, and Francisco Rodriguez Fernandez

The Case of the Undying Debt **WP-09-12**
François R. Velde

Paying for Performance: The Education Impacts of a Community College Scholarship Program for Low-income Adults **WP-09-13**
Lisa Barrow, Lashawn Richburg-Hayes, Cecilia Elena Rouse, and Thomas Brock

Establishments Dynamics, Vacancies and Unemployment: A Neoclassical Synthesis **WP-09-14**
Marcelo Veracierto

Working Paper Series *(continued)*

Working Paper Series *(continued)*

Robustness and Macroeconomic Policy **WP-10-04**
Gadi Barlevy

Benefits of Relationship Banking: Evidence from Consumer Credit Markets **WP-10-05**
Sumit Agarwal, Souphala Chomsisengphet, Chunlin Liu, and Nicholas S. Souleles

The Effect of Sales Tax Holidays on Household Consumption Patterns **WP-10-06**
Nathan Marwell and Leslie McGranahan

Gathering Insights on the Forest from the Trees: A New Metric for Financial Conditions **WP-10-07**
Scott Brave and R. Andrew Butters

Identification of Models of the Labor Market **WP-10-08**
Eric French and Christopher Taber

Public Pensions and Labor Supply Over the Life Cycle **WP-10-09**
Eric French and John Jones

Explaining Asset Pricing Puzzles Associated with the 1987 Market Crash **WP-10-10**
Luca Benzoni, Pierre Collin-Dufresne, and Robert S. Goldstein

Prenatal Sex Selection and Girls' Well‐Being: Evidence from India **WP-10-11**
Luojia Hu and Analía Schlosser

Mortgage Choices and Housing Speculation **WP-10-12**
Gadi Barlevy and Jonas D.M. Fisher

Did Adhering to the Gold Standard Reduce the Cost of Capital? **WP-10-13**
Ron Alquist and Benjamin Chabot

Introduction to the *Macroeconomic Dynamics*:
Special issues on money, credit, and liquidity **WP-10-14**
Ed Nosal, Christopher Waller, and Randall Wright

Summer Workshop on Money, Banking, Payments and Finance: An Overview **WP-10-15**
Ed Nosal and Randall Wright

Cognitive Abilities and Household Financial Decision Making **WP-10-16**
Sumit Agarwal and Bhashkar Mazumder

Complex Mortgages **WP-10-17**
Gene Amromin, Jennifer Huang, Clemens Sialm, and Edward Zhong

The Role of Housing in Labor Reallocation **WP-10-18**
Morris Davis, Jonas Fisher, and Marcelo Veracierto

Why Do Banks Reward their Customers to Use their Credit Cards? **WP-10-19**
Sumit Agarwal, Sujit Chakravorti, and Anna Lunn

Working Paper Series *(continued)*

The impact of the originate-to-distribute model on banks
before and during the financial crisis WP-10-20
Richard J. Rosen

Simple Markov-Perfect Industry Dynamics WP-10-21
Jaap H. Abbring, Jeffrey R. Campbell, and Nan Yang

Commodity Money with Frequent Search WP-10-22
Ezra Oberfield and Nicholas Trachter

Corporate Average Fuel Economy Standards and the Market for New Vehicles WP-11-01
Thomas Klier and Joshua Linn

The Role of Securitization in Mortgage Renegotiation WP-11-02
Sumit Agarwal, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff

Market-Based Loss Mitigation Practices for Troubled Mortgages
Following the Financial Crisis WP-11-03
Sumit Agarwal, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff

Federal Reserve Policies and Financial Market Conditions During the Crisis WP-11-04
Scott A. Brave and Hesna Genay

The Financial Labor Supply Accelerator WP-11-05
Jeffrey R. Campbell and Zvi Hercowitz

Survival and long-run dynamics with heterogeneous beliefs under recursive preferences WP-11-06
Jaroslav Borovička

A Leverage-based Model of Speculative Bubbles (Revised) WP-11-07
Gadi Barlevy

Estimation of Panel Data Regression Models with Two-Sided Censoring or Truncation WP-11-08
Sule Alan, Bo E. Honoré, Luojia Hu, and Søren Leth–Petersen

Fertility Transitions Along the Extensive and Intensive Margins WP-11-09
Daniel Aaronson, Fabian Lange, and Bhashkar Mazumder

Black-White Differences in Intergenerational Economic Mobility in the US WP-11-10
Bhashkar Mazumder

Can Standard Preferences Explain the Prices of Out-of-the-Money S&P 500 Put Options? WP-11-11
Luca Benzoni, Pierre Collin-Dufresne, and Robert S. Goldstein

Business Networks, Production Chains, and Productivity:
A Theory of Input-Output Architecture WP-11-12
Ezra Oberfield

Equilibrium Bank Runs Revisited WP-11-13
Ed Nosal

Working Paper Series *(continued)*

Are Covered Bonds a Substitute for Mortgage-Backed Securities? **WP-11-14**
Santiago Carbó-Valverde, Richard J. Rosen, and Francisco Rodríguez-Fernández

The Cost of Banking Panics in an Age before "Too Big to Fail" **WP-11-15**
Benjamin Chabot

Import Protection, Business Cycles, and Exchange Rates:
Evidence from the Great Recession **WP-11-16**
Chad P. Bown and Meredith A. Crowley

Examining Macroeconomic Models through the Lens of Asset Pricing **WP-12-01**
Jaroslav Borovička and Lars Peter Hansen

The Chicago Fed DSGE Model **WP-12-02**
Scott A. Brave, Jeffrey R. Campbell, Jonas D.M. Fisher, and Alejandro Justiniano

Macroeconomic Effects of Federal Reserve Forward Guidance **WP-12-03**
Jeffrey R. Campbell, Charles L. Evans, Jonas D.M. Fisher, and Alejandro Justiniano

Modeling Credit Contagion via the Updating of Fragile Beliefs **WP-12-04**
Luca Benzoni, Pierre Collin-Dufresne, Robert S. Goldstein, and Jean Helwege

Signaling Effects of Monetary Policy **WP-12-05**
Leonardo Melosi

Empirical Research on Sovereign Debt and Default **WP-12-06**
Michael Tomz and Mark L. J. Wright

Credit Risk and Disaster Risk **WP-12-07**
François Gourio

From the Horse's Mouth: How do Investor Expectations of Risk and Return
Vary with Economic Conditions? **WP-12-08**
Gene Amromin and Steven A. Sharpe

Using Vehicle Taxes To Reduce Carbon Dioxide Emissions Rates of
New Passenger Vehicles: Evidence from France, Germany, and Sweden **WP-12-09**
Thomas Klier and Joshua Linn

Spending Responses to State Sales Tax Holidays **WP-12-10**
Sumit Agarwal and Leslie McGranahan

Micro Data and Macro Technology **WP-12-11**
Ezra Oberfield and Devesh Raval

The Effect of Disability Insurance Receipt on Labor Supply: A Dynamic Analysis **WP-12-12**
Eric French and Jae Song