Social Security Disability Reform and Implications for Employment

Hilary Hoynes
UC Berkeley and NBER
hoynes@berkeley.edu

Nicole Maestas
Harvard University and NBER
maestas@hcp.med.harvard.edu

Alexander Strand
Social Security Administration
Alexander.Strand@ssa.gov

September 2023

ABSTRACT

Beginning in 2010, the Social Security Administration began implementing a series of data-driven policy initiatives to improve the quality and consistency of disability case reviews performed by Administrative Law Judges (ALJs). The policy initiatives included revised training curricula for ALJs, new decision-support tools, and direct feedback about common decision errors. Around this time, the hearing-level allowance rate dropped by 22 percentage points and employment among people with disabilities rose for the first time in decades. However, it is not known what, if any, role the policy initiatives played in these changes. This paper investigates the impacts of the policy initiatives on the hearing-level allowance rate and the effects of these changes in allowance rates on the employment and earnings of applicants. We conclude that the policy initiatives accounted for 28-36% of the decline in the hearing-level allowance rate and, as the margin of allowance gradually tightened, led to increased work activity among many denied applicants.

I. Introduction

Labor force participation rates for men have been declining for decades while for women they have declined since 2000 (e.g., Abraham and Kearney 2020, Black et al, 2017). The downward trend has been especially pronounced for people with a work disability, who experienced a 40 percent reduction in their employment rate between 1988 and 2014 (Figure 1). These trends renewed attention to the possible role of the Social Security Disability Insurance (SSDI) program, which tripled in size over the same period (see e.g., Autor & Duggan 2003, Daly and Burkhauser 2011).[1] However, it is little-noticed that SSDI caseloads began to decline in 2015 after a five-decade increase (Figure 2). In fact, new SSDI awards began to decline in 2010 (Figure 3). At the same time, the early 2010s saw a turnaround in employment among disabled persons that continues today (Figure 1). In this paper, we investigate the reasons for the decline in SSDI beneficiaries and the possible role in the employment of people with disabilities.

Beginning in 2010, the Social Security Administration began implementing a series of data-driven policy initiatives to improve the quality and consistency of disability case reviews performed by Administrative Law Judges (ALJs) (Ray and Lubbers 2015). The policy initiatives included new training curricula for ALJs, new decision-support tools, and direct feedback about common decision errors. Around this time, the hearing-level allowance rate dropped by more than 20 percentage points (Figure 4). However, it is not known what, if any, role SSA policy played in the decline in allowance rates and the increase in employment among people with

---

[1] An earlier literature investigated SSDI caseload growth in the 1960s and 1970s (see e.g., Parsons 1980, 1982; Haveman and Wolfe 1984a, 1984b; Bound 1989; Bound and Waidmann 1992).

disabilities. Our paper fills this gap.

In this paper, we describe the specific steps taken by SSA to improve the quality and consistency of decisions before ALJs. Using SSA administrative data, we build the first comprehensive measurement of the SSDI policy reforms at the ALJ judge and ALJ office level. We then use SSA administrative data on all appellate decisions between 2007 and 2015 to estimate the impact of the policy reforms on appellate allowance rates and determine what share of the decline in allowance rates was due to policy reforms implemented by SSA. Taking advantage of the random assignment of judges to cases and combining the appellate data with data from the SSA Master Earnings File, we use instrumental variables to estimate the effect of the decline in allowance rates on the employment and earnings of applicants. We then apply these results to inform the causes of the turnaround in employment rates among people with disabilities.

We find a 22 percentage point reduction in appellate allowance rates (from a peak of 69% in 2008). The SSA policy reforms explain 28-36% of this decline; and we find little role of changes in caseload characteristics or labor market cycles. Our instrumental variables estimates show that SSDI allowance leads to a 22 percentage point decline in annual employment, a 14 point decline in employment above SGA, and a $4,100 decline in annual earnings four years after the decision. Because the reform reduced the SSDI allowance rate, our estimates imply that as the reforms played out, many denied claimants on the margin of allowance increased work activity. The increase in labor supply is large in the context of low average labor supply among SSDI applicants; however, the majority of applicants did not return to work and a substantial

2

fraction of those who did work earned below SGA.

Our estimates of the causal impact on earnings are 12-15 percent smaller than those of French and Song (2014), who estimated the effect of SSDI allowance on the labor supply outcomes of appellants in the 1990s. This is consistent with the policy reforms having made the eligibility standard stricter; as the program was gradually tightened over a period of several years, those on the margin of allowance had lower work capacity and therefore a smaller SSDI treatment effect than the marginal claimant before the reforms.

This paper contributes to three areas. First, there is a significant literature on SSDI caseload growth, the role of labor markets, and moral hazard (Parsons 1980, 1982; Haveman and Wolfe 1984a, 1984b; Bound 1989; Bound and Waidmann 1992; Autor & Duggan 2003, Black et al. 2002; Maestas, Mullen & Strand 2013; French & Song 2014; Liebman 2015; Charles et al. 2018; Maestas, Mullen & Strand 2021). All of these studies investigate the period of SSDI program expansion. Second, there is a growing literature on the value of DI and targeting (Low & Pistaferri, 2015, Gelber, Moore & Strand 2017; Deshpande, Gross, & Su, 2021; Deshpande & Lockwood 2022; Deshpande & Mueller-Smith 2022). Third, we make a significant contribution to the small literature on caseload decline. To date this work has focused primarily on caseload decline in other countries, most prominently disability reform in the Netherlands (e.g., Koning & Lindeboom 2015; Degroot & Koning 2015), with one recent paper on SSDI reform (Liu & Quimby, 2023).[2] We advance these literatures by comprehensively examining the first major

---

[2] Liu & Quimby (2023) use and approach similar to Liebman (2015) and simulate the role of demographic, labor markets and changes in SSDI.

reform of the SSDI program in decades, by analyzing a period of program *contraction* without precedent in the history of SSDI, and by using these reforms to generate new causal estimates of the impact of SSDI on earnings and employment.

The rest of the paper proceeds as follows. In Section II we describe the SSDI program and in Section III we describe the SSDI reforms. In Section IV we describe our data and in Section V we present out empirical methods. We present results in Section VI and conclude in Section VII.

II. Social Security Disability Insurance Program

The U.S. Social Security Disability Insurance (SSDI) program is a contributory social insurance program that pays monthly cash benefits to individuals who become disabled, provided they are insured for benefits and meet the program's medical and vocational criteria. SSDI is closely related to the Supplemental Security Income (SSI) program, which is a federal welfare program that pays cash benefits to individuals who are disabled (in addition to elderly adults), provided they have very low income and assets and meet the same medical and vocational criteria as SSDI. Applications for SSDI and SSI benefits are processed by the same administrative review structure and are evaluated concurrently for both programs when the non-medical eligibility criteria are met for both.

Individuals are insured for SSDI benefits if they have earned at least one work credit (and made corresponding payroll tax contributions) for every calendar year elapsing between age 21 and the year they became disabled, and if half of the work credits were earned recently. SSDI benefits are a function of average career earnings and the average monthly benefit for disabled

4

workers was $1,282 in 2021 (SSA 2021). In addition to cash benefits, SSDI recipients receive health insurance through Medicare.

The process of claiming disability benefits can be long and complicated, and a favorable outcome is not assured—even for people who have made the required payroll tax contributions to become insured for SSDI. SSA must determine that the applicant is "unable to engage in any substantial gainful activity (SGA) because of a medically determinable physical or mental impairment(s) that is expected to result in death or that has lasted or is expected to last for a continuous period of at least 12 months." The practical application of this standard to modern disability cases requires complex judgments, and results in substantial variation in outcomes across disability adjudicators (Maestas, Mullen and Strand, 2013).

To apply for SSDI or SSI benefits, applicants complete paperwork online or in person at a local Social Security Field Office. Field office staff first verify that the applicant is not currently engaging in substantial gainful activity—defined as earning $1,350 per month or more in 2022. Next field office staff assist applicants in developing their applications including information on personal information about the applicant, including English language proficiency, current work activity, job history (over the last 15 years), a report of medical conditions, a list of medications, and a description of medical treatments received from each medical provider. Medical records from relevant providers can be submitted with the application.

The application is sent to the state office of the Disability Determination Service (DDS), where it is assigned to a disability examiner. The examiner uses a five-step sequential decision process to decide whether the applicant meets SSA's disability standard. The first step is to

verify that the applicant is not earning above the SGA level and as noted above, this step is done by field office staff. In Step 2, the examiner screens the application for short-duration or non-severe impairments and denies those cases. In Step 3, the examiner checks the applicant's medical diagnoses against SSA's Listing of Impairments (severe medical conditions for which it is presumed the claimant has no residual work capacity) and allows those cases. The remaining cases proceed to Steps 4 and 5, where the examiner uses vocational information—education, training, work history, and age—in combination with an assessment of the claimant's physical and/or mental residual functional capacities. In Step 4, the examiner considers the applicants' work history and assesses whether the applicant has the functional capacity to perform any of his or her past jobs. If yes, the application is denied. If no, then the examiner proceeds to Step 5, in which the applicant's education, training, work experience, age and residual functional capacity are used to determine if the applicant has the capacities to perform any job in the national economy. If so, the applicant is denied. If not, the applicant is allowed.

Applicants who are denied benefits in their initial review have the option of appealing. The first level of appeal is called Reconsideration and entails a second review of the application by another examiner in the same DDS office. If the application is denied on Reconsideration, the applicant can request a hearing before an Administrative Law Judge (ALJ). Most hearings are conducted at the nearest Social Security regional hearing office in the presence of the claimant and any legal representative enlisted by the claimant. If an applicant is denied by an ALJ, a further appeal can be pursued with the SSA Appeals Council, and then Federal Court. At the appellate level, the same five-step decision process is used; that said, appeals often involve the submission of updated medical evidence, which if the applicant's condition has worsened, can

increase the likelihood of an award.

If a claimant's application is successful, they receive any past-due benefits in the form of a lump-sum back payment. The back payments cover the number of months between the established date of disability onset and the date of approval minus a five-month waiting period. Retroactive benefits are available for SSDI awardees up to a maximum of 12 months. Overall, initial determinations are completed in about three months' time but the appeal process extends total processing time for appellants to nearly two years on average.

III. SSDI Policy Environment and Reforms

In this paper, we focus on policy changes taking place at the appellate hearing level to improve consistency among ALJ decisions. Here we provide background on the ALJs and we describe the three sets of policy reforms they faced.

First, from 1999 through 2007, there was a disruption in the hiring of ALJs across the federal government, which heavily impacted the Social Security Administration, the largest employer of ALJs (GAO 2010). Initially, the Office of Personnel Management (OPM) was involved in a lawsuit concerning how the veterans' preference should be counted in the ALJ hiring process. During 1999 to 2003, the list of qualified candidates—known as the *ALJ Register*—was suspended except for brief periods. After resolution of the lawsuit, the register was reopened but OPM did not allow new candidates to take the examination required to enter the register. In October 2007, OPM established a new register with a new exam and federal agencies began hiring from this register shortly thereafter. In fiscal year 2008, SSA hired 185 ALJs, adding to 1,007 incumbents (GAO 2010).

The effect of this near total suspension in hiring was that the ALJ workforce became skewed towards the near-retirement ages. Even when including the new hires, 51 percent of all ALJs were eligible to retire by the end of 2008 (GAO 2010). Further, there was a large gap in tenure between incumbents and the new cohorts of ALJs.

With the suspension of ALJ hiring, the appellate backlog—i.e., applicants awaiting a hearing before an ALJ—increased substantially (Figure 5) from 10.7 months in 2004 to 15.6 months in 2007.[3] After the start of hiring from the new register, the backlog declined quickly, before stabilizing at 10-11 months on average by 2010. At this same time, the Great Recession led to increases in unemployment, peaking in October 2009. The rise in unemployment rates during the Great Recession led to an increase in disability applications, which followed with a lag and peaked six months later, in April 2010 (Maestas, Mullen & Strand 2021).

Second, shortly after the resumption of hiring from the new register, SSA launched additional initiatives designed to reduce the backlog. In November 2007, SSA implemented the Senior Adjudicative Attorney (SAA) program with a goal of addressing older pending claims (SSA:OIG 2016). SAAs have the authority to issue "on-the-record" allowances. These are decisions where the existing documentary evidence can support an allowance and, thus, a hearing with an ALJ is not necessary. A typical SAA case starts with a rejection of a claim at the initial level and then additional medical evidence comes in while a claimant is awaiting a hearing with an ALJ, particularly when the wait is long and perhaps with the assistance of a

---

[3] In this figure, the backlog is defined as the number of pending hearings divided by the average hearings per month. The resulting measure is the number of months it would take to eliminate the backlog if counterfactually there were no further applications.

representative (Hoynes, Maestas and Strand 2022). An SAA looks over a case and if it is complete, they can allow it; otherwise, it goes on to a hearing with an ALJ.[4] An audit by SSA's Office of the Inspector General found that 9 percent of on-the-record allowances had medical evidence that was received while the claim was at the initial level of review but after the initial decision had been made (SSA:OIG 2016). Figure 6 shows that on-the-record allowances by SAAs increase after the start of the program in 2007 peaking at 45,000 cases per year in 2010. A decline in OTR decisions by SAAs started in 2010 and, simultaneously, the number issued by ALJs also began a steady decline at that time. The decline in use of OTR decisions was due to efforts to create consistency in appellate allowances; we discuss this next.

Third, shortly after the peak of use of on-the-record allowances, criticisms emerged about the use of SAAs and, importantly, appellate decision making more generally. In May 2011, the *Wall Street Journal* profiled an ALJ with a nearly 100 percent allowance rate on SSDI applications (Paletta 2011). Further, rather than scheduling hearings through the usual first-in, first-out system, the judge was reviewing a disproportionate number of applications from one particular attorney. A hearing by the House Ways and Means Subcommittee followed in July 2011 where members emphasized inconsistency of decision making, quality control issues, and the lack of accountability for judges (U.S. House of Representatives 2011). The *Wall Street Journal* article and subsequent Committee hearing brought extreme outliers in ALJ allowance rates to the public's attention. It also raised questions about possible criminal conduct.

Even before this publicity, however, SSA had begun capturing structured decision data

---

[4] ALJs can also make on-the-record determinations.

and applying data analytics to investigate the quality and consistency of appellate decisions (Ray and Lubbers 2015). The effort, which took place between 2009 and 2011, revealed types of decisions where adjudicators (e.g., ALJs, SAAs) made systematic errors by incorrectly applying agency policy. It also identified "outlier" judges who were awarding or denying benefits in a high percentage of cases. Further, some administrators believed that the use of SAAs and on-the-record allowances increased the allowance rate over what it would have been in their absence (Ray 2018).[5]

Three specific actions followed. First, SSA created and rolled out a new system of ALJ trainings that aimed to put decision making on a more policy compliant path. Incumbent ALJs and initial cohorts of ALJs that were hired off the new register took the new curriculum at summer meetings, most during 2011-2013 (we refer to this as "summer training"). All ALJs hired starting in the summer of 2010 and later received the new training as part of orientation training. We exploit this variation in the timing of the new training to estimate its (direct) impact on ALJ allowances.

Second, SSA conducted *focused reviews* of judges who were outliers in past decisions. The average number of cases reviewed per focused review was 99 (authors' calculations). The information was used to give feedback to judges or offices, or to refer the case to the Office of the Inspector General. Some administrators believed that focused reviews led to changes in behavior, early retirement of judges, and network effects on the peers of the judge that was

---

[5] Gerald Ray is former Assistant Deputy Chair of the Appeals Council, former Deputy Executive Director of the Office of Appellate Operations, and a former Administrative Appeals Judge.

reviewed (Ray & Lubbers 2015; Ray 2018). In our empirical analysis, we allow for both direct and network (any focused review in an office) effects of these focused reviews.

Third, SSA created software applications by which judges could compare their decision making to others. A prominent application called "How MI Doing?" came online in 2011. It included feedback mechanisms, training modules and hyperlinks to detailed agency policy information. Also, this application was used to identify types of claims where policy clarification could be effective (Ray & Lubbers 2015).

There is evidence that these initiatives narrowed the distribution of allowance rates across judges (GAO 2017). Our interest is in whether these initiatives additionally decreased the level of allowances.

Figure 4 shows that between 2008 and 2014 the hearing-level allowance rate fell by more than 20 percentage points, from above 70% to around 50%. This magnitude of decline was not observed at the initial level, where the allowance rate hovered in the range of 35-38% (U.S. Social Security Administration 2021).

We discuss how we quantify the reforms after we present the data and methods, below.

IV. Data

We employ the structured data systems that were developed by SSA and used in these policy initiatives. Several relational databases contain information about claimants, their impairment, and their characteristics. Further, they record the identity of the decision maker (e.g., ALJ, SAA) and their affiliated office. The Case Processing and Management System is our

primary database and we supplement this with information from the Management Information Electronic Disability Folder, the Appointed Representative Data Base, the Modernized Claim System, the 831 files, and the Summary Earnings Record. In addition, we utilize records of when specific judges received training and the dates that they were the subject of focused reviews.[6]

We include the universe of SSDI decisions made at the appellate level between 2007 and 2015, dropping only those with incomplete information.[7] There are a total of 3,433 decision makers—including both ALJs and SAAs—that made a total of 4,128,896 allowance decisions during this period.

We merge this appellate data with an extract from the SSA Master Earnings File, which has earnings for 1979 to 2019. This allows us to observe earnings for ten years prior to the application and for four years after the appellate decision for the full sample. We construct annual outcomes commonly used in the SSDI literature: employment (annual earnings > 0, annual earnings >$1,000), employment above substantial gainful activity (where the SGA threshold is that defined by SSA in a given year, e.g., $1,000/month in 2010), and annual earnings. We also construct variables for the number of employers per year (a proxy for number of work attempts), presence of self-employment income and amount of self-employment income. All earnings amounts are in 2014 dollars and include those with zero earnings.

---

[6] These records were obtained from the Office of Appellate Operations.

[7] We include SSDI applications that are concurrently considered for Supplemental Security Income (SSI) but not those that are for SSI alone.

V. Empirical Methods

In the first stage, we estimate the effects of the policy initiatives on the probability of allowance at the appellate level. We fit models of the following form:

(1) $$D_{ijdt} = \alpha + \gamma P_{jdt} + \pi X_{ijdt} + \eta_t + \theta_d + \varepsilon_{ijdt}$$

where $D_{ijdt} = 1$ if individual i's case heard by judge j in hearing office d in period (time of appellate determination) t is allowed (0 if denied or dismissed). We use the term "judge" but technically this refers to any adjudicator, ALJ or SAA. P_{jdt} is the vector of SSA policy reforms in place for judge j in hearing office d in period t. X_{ijdt} includes controls at the applicant, claim, and place level. This includes the unemployment rate measured at the state level (Maestas, Mullen & Strand 2021, Liu & Quimby 2023) at the time of the initial filing and the unemployment rate interacted with the diagnosis group. We include fixed effects for year (or month) of appellate determination (η_t) and hearing office (θ_d). The coefficients of interest are the γ's, which capture the causal effects of the various policy initiatives on the appellate-level allowance rate. We cluster the standard errors on judge.

In the second stage, using the first stage instruments, we estimate the effects of case allowance on employment and earnings (Y_{ijdt}) using two stage least squares:

(2) $$Y_{ijdt} = \alpha + \beta D_{ijdt} + \vartheta X_{ijdt} + \eta_t + \theta_d + \varepsilon_{ijdt}$$

We construct the instruments, P_{jdt}, leveraging the random assignment of judges to appellate cases conditional on assignment variables. Cases are not randomly assigned to SAAs. Further, during the first 9 months as a new ALJ, judges are not randomly assigned cases. Thus, our

assignment variables include indicator variables for the claim being assigned to an SAA or a trainee, along with the fixed effects for office and time (Table 1). The instruments then consist of the judge-level policy variables. This includes a set of judge hiring cohort indicators (incumbent judges present in 2007 are the omitted group).[8] We observe the training date of all judges and create a set of "post-training" dummies, separately for incumbent judges and those hired in 2008, 2009, and 2010. (Those hired in 2011 and later are hired with the new training.) We also include a dummy for the period following an own focused review (for the small share of judges who received a focused review). In some specifications, we also include office-month policy variables including the share of judges making 1 or more on the record decisions and the ratio of the number of SAAs to the number of ALJs, to capture the capacity to conduct on-the-record decisions. Finally, we include an indicator variable equal to 1 if any judge in the office received a focused review, to capture possible network effects.[9]

Control variables for individual characteristics include: gender, age categories (below 30, 30-39, 40-49, 50-54, 55-61, 62 plus), major class of body system for primary diagnosis, and whether there is an alleged secondary diagnosis. Claim characteristics include: concurrent for Supplemental Security Income (SSI), claim filed in a prototype state, claim flagged as noncitizen, jail, terminal illness, suicidal risk, or dire need, claim flagged at the initial level for expedited review or Quality Assurance review, claimant has a representative (fee or non-fee) before the claim is appealed, the denial step of the initial decision (step 2, 4, or 5), time from onset to filing and the allowance rate of the Disability Determination Service that made the

_____

[8] In our definition, the 2008 cohort starts in October 2007 to correspond to the start of hiring from the new register.
[9] We do not include a variable for "How MI Doing" as this was rolled out at the same time throughout all offices.

initial determination during the month of initial decision.

VI. Results

Descriptive Statistics

Table 2 illustrates changes in the workforce over the analysis period by showing descriptive statistics at the beginning (2007), midpoint (2011) and end (2015). The number of judges grew after the start of the new register (1,132 in 2007 to 1,564 in 2011) while the number of decisions per judge stayed relatively constant. As a result, the proportion of decisions made by incumbent judges—those hired before the start of the new register—declined from 100% in 2007 to 56% in 2011. Table 2 shows that the proportion of judges that received the policy compliant training (PCT) increased from 0% in 2007 to 41% in 2011 to 98% in 2015. Figure 7 further illustrates the changing composition of the adjudicators. The figure displays a histogram of all appellate cases between 2007 and 2015, identifying five groups: incumbent judge pre-PCT, incumbent judge post-PCT, new judge pre-PCT, new judge post-PCT, and SAA. This shows the evolution from incumbent judges to new judges, and the rollout of the PCT. This, along with the random assignment of judges to cases, is the basis of our policy instruments.

The trends in allowance rates are also visible in Table 2. Allowance rates declined across the distribution. For example, the median allowance rate fell from 65% in 2007, to 54% in 2011, to 45% in 2015 – a 20 percentage point decline. Similarly, the 25th (75th) percentile allowance fell by 18 (22) percentage points. Although the distribution narrowed (GAO 2017), the entire distribution shifted towards lower allowance rates. Figure 8 presents the mean allowance rates by judge cohort and month and shows two striking features. First, reductions in allowance rates are

present within virtually all cohorts of judges. Second, allowance rates decline nearly sequentially across successive judge cohorts from 2008 and 2013.

First Stage: Effects of SSDI Reforms on Allowance Rates

We show the results of the first stage regression in Table 3. The dependent variable equals one if the case is allowed at the appellate level, and otherwise zero. As noted above, the model also includes hearing office fixed effects, time fixed effects, state unemployment rates (allowed to vary by type of impairment), and a comprehensive set of applicant characteristics. The first specification includes fixed effects by year and the second specification includes fixed effects for each year-month. The third specification adds the office environment variables.

Our base specification is column 2, with year-month fixed effects. The first stage shows reductions in allowance rates across the judge cohort (as also seen in Figure 8). The first new cohort after the ALJ hiring resumed, those hired in 2008, had allowance rates that were 3.4 percentage points lower than incumbents (the omitted group). This is relative to the mean allowance rate of 57.7 percent for the full sample period. Similarly, those hired in 2009 (2010) had allowance rates that were 3.6 (4.9) percentage points lower than incumbents. Beginning in 2011, judges were hired with the new PCT and the estimated judge cohort fixed effects show allowance rates 7.3, 7.7, 10.2, 12.9 and 13.9 percentage points lower than incumbents (for 2011-2015 cohorts). This is a striking decline in allowance rates in a very short period of time.

The first stage also shows the effects of the new training for those hired prior to 2011. The training reduced allowance rates by 2.1 percentage points for incumbents, 1.3 percentage points for the 2009 cohort, and 1.8 percentage points for the 2010 cohort (with insignificant

16

effects for the 2008 cohort). A judge receiving a focused review experienced a 4.5 percentage point reduction in allowance rates. These are significant impacts of training, though smaller than the overall effects across successive judge hiring cohorts.

The third column shows the impacts of adding the three office environment variables. These variables vary continuously over offices and months and reflect characteristics of the office as a whole.[10] An office that has a higher portion of judges that perform on-the-record (OTR) decisions (in a given month) has a higher allowance rate. This confirms the belief of some administrators that the use of OTR increased allowances (Ray 2018). Interestingly, holding OTR decisions constant, having more SAAs per judge in an office leads to lower allowance rates. Finally, we find no evidence of network or spillover effects of focused reviews; having a judge in the office receive a focused review has a very small and statistically insignificant effect on allowance rates.

In Figure 9, we illustrate the importance of the policy variables relative to the total reduction in allowance rates, by adding them one at a time to the model (each is cumulative). We use the specification in column 2 of Table 3, with year-month dummies. In each row of the figure, we present the regression-estimated change in the allowance rate, on average, between 2007 and 2015. The first row shows that the allowance rate declined by 22 percentage points between 2007 and 2015; this is the unconditional change without any controls (consistent with Figure 3). Adding office fixed effects has little effect on the decline (reducing it to 21 percentage

---

[10] We measure the office variables using a "leave out" approach, dropping the current observation in constructing the office-month variable.

points). Interestingly, adding the state unemployment rate at the time of the original filing (and interactions of unemployment and diagnosis) increases the reduction to 23 percentage points. This result is compatible with Maestas, Mullen & Strand, 2021.[11] Next, adding the large set of personal and case characteristics has almost no impact on the change in allowance rates, suggesting the composition of cases that come to the appellate level are not part of the story. In the 5th row we add the remaining assignment variables, the indicator variables for SAA and the ALJ trainee period. Overall, adding these controls and assignment variables has little impact on the estimated change in the allowance rate – the decline goes from 22 percentage points in the unconditional model and 21.2 percentage points in the model with the full set of controls.

The policy instruments, in contrast, explain a significant amount of the time series variation in allowance rates. Adding the judge cohort variables explains 4 percentage points of the decline and adding the PCT dummies explains another 2 percentage points. Adding the office environment explains another 2 percentage points. Overall, we find that the turnover of judges, training of judges, and other policy changes explain 6.2 to 8.0 percentage points of the 22 percentage point decline in allowance rates or 28-36% of the decline. On the other hand, we find little role of the economic cycle or changes in caseload.

Validity of IV and Possible Threats to Exogeneity

_____

[11] According to estimates by Maestas, Mullen & Strand (2021), the Great Recession strongly affected the volume of SSDI application and awards. However, there are two reasons why the effect would be muted in the comparison in Figure 9. First, the applications induced by the Great Recession were disproportionately denied at the initial level of review and less than half of those proceeded to the hearing level (see their Table 4). Second, the effect of the Great Recession had largely dissipated by 2013 (see their Figure A3). As a result, the endpoints of our analysis period, 2007 and 2015, were largely outside the period affected by the Great Recession.

Before presenting the instrumental variables estimates, we provide evidence in support of the conditions necessary to interpret them as local average treatment effects (LATEs). First, the instruments must be relevant; the first stage F statistics range from 15.2 to 16.9 (Table 3). Second, we need independence of instruments; we explore that here with a series of balance tests. Third, we investigate the exclusion restriction by presenting evidence on the effect of instruments on claims.

To investigate the independence of the instruments, we identify a set of predetermined and "hold-out" characteristics of each case, annual earnings averaged across the 6 to 10 year period prior to filing for DI, completed years of education of claimant, and day of the week and day of the month of initial filing. We regress each of these characteristics on our policy instruments and controls, as in equation 1. We present the estimates of this balance test in Figure 10. There are four panels in the figure, one for each of the balance variables. Each of the panels plots the coefficient and 95 percent confidence interval for each of the 13 policy instruments (8 judge cohort dummies, 4 post-PCT dummies, one post-focused review). To put them on a similar scale we plot the effect of a 1 SD change in the instrument on the outcomes. Of the 52 coefficients (13 instruments x 4 balance variables) we have 5 statistically significant at the 5 percent level. Overall, the figure shows very small magnitudes – the instruments are not associated with changes in the composition or characteristics of claims. We take this as confirming of a key assumption of the IV approach – that the instruments are not correlated with the unobservable determinants of case outcomes.

To investigate the exclusion restriction, we estimate impacts of the policy instruments on

new claims for SSDI. Here we use SSA administrative data on new applications for SSDI (Hoynes, Maestas and Strand, 2022) covering 1989 to 2021 and collapse to the office by month level. We then regress the log of applications on our policy instruments, including fixed effects for office and year-month, along with office-month means of each of the covariates in equation 1. We weight the regression by the office-area population and cluster on office.[12] The results are presented in Table 4. Each column is a different regression, tracing out the leads and lags of the policy instruments on the log of claims. Overall, the instruments have little impact on SSDI applications. Figure 11 illustrates this by presenting the coefficient and 95 percent confidence interval of a 1 standard deviation change in each of the instruments. Some are positive, some are negative, and few are statistically significant. Additionally, the point estimates are not of economic significance. In sum, the data on applications provides supportive evidence that our instruments are not affecting outcomes through the pathway of affecting the quantity of applications.

IV Estimates on Labor Market Outcomes

Table 5 provides instrumental variables estimates of the effect of appellate allowance rates on claimant labor market outcomes measured four years after the appellate decision. Receiving an allowance leads to a $4,140 decline in annual earnings and a 22.1 percentage point reduction in employment. These are large effects relative to the low mean earnings and employment of all applicants four years after the appellate decision (mean annual earnings $2,871, mean annual employment 20.6 percent).

---

[12] We drop the small number of observations with applications in that cell.

How do our LATEs from these policy reforms compare to prior work? The policy reforms we study lead to a reduction in DI program generosity, akin to moving the program's eligibility threshold to a lower point in the distribution of residual work capacity. This implies that over this period of gradual program tightening the marginal claimant would have had *lower* work capacity and therefore a smaller SSDI treatment effect than the marginal claimant denied before the policy reforms. Consistent with this, our estimates are about 12-15 percent smaller than French and Song (2014) who use random assignment of judges and SSA data on claims assigned to ALJs in 1990-1999 to estimate causal effects of SSDI allowance on labor supply three years after the ALJ assignment.

We present other measures of labor market impacts in Table 5. We find that an allowance leads to a 14 percentage point decline in annual earnings being above SSA's measure of substantial gainful activity (SGA). This accounts for nearly two-thirds of the 22 percentage point reduction in employment and implies there is some work capacity among those on the margin of the policy reforms, including for some, work capacity above SGA. Consistent with this, we find that an allowance leads to having a smaller number of employers per year, which may reflect a reduction in the number of work attempts. This effect is also large relative to the mean among all applicants. We find a much smaller effect on self-employment income, where an allowance leads to a 1.7 percentage point decline in the presence of self-employment income.

Since the policy reforms led to a reduction in allowances (and therefore an increase in denials), multiplying our estimates by -1 gives the impact of the policy reform. Annual earnings of denied applicants on the eligibility margin more than doubled relative to mean earnings for the

group as a whole, as did the employment rate (from 0.206 to 0.427 [0.206+0.221]), and nearly two-thirds of the increase in employment was associated with earnings above SGA. Nonetheless, the majority of applicants did not return to work (1-0.427), and even among those who did, their average earnings were relatively low, possibly indicating part-time work.

Table 6 shows the robustness of the IV impacts on earnings to the alternative specifications presented earlier in Table 3. Including year or year-month time dummies (column 1, 2) or adding the office environment variables (column 2) has little impact on the IV estimates.

VII. Conclusion

We document recent reforms to the appellate process in the SSDI program including the opening of the new ALJ register, the corresponding turnover in ALJs, and other training and policy initiatives. These are possibly the most important (and least heralded) SSDI policy reforms of our time. We show that turnover among judges and the implementation of policy compliant training led to a significant decline in the allowance rate, explaining 28-36 percent of the 22 percentage point decline in the allowance rate between 2007 and 2015. Using the policy reforms as instruments, we estimate causal effects of allowance on labor supply outcomes with estimates about 12-15 percent lower than French and Song (2014). In ongoing work, we examine what these estimates say about the nature of the reforms (e.g., did they improve targeting?), the role of program generosity, and to what extent they contribute to the recent increases in employment among persons with disabilities.

References

Abraham, Katharine G., and Melissa S. Kearney. 2020. "Explaining the Decline in the US Employment-to-Population Ratio: A Review of the Evidence." *Journal of Economic Literature*, 58 (3): 585-643.

Autor, David H., and Mark G. Duggan. 2003. "The Rise in the Disability Rolls and the Decline in Unemployment." *The Quarterly Journal of Economics*, 118(1): 157-206, https://doi.org/10.1162/00335530360535171

Black, Dan, Kermit Daniel and Seth Sanders. 2002. "The Impact of Economic Conditions on Disability Program Participation: Evidence from Coal Boom and Bust" *American Economic Review* 92(1): 27-50.

Black, Sandra E., Diane Whitmore Schanzenbach, and Audrey Breitwieser. 2017. "The Recent Decline in Women's Labor Force Participation." in The 51%: Driving Growth Through Women's Economic Participation, by Diane Whitmore Schanzenbach and Ryan Nunn, Brookings Institution Press: Washington, DC.

Bound, John. 1989. "The Health and Earnings of Rejected Disability Insurance Applicants," *American Economic Review*, 79(3): 482-503. http://www.jstor.org/stable/1806858

John Bound, and Timothy Waidmann. 1992. "Disability Transfers, Self-Reported Health, and the Labor Force Attachment of Older Men: Evidence from the Historical Record." *The Quarterly Journal of Economics*, 107(4): 1393–1419. https://doi.org/10.2307/2118393

Charles, Kerwin, Yiming Li and Melvin Stephens, Jr. 2018. "Disability Benefit Take-Up and Local Labor Market Conditions," *The Review of Economics and Statistics* 100(3): 416-423.

Burkhauser, Richard V., and Mary Daly. 2011. *The Declining Work and Welfare of People with Disabilities: What Went Wrong and a Strategy for Change.* American Enterprise Institute Press.

Degroot, Nynke de, and Pierre Koning. 2016. "Assessing the Effects of Disability Insurance Experience Rating. The Case of The Netherlands." *Labour Economics,* 41: 304-317. https://doi.org/10.1016/j.labeco.2016.05.007

Deshpande, Manasi, Tal Gross, and Yalun Su. 2021. "Disability and Distress: The Effect of Disability Programs on Financial Outcomes." *American Economic Journal: Applied*

Economics, 13(2): 151-78. https://doi.org 10.1257/app.20190709

Deshpande, Manasi, and Lee M. Lockwood. 2022. "Beyond Health: Nonhealth Risk and the
Value of Disability Insurance." *Econometrica,* 90(4): 1781-1810

Deshpande, Manasi and Michael Mueller-Smith. 2022. "Does Welfare Prevent Crime? the
Criminal Justice Outcomes of Youth Removed from SSI." *The Quarterly Journal of
Economics*, 137(4): 2263–2307. https://doi.org/10.1093/qje/qjac017.

French, Eric, and Jae Song. 2014. "The Effect of Disability Insurance Receipt on Labor
Supply." *American Economic Journal: Economic Policy*, 6(2): 291-337.
https://doi.org/10.1257/pol.6.2.291

Gelber, Alexander, Timothy J. Moore, and Alexander Strand. 2017. "The Effect of Disability
Insurance Payments on Beneficiaries' Earnings." *American Economic Journal: Economic
Policy*, 9 (3): 229-61. https://doi.org/10.1257/pol.20160014

Haveman, Robert H., and Barbara L. Wolfe. 1984a. " Disability transfers and early retirement: a
causal relationship?" *Journal of Public Economics,* 24(1): 47-66. https://doi.org/10.1016/0047-
2727(84)90004-5

Haveman, Robert H., and Barbara L. Wolfe. 1984b. "The Decline in Male Labor Force
Participation: Comment." *Journal of Political Economy,* 92(3): 532-541.
https://doi.org/10.1086/261241

Hoynes, Hilary, Nicole Maestas and Alexander Strand, 2022. "Legal Representation in Disability
Claims." NBER Working paper 29871, March 2022.

Koning, Pierre, and Maarten Lindeboom. 2015. "The Rise and Fall of Disability Insurance
Enrollment in the Netherlands." *Journal of Economic Perspectives*, 29(2): 151-72.
https://doi.org/ 10.1257/jep.29.2.151

Liebman, Jeffrey B. 2015. "Understanding the Increase in Disability Insurance Benefit Receipt in
the United States." *Journal of Economic Perspectives*, 29 (2): 123-50.
https://doi.org/10.1257/jep.29.2.123

Liu, Siyan and Laura D. Quinby. 2023. "What Factors Explain the Drop in Disability Insurance
Rolls from 2015 to 2019?" Boston College Center for Retirement Research, Working Paper
2023-7.

Low, Hamish, and Luigi Pistaferri. 2015. "Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off." *American Economic Review*, 105(10): 2986-3029**.**

Parsons, Donald O. 1980. "The Decline in Male Labor Force Participation." *Journal of Political Economy* 88(1): 117–34. http://www.jstor.org/stable/1830962.

Parsons, Donald O. 1982. "The Male Labour Force Participation Decision: Health, Reported Health, and Economic Incentives." *Economica,* 49(193): 81–91. https://doi.org/10.2307/2553527.

National Academies of Sciences, Engineering, and Medicine. 2022. Understanding the Aging Workforce: Defining a Research Agenda. Washington, DC: The National Academies Press. https://doi.org/10.17226/26173.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand. 2013. "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt" *American Economic Review*, v. 103, iss. 5, pp. 1797-1829.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand. 2021. "The Effect of Economic Conditions on the Disability Insurance Program: Evidence from the Great Recession." *Journal of Public Economics*, Volume 199, July 2021, 104410.

Paletta, Damian. 2011. "Disability-Claim Judge Has Trouble Saying 'No': Near-Perfect Approval Record; Social-Security Program Strained." *The Wall Street Journal*, May 11, 2011.

Ray, Gerald K. and Jeffrey S. Lubbers. 2015. "A Government Success Story: How Data Analysis by the Social Security Appeals Council (with a Push from the Administrative Conference of the United States) Is Transforming Social Security Disability Adjudication" *The George Washington Law Review*, Vol 83, 1575-1608.

Ray, Gerald K. 2018. Conversations with the Authors.

U. S. Government Accountability Office. 2010. "Results-Oriented cultures: Office of Personnel Management Should Review Administrative Law Judge Program to Improve Hiring and Performance Management." Washington, DC: U.S. Government Accountability Office.

U. S. Government Accountability Office. 2017. "Social Security Disability: Additional Measures and Evaluation Needed to Enhance Accuracy and Consistency of Hearings Decisions." Washington, DC: U.S. Government Accountability Office.

U.S. House of Representatives, Committee on Ways and Means. 2011. "Role of Social Security Administrative Law Judges." Serial no. 112-67, Serial No. 112-SS8. Washington, DC: U.S. Government Printing Office. (here)

U. S. Social Security Administration, Office of the Inspector General. 2016. "On-the-Record Favorable Decisions Processed at Hearing Offices Within 100 Days of Receipt." Baltimore, Maryland: U. S. Social Security Administration.

U. S. Social Security Administration. 2007, 2009, 2010, 2013 & 2016. "Annual Statistical Supplement to the Social Security Bulletin." Baltimore, Maryland: U. S. Social Security Administration.

U. S. Social Security Administration. 2021. "Annual Statistical Report on the Social Security Disability Insurance Program, 2020." Baltimore, Maryland: U. S. Social Security Administration.

Figure 1. Employment Rate Relative to 1988 of People with and without a Work Disability (1988=1)
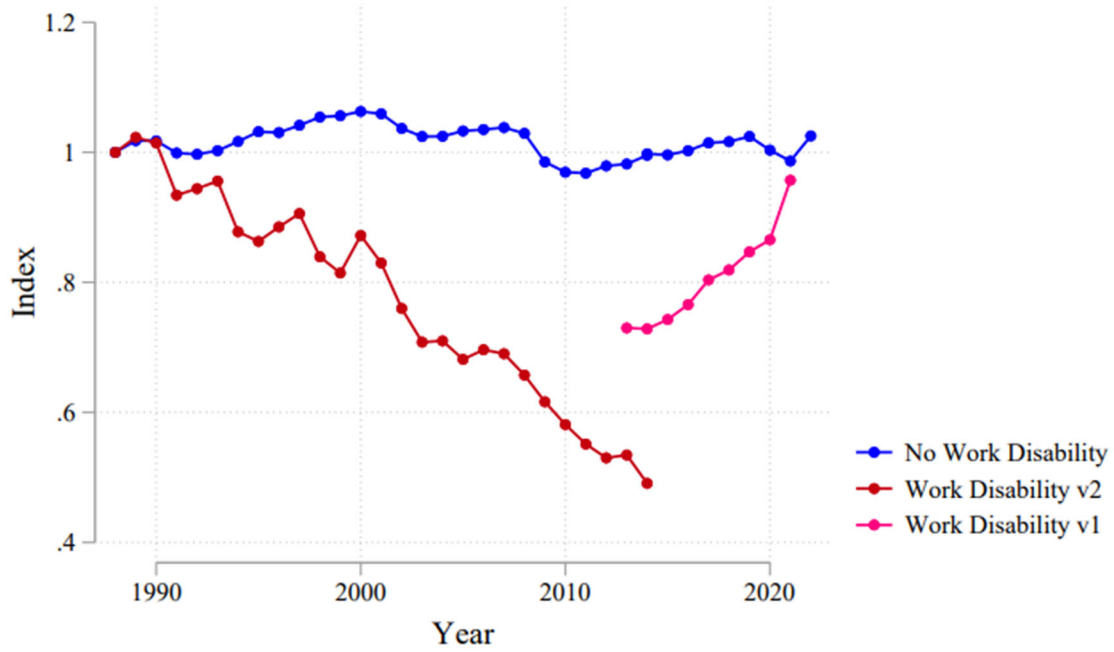
Notes: Authors' tabulations of the IPUMS-CPS, ASEC data. Respondents are considered to have a work disability if they have a "Health problem or disability which prevents from working or which limits the kind or amount of work." Break in series occurs because the question wording changed in 2014/2015 to include temporary disabilities ("even for a short time") and to switch the reference period from "currently" to "any time in last year."

Figure 2. Number of SSDI Disabled Worker Beneficiaries 1960-2020, in millions
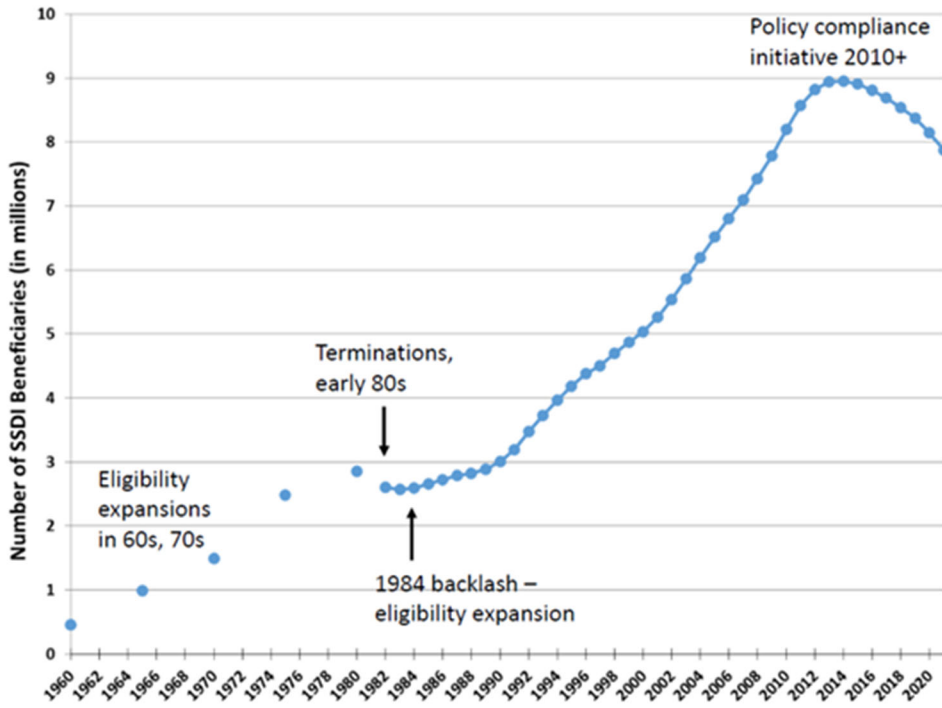
Source: SSA, Annual Statistical Report on the Social Security Disability Insurance Program, 2021, Table 1

Figure 3. Number of SSDI Disabled Worker Beneficiaries, Awards and Exits, 2001-2021
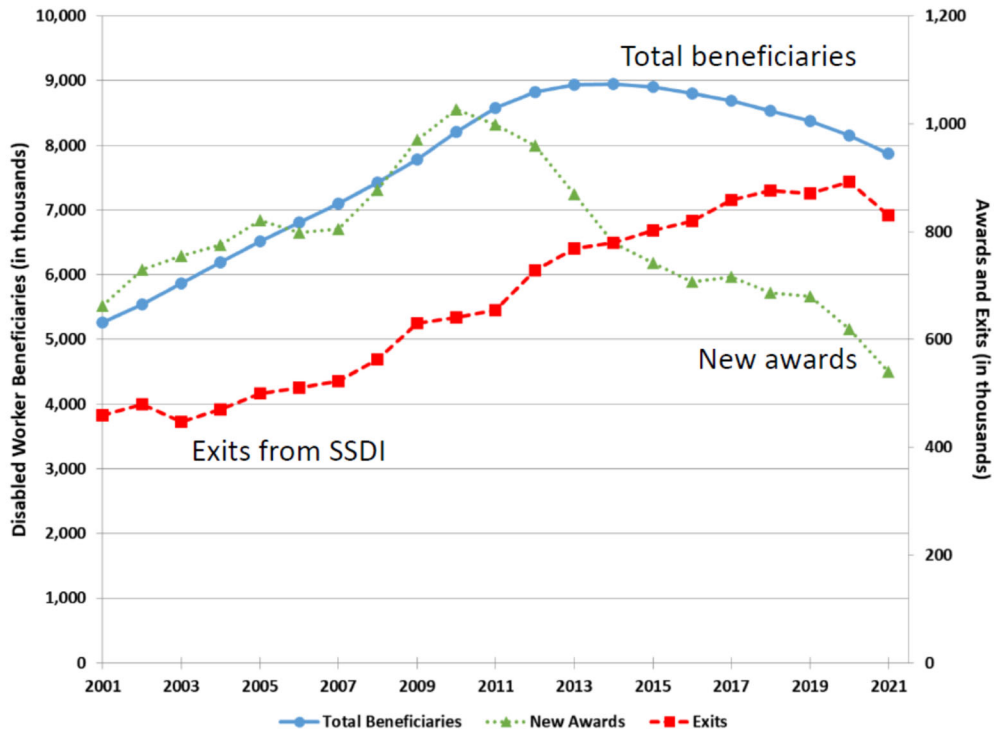
Source: SSA, Annual Statistical Report on the Social Security Disability Insurance Program, 2021, Tables 1, 35, and 49

Figure 4. SSDI Allowance Rates, Appellate Level
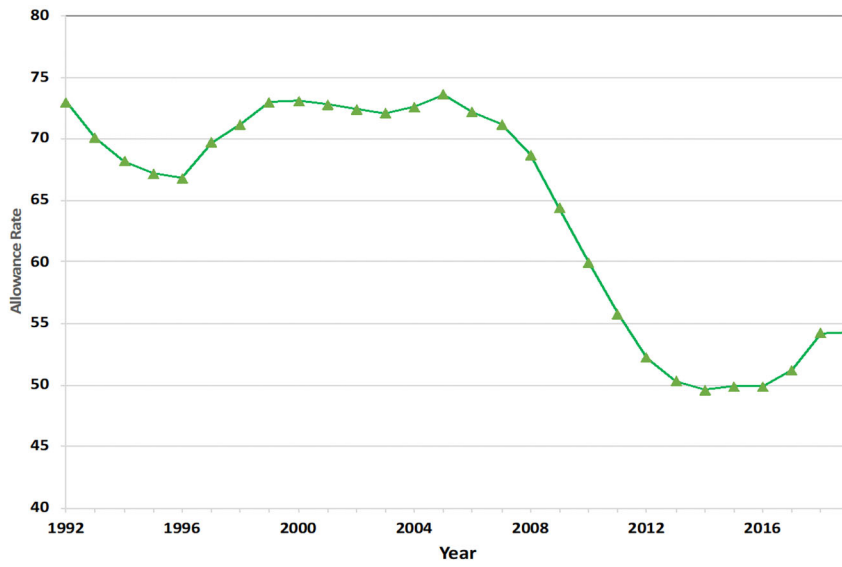
Source: Annual Statistical Report on the Social Security Disability Insurance Program, 2020. Tables 60 and 63.

Figure 5. Backlog of Disability Applications, by year

Appellate Case backlog
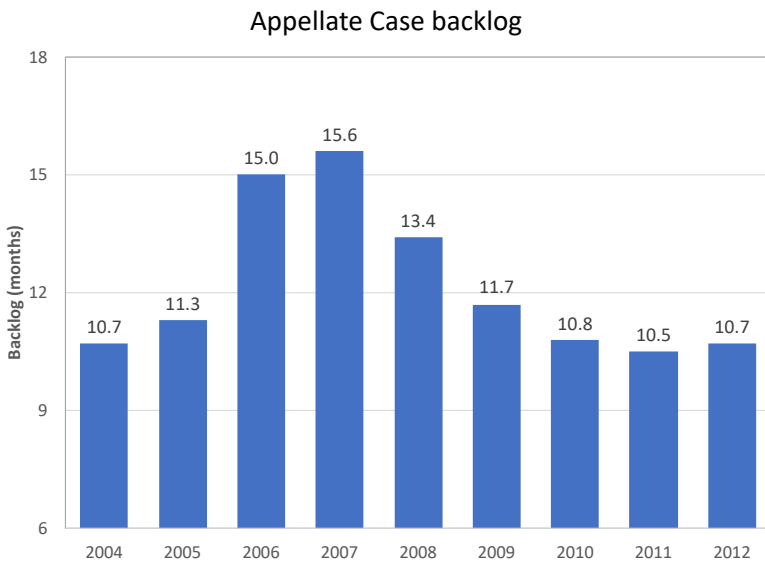
Source: Annual Statistical Supplement to the Social Security Bulletin, various years.

Note: the backlog is defined as the number of pending hearings divided by the average hearings per month. The resulting measure is the number of months it would take to eliminate the backlog if counterfactually there were no further applications.

Figure 6. Number of On-The-Record Decisions, 2007-2015
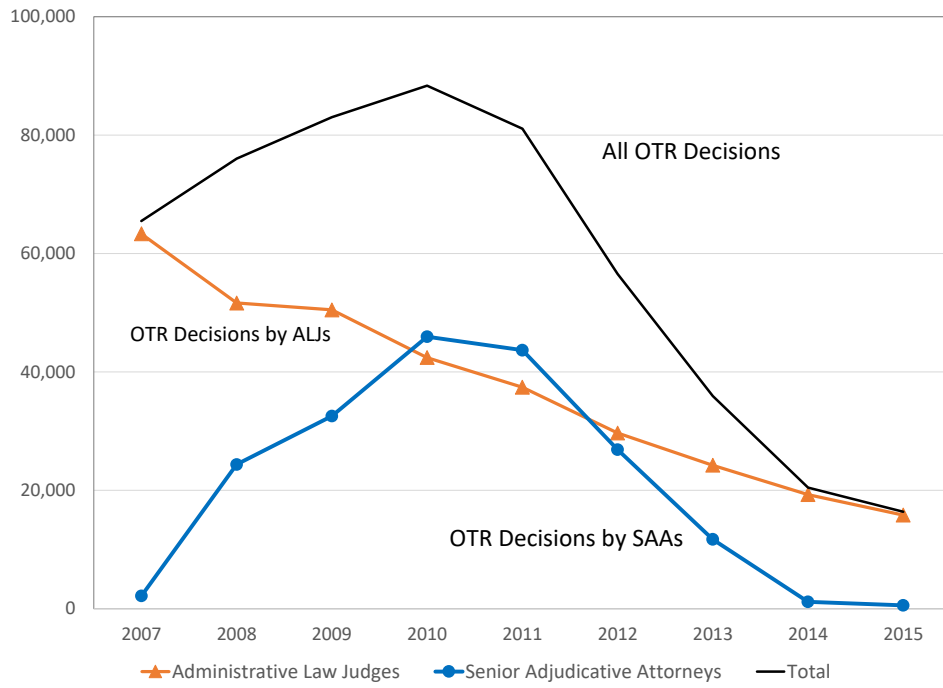
Source: Authors' calculation using the Case Processing & Management System.

Figure 7. Number of Appellate Decisions by Type of Decision Maker and Policy Compliant Training
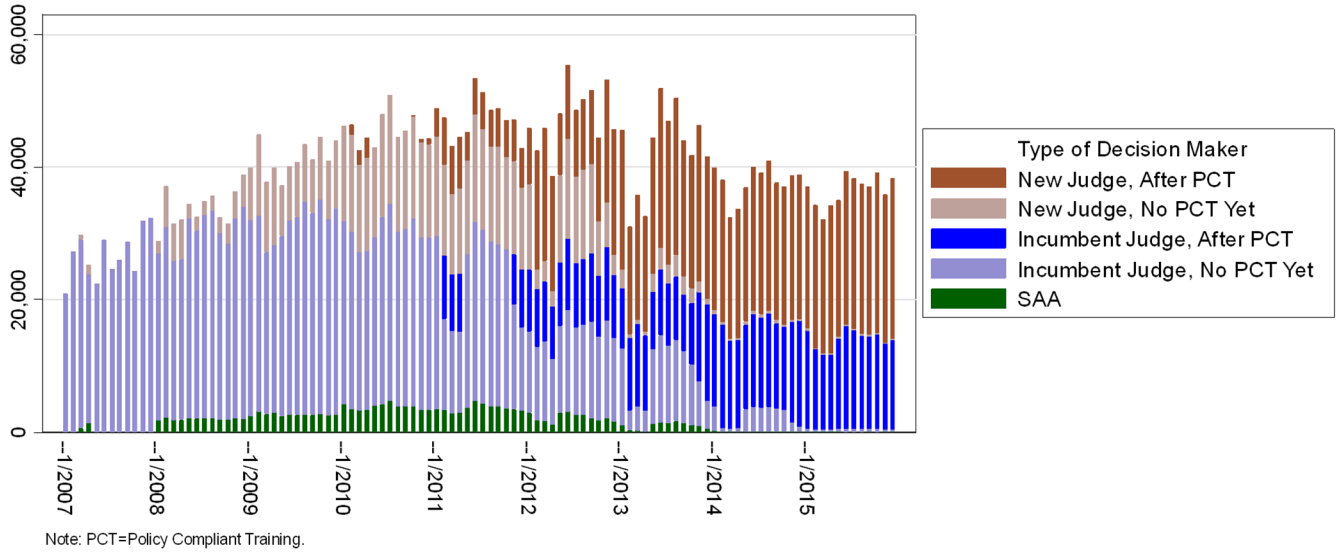
Note: PCT=Policy Compliant Training.

Notes: Based on authors' tabulations of appellate cases recorded in Social Security Administration databases for years 2007-2015 (data sources described in Section IV). Incumbent judges are those in place before the ALJ register reopened in October 2007.

Figure 8. Allowance Rate by Judge Cohort
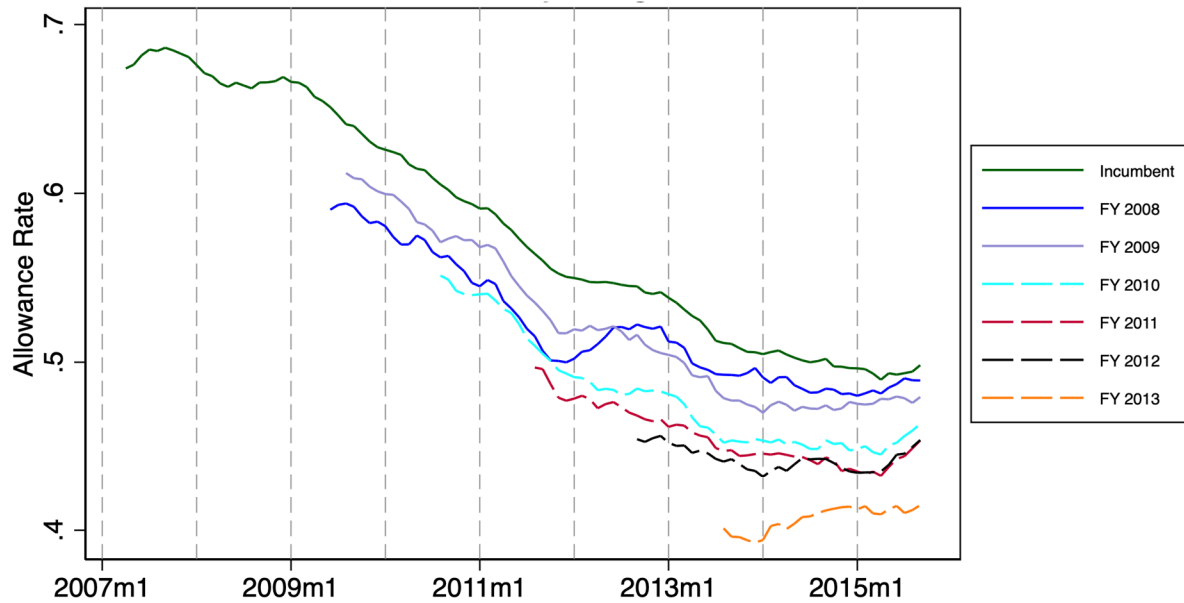
Note: trainees and month-cohort cells with <800 determinations are excluded.
Allowance rates are smoothed with a seven-month moving average.
Dashed lines indicate cohorts hired with policy-compliant training.

Notes: Based on authors' tabulations of appellate cases recorded in Social Security Administration databases for years 2007-2015 (data sources described in Section IV). Incumbent judges are those in place before the ALJ register reopened in October 2007.

Figure 9. Estimated Difference in Allowance Rate in 2015 Relative to 2007, Effects of Instruments and Controls

Difference in 2015 Average Allowance Rate Relative to 2007 Average With and Without Controls (percentage points)

Label	Value
Year-Month Dummies Only	-22.0
Plus Office Dummies	-21.1
Plus UR and UR * Impairment Class	-23.1
Plus Personal and Case Char.	-22.7
Plus Other Cond. Random Assign. Vars.	-21.2
Plus Judge Cohort	-17.6
Plus Post PCT & Post FR	-15.8
Plus Office Environment	-14.0
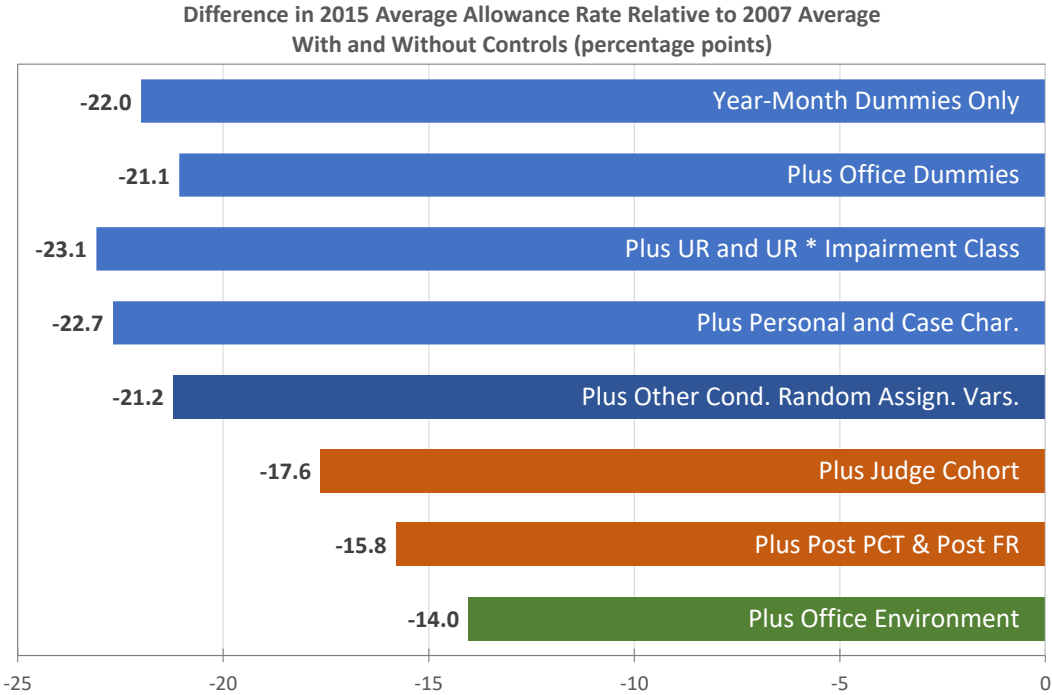
Figure 10. Balance Test: Regressing Holdout Variables on Instrumental Variables
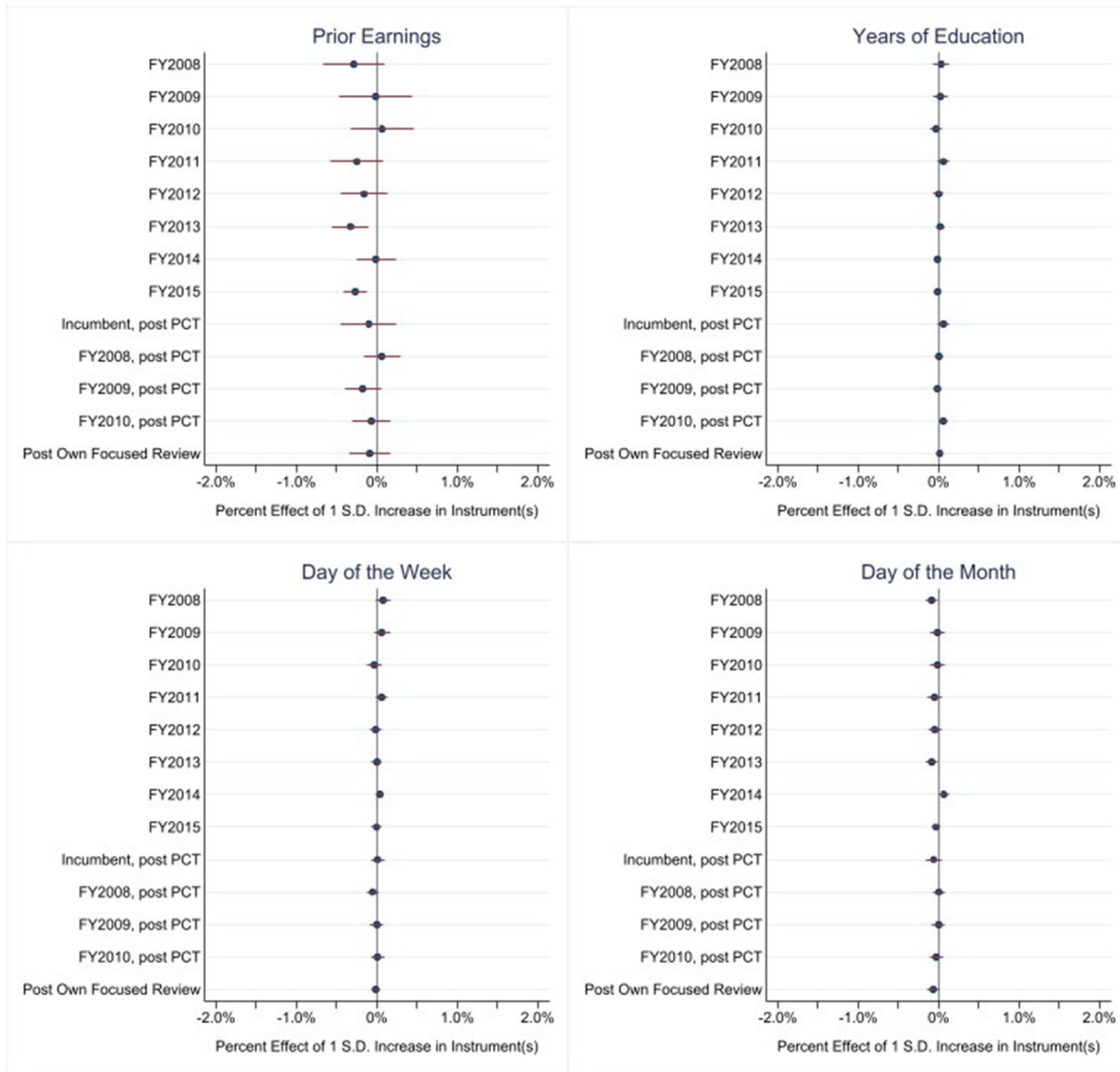
Notes: Estimates of equation (1) where outcome variable is a holdout balance test variable. All models include fixed effects for year-month and DDS office as well as individual and claim characteristics. Here we simulate a 1 SD increase in each instrument and its 95 percent confidence interval.

Figure 11. Validity Test: Regressing Log Number of Claims on Instrumental Variables
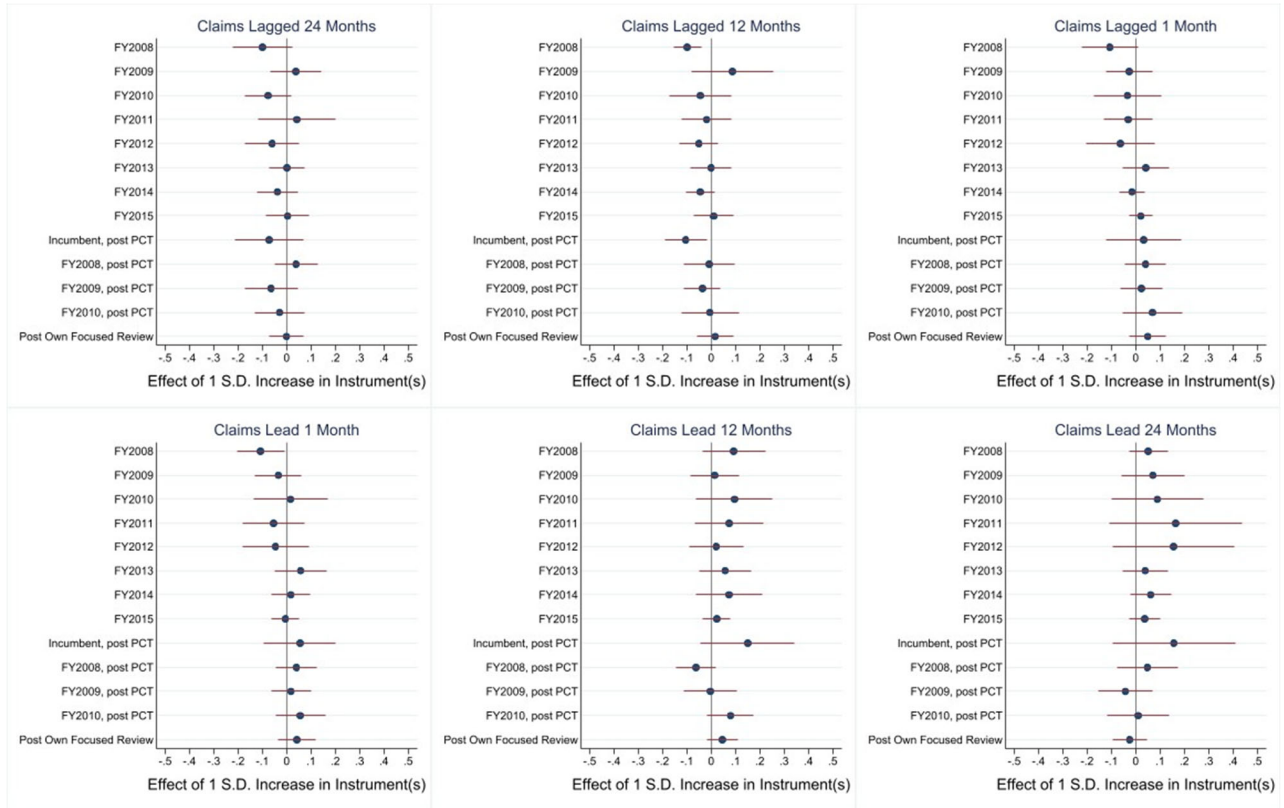
Notes: Applications to SSDI is collapsed to DDS by month cells. 0.86 percent of cells are lost due to zero observations. Covariates include fixed effects for year-month and DDS office. Regressions are weighted by DDS-area population. Here we simulate a 1 SD increase in each instrument and its 95 percent confidence interval.

Table 1. Instruments, Assignment Variables, and Controls

Instruments	Assignment Variables	Controls
Judge Cohort (omitted = incumbent)	SAA (=1 if SAA)	Female
FY2008	Trainee (=1 if ALJ in initial 9-mo	Age category dummies
FY2009	trainee period)	Body system, major class
FY2010	Office fixed effects	Case characteristics (41 vars)
FY2011	Time fixed effects	Labor market controls
FY2012		State-month UR
FY2013		State-month UR * body system class
FY2014		
FY2015		
Post Policy Compliant Training		Office-month environment
(FY2011+ hired with PCT)		Share with OTR decisions
Incumbent, post PCT		Number SAAs / Number ALJs
FY2008, post PCT		Post focused review in office
FY2009, post PCT		
FY2010, post PCT		
Post Own Focused Review		

Table 2. Characteristics of Judges and Senior Adjudicative Attorneys

	2007	2011	2015
Judges			
Number	1,132	1,564	1,633
Number of Disability Insurance Decisions, Mean	282	308	268
Allowance Rate			
1st Percentile	0.16	0.18	0.00
5th Percentile	0.36	0.30	0.22
10th Percentile	0.43	0.36	0.28
25th Percentile	0.54	0.44	0.36
Median	0.65	0.54	0.45
75th Percentile	0.76	0.65	0.54
90th Percentile	0.84	0.74	0.63
95th Percentile	0.89	0.79	0.68
99th Percentile	0.96	0.89	0.80
Proportion that make On-The-Record Decisions	0.97	0.95	0.90
Proportion Incumbent	1.00	0.56	0.36
Policy-Compliant Training, Proportion			
Summer Training	0.00	0.21	0.64
New-Hire Training	0.00	0.20	0.44
Total	0.00	0.41	0.98
Had a Focused Review	0.00	0.01	0.03
Office Had a Focused Review	0.00	0.19	0.50
Senior Adjudicative Attorneys			
Number	299	751	98
Number of Disability Insurance Decisions, Mean	7	58	6
Allowance Rate, Mean	1	1	1

Notes: Based on authors' tabulations of appellate cases recorded in Social Security Administration databases for years 2007-2015 (data sources described in Section IV). Counts reflect only decisions made on Title II (Disability Insurance) disabled worker claims. Incumbent judges are those in place before the ALJ register reopened in October 2007.

Table 3. First Stage: Effects of Policy Reforms on Appellate Allowances

	(1)		(2)		(3)	
Judge Cohort (omitted = incumbent)						
FY2008	-0.035**	(0.014)	-0.034**	(0.014)	-0.034**	(0.014)
FY2009	-0.037***	(0.010)	-0.036***	(0.010)	-0.036***	(0.010)
FY2010	-0.051***	(0.008)	-0.049***	(0.008)	-0.049***	(0.008)
FY2011	-0.077***	(0.008)	-0.073***	(0.009)	-0.073***	(0.009)
FY2012	-0.082***	(0.010)	-0.077***	(0.010)	-0.077***	(0.010)
FY2013	-0.108***	(0.011)	-0.102***	(0.012)	-0.102***	(0.012)
FY2014	-0.136***	(0.025)	-0.129***	(0.025)	-0.127***	(0.025)
FY2015	-0.144***	(0.012)	-0.139***	(0.012)	-0.136***	(0.012)
Post Policy Compliant Training *(FY2011-2015 hired with PCT)*						
Incumbent, post PCT	-0.026***	(0.006)	-0.021***	(0.006)	-0.021***	(0.006)
FY2008, post PCT	0.0003	(0.011)	0.004	(0.011)	0.004	(0.011)
FY2009, post PCT	-0.016**	(0.008)	-0.013*	(0.008)	-0.013*	(0.008)
FY2010, post PCT	-0.021***	(0.007)	-0.018**	(0.007)	-0.018***	(0.007)
Post Own Focused Review	-0.047**	(0.020)	-0.045**	(0.020)	-0.045**	(0.020)
Office-Month Environment						
Share with OTR decisions					0.059***	(0.005)
Number SAAs / Number ALJs					-0.047***	(0.009)
Post focused review in office					0.001	(0.005)
Time fixed effects	year		yr-month		yr-month	
Office-Month Environment Variables	No		No		Yes	
Number of observations	4,128,896		4,128,896		4,128,896	
R Squared	0.1615		0.1619		0.1622	
F Statistic	16.85		15.36		15.21	

Source: Case Processing and Management System and other administrative data sources.

Notes: *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Table 4. Validity Test: Regressing Leads and Lags of Log Number of Claims on Instruments

	Models with Lags						Models with Leads					
	5 years	4 years	3 years	2 years	1 year	1 month	1 month	1 year	2 years	3 years	4 years	5 years
Judge Cohort (omitted = incumbent)												
FY2008	-0.125	-0.191	-0.501**	-0.528	-0.521***	-0.558*	-0.563**	0.491	0.273	-0.279	0.660***	0.263*
	(0.395)	(0.387)	(0.191)	(0.335)	(0.152)	(0.313)	(0.261)	(0.349)	(0.214)	(0.204)	(0.185)	(0.155)
FY2009	0.575*	0.271	0.289	0.149	0.337	-0.105	-0.131	0.0552	0.272	-0.111	-0.324	-0.482
	(0.300)	(0.208)	(0.265)	(0.206)	(0.333)	(0.188)	(0.188)	(0.193)	(0.252)	(0.216)	(0.291)	(0.389)
FY2010	0.141	-0.129	-0.542	-0.276	-0.161	-0.122	0.0587	0.347	0.322	-0.145	-0.298	-0.0569
	(0.256)	(0.318)	(0.394)	(0.173)	(0.232)	(0.258)	(0.277)	(0.290)	(0.348)	(0.242)	(0.273)	(0.408)
FY2011	0.0336	-0.038	0.0774	0.162	-0.0715	-0.119	-0.208	0.287	0.633	-0.0452	-0.467	-0.525
	(0.333)	(0.419)	(0.187)	(0.312)	(0.198)	(0.201)	(0.252)	(0.280)	(0.540)	(0.193)	(0.462)	(0.568)
FY2012	-0.402	-1.021*	-0.776*	-0.282	-0.238	-0.296	-0.217	0.0985	0.728	0.156	-0.344	-0.695
	(0.340)	(0.536)	(0.448)	(0.264)	(0.187)	(0.338)	(0.325)	(0.262)	(0.601)	(0.283)	(0.331)	(0.518)
FY2013	-0.0305	0.0182	0.0591	0.00974	-0.00163	0.248	0.346	0.347	0.236	-0.431	-0.453	-0.362
	(0.349)	(0.249)	(0.404)	(0.227)	(0.255)	(0.289)	(0.331)	(0.323)	(0.290)	(0.458)	(0.334)	(0.330)
FY2014	-1.448	-0.683	-0.135	-0.627	-0.751	-0.25	0.284	1.219	1.024	-0.609	-0.799	-0.872
	(1.166)	(0.532)	(0.551)	(0.715)	(0.496)	(0.455)	(0.660)	(1.165)	(0.721)	(0.770)	(0.585)	(0.820)
FY2015	0.498	0.184	-0.186	0.0443	0.126	0.238	-0.0608	0.256	0.413	-0.188	-0.132	-0.141
	(0.522)	(0.369)	(0.295)	(0.513)	(0.474)	(0.267)	(0.328)	(0.331)	(0.369)	(0.271)	(0.390)	(0.361)
Post Policy Compliant Training (FY2011-2015 hired with PCT)												
Incumbent, post PCT	0.019	-0.102	0.0225	-0.204	-0.295**	0.0922	0.156	0.422	0.442	-0.169	-0.401	-0.218
	(0.297)	(0.329)	(0.115)	(0.205)	(0.124)	(0.221)	(0.212)	(0.277)	(0.363)	(0.173)	(0.290)	(0.259)
FY2008, post PCT	-0.217	-0.676	0.0139	0.339	-0.0751	0.356	0.347	-0.544	0.424	-0.0801	-0.896*	-0.302
	(0.375)	(0.469)	(0.485)	(0.399)	(0.464)	(0.379)	(0.387)	(0.361)	(0.554)	(0.462)	(0.453)	(0.332)
FY2009, post PCT	0.0897	-0.13	-0.115	-0.379	-0.21	0.137	0.103	-0.0176	-0.247	-0.578	-0.11	0.0899
	(0.322)	(0.408)	(0.387)	(0.326)	(0.225)	(0.256)	(0.247)	(0.324)	(0.335)	(0.389)	(0.275)	(0.271)
FY2010, post PCT	-0.28	-0.236	0.13	-0.152	-0.0295	0.362	0.294	0.42	0.0556	-0.1	-0.231	-0.254
	(0.270)	(0.459)	(0.437)	(0.279)	(0.317)	(0.324)	(0.276)	(0.258)	(0.342)	(0.256)	(0.291)	(0.231)
Post Own Focused Review	-0.404	-0.779*	-0.720**	-0.00924	0.172	0.503	0.431	0.46	-0.247	0.0946	-0.357	-0.501
	(0.566)	(0.424)	(0.346)	(0.359)	(0.394)	(0.398)	(0.395)	(0.322)	(0.366)	(0.341)	(0.281)	(0.416)
Observations	10,771	10,814	10,845	10,868	10,874	10,875	10,876	10,878	10,874	10,863	10,854	10,843
R-squared	0.885	0.892	0.902	0.902	0.902	0.898	0.896	0.901	0.908	0.914	0.916	0.916

Covariates include assignment variables and office environment variables as the main specification. Weighted by state population. Robust standard errors in parentheses, clustered by DDS office. *p<0.10   ** p<0.05   *** p<0.01

Table 5. Instrument Variables Estimates of SSDI Allowance on Labor Market Outcomes, Four Years after Decision

VARIABLES	Earnings	Earnings > $0	Earnings > $1,000	Earnings > SGA	Number of employers	SE Income	SE income > $0
Appellate Allowance	-4,140.455***	-0.221***	-0.220***	-0.141***	-0.430***	-207.379***	-0.017***
	(259.106)	(0.013)	(0.012)	(0.009)	(0.028)	(60.067)	(0.004)
	-15.980	-17.407	-17.843	-16.338	-15.274	-3.452	-4.244
Observations	4,128,896	4,128,896	4,128,896	4,128,896	4,128,896	4,128,896	4,128,896
R-squared	0.093	0.126	0.124	0.085	0.110	0.005	0.013
MeanY	$2,871	0.206	0.178	0.0860	0.380	$212	0.0250

Notes: All models include fixed effects for year-month and DDS office as well as individual and case characteristics. Standard errors in parentheses, clustered by DDS office. *** p<0.01, ** p<0.05, * p<0.1, Earnings in 2014 dollars.

Table 6. IV Estimates on Labor Market Outcomes, Robustness

	BASE MODEL		
	(1)	(2)	(3)
Appellate Allowance	-$4,259	-$4,140	-$4,149
	(251)	(259)	(259)
Time fixed effects	year	yr-month	yr-month
Office-Month Env Var	no	no	yes
Observations	4,128,896	4,128,896	4,128,896
R-squared	0.093	0.093	0.093
MeanY	$2,871	$2,871	$2,871

Notes: Standard errors in parentheses, clustered by DDS office. *** p<0.01, ** p<0.05, * p<0.1, Earnings in 2014 dollars.