Social Security Disability Reform and Implications for Targeting and Employment

Hilary Hoynes, University of California, Berkeley and NBER

Nicole Maestas, Harvard University and NBER

Alexander Strand, Social Security Administration

Introduction

- Social Security Disability Insurance (SSDI) provides benefits to workers **who are unable to work due to a medical condition** expected to last at least one year or result in death

- Previous work shows SSDI provides important insurance that mitigates financial and health risks (see e.g., Deshpande, Gross, & Su, 2021; Deshpande & Lockwood 2022)

- After decades of growth, the **SSDI caseload has declined for the last decade** due in part to **declines in allowance rates at the appellate level**

- In this paper, we study the **targeting of SSDI** leveraging the recent period of declining allowance rates and program generosity to investigate:
  - What is the underlying work capacity of denied workers?
  - How does this change as the program becomes less generous?

- We explore mechanisms for the decline in allowance rates by documenting and examining **little-noticed SSA administrative policy reforms**

Main approach

- Combine SSA administrative data covering appellate cases from 2007–2015 with Master Earnings File (from W-2 and 1099s) 1979–2019

- Using random assignment of judges in appellate process, estimate effects of SSDI allowance on subsequent labor supply using **judge IV models** across period of declining allowance rates

- Use untreated (denied) complier means to characterize **how program targeting evolves as allowance rates decline**

- Explore mechanisms for the decline in allowance rates by building the first data set of **little known SSDI reforms at the judge level**, including the timing of hiring and re-training of each judge

- Use **SSDI reforms as IVs** to estimate effects of SSDI allowance on subsequent labor supply – identify earnings capacity of applicants affected by reforms

Main Findings

- We document the recent decline in SSDI caseload and central role of appellate allowance rate

- Judge IV models show that SSDI allowance leads to large reductions in labor supply

- However, untreated complier means show that denied compliers have very low work capacity in the current SSDI program

- Unpacking the IV estimates by year show that the impact of SSDI allowance on labor supply is declining over the period of eligibility tightening, consistent with a declining work capacity

- To speak to mechanisms, we document previously unnoticed SSA policy reforms that explain about 28-36% of the decline in ALJ allowances

- Find policy reforms targeted those applicants with relatively greater work capacity, but eligibility threshold was already below the SGA target, and reforms moved it further away

Prior work and our contributions

- <u>Value of SSDI and targeting</u> – benefits exceed distortionary cost; benefits include reduced crime, financial distress, mortality (Low & Pistaferri, 2015, Deshpande, Gross, & Su, 2021; Deshpande & Lockwood 2022; Deshpande & Mueller-Smith 2022; Gelber et al 2023)

- <u>SSDI and labor supply</u> – SSDI receipt reduces labor supply (Parsons 1980, 1982; Haveman and Wolfe 1984a, 1984b; Bound 1989; Bound and Waidmann 1992; Autor & Duggan 2003, Daly and Burkhauser 2011; Black et al. 2002; Charles et al. 2018; Maestas, Mullen & Strand 2013; French & Song 2014; Liebman 2015; Gelber, Moore & Strand 2017; Maestas, Mullen & Strand 2021)

- <u>DI caseload decline</u> –some work on disability reform in the Netherlands (e.g., Koning & Lindeboom 2015; Degroot & Koning 2015); little work on U.S. SSDI decline (Liu & Quimby 2023)
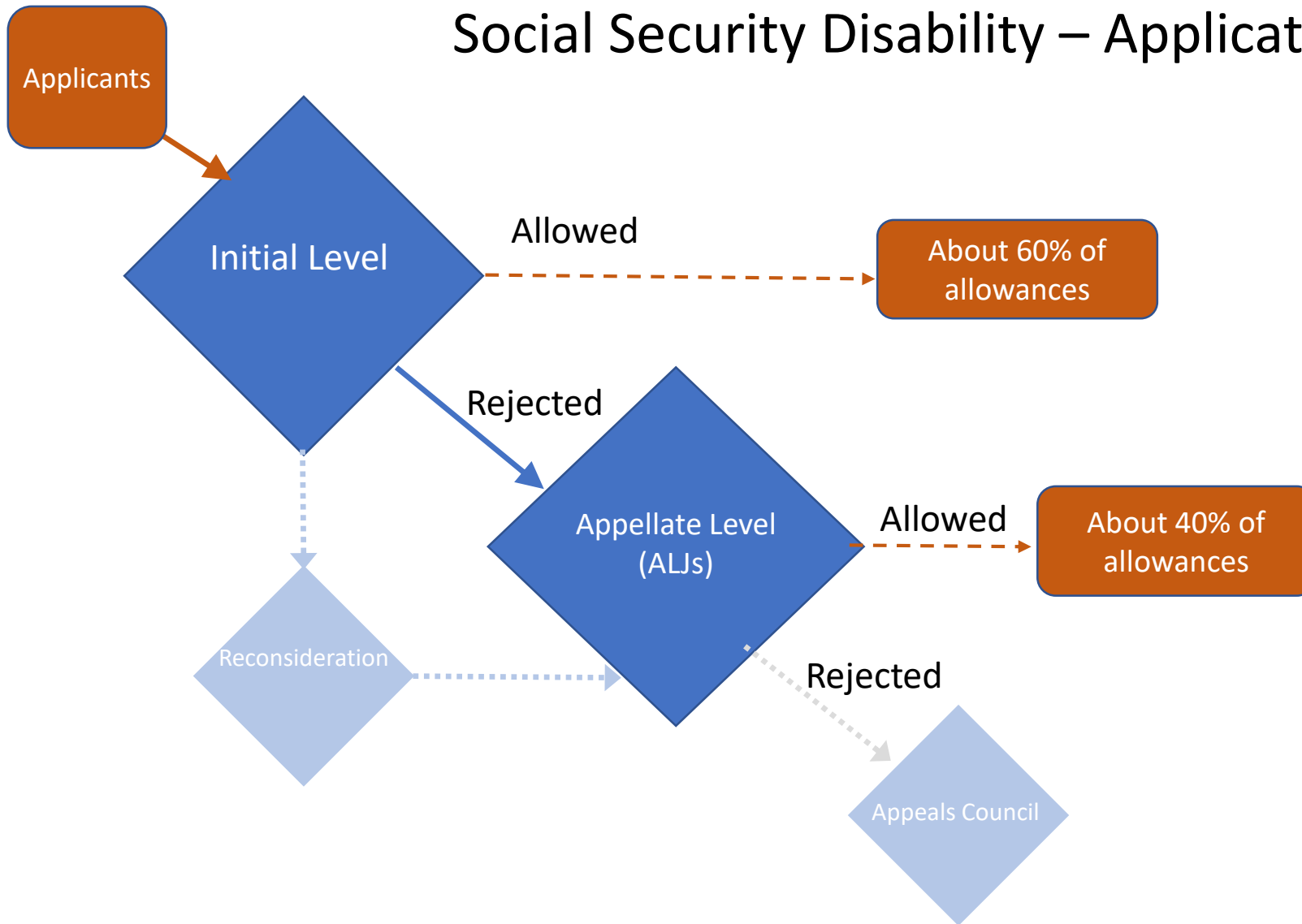
We show that as the SSDI caseload has declined, the program's labor supply disincentive effect has declined, and thus, SSDI has become narrowly targeted at those with very low work capacity. We document an important mechanism for caseload decline.

The SSDI Program and New Facts in the Current Context

SSDI

- 7.6M disabled workers & dependents receive $160B in cash benefits

- Contributory social insurance program; eligibility requires:

  - **Insured status**: work history requirement (paying into contributory system)

  - **Work disabling condition**: unable to engage in any substantial gainful activity because of a medically-determinable physical or mental impairment that is expected to last 12+ months (or result in death)

- If applicant is successful, they receive:

  - Monthly cash benefit - avg. $1,358/mo, until full retirement age

  - Medicare (after 29-month waiting period)

Social Security Disability – Application Process

Applicants → Initial Level

Initial Level —Allowed→ About 60% of allowances

Initial Level —Rejected→ Appellate Level (ALJs)

Reconsideration → Appellate Level (ALJs)

Appellate Level (ALJs) —Allowed→ About 40% of allowances

Appellate Level (ALJs) —Rejected→ Appeals Council

Eligibility Process

Screen out if:
1. Don't meet work history requirement
2. If short duration impairments
3. If earnings > SGA

Allowance can occur if diagnosis meets "**the listings**" (medical evidence to support diag.) or or if residual functional capacity — combined with skills, education, and age — indicate skills not transferable to other jobs
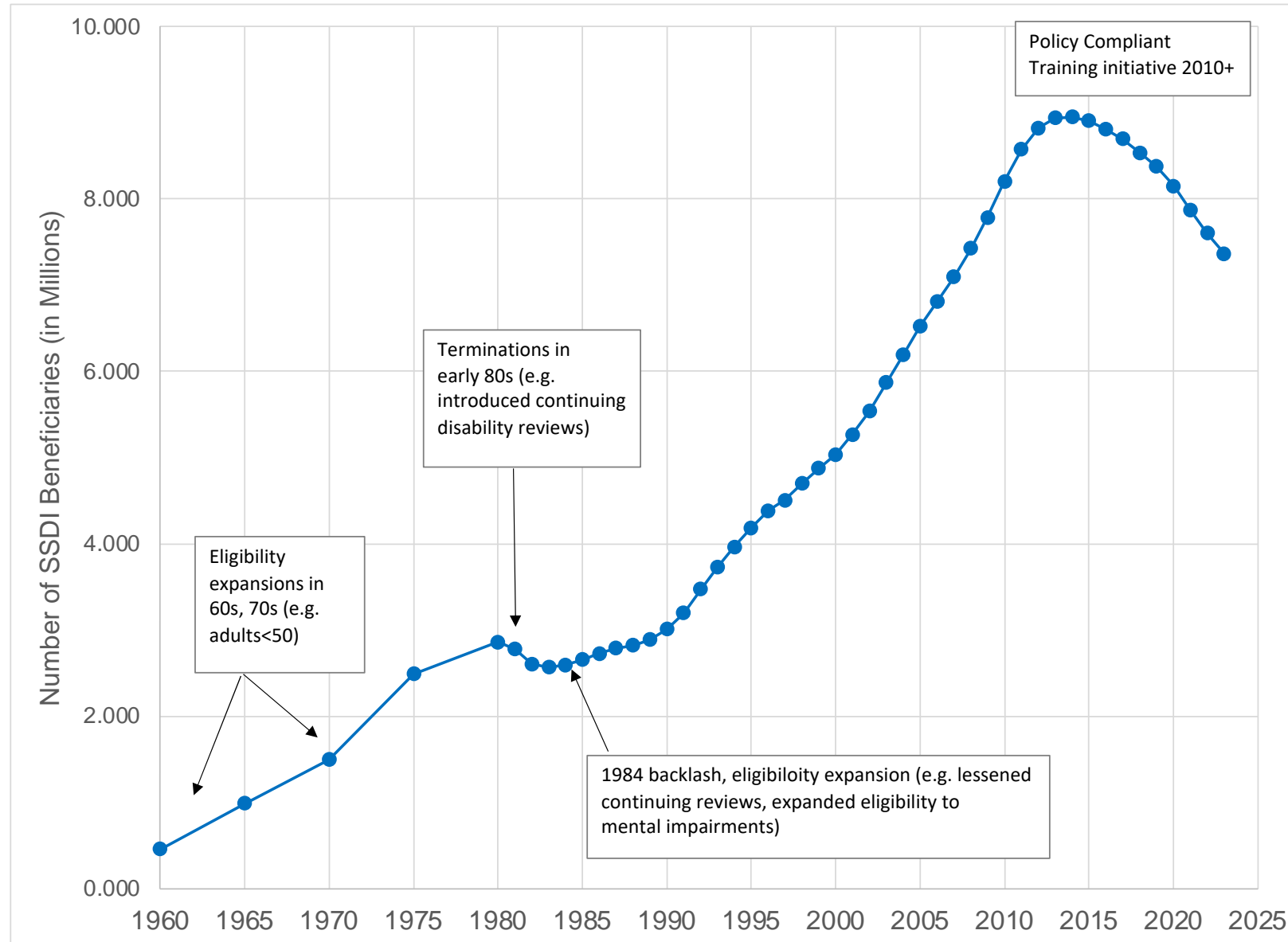
Initial Stage – handled by Disability Determination Service (no in person needed)

Appellate Stage - hearing before Administrative Law Judge (ALJ)

Same eligibility criteria for both

Number of SSDI Disabled Worker Beneficiaries, 1960-2023

Policy Compliant
Training initiative 2010+

Terminations in
early 80s (e.g.
introduced continuing
disability reviews)

Eligibility
expansions in
60s, 70s (e.g.
adults<50)

1984 backlash, eligibiloity expansion (e.g. lessened
continuing reviews, expanded eligibility to
mental impairments)

Against the
historical SSDI
increases, recent
period is significant
break with trend

Source: SSA, Annual Statistical Report on the Social Security Disability Insurance Program, 2023, Table 1
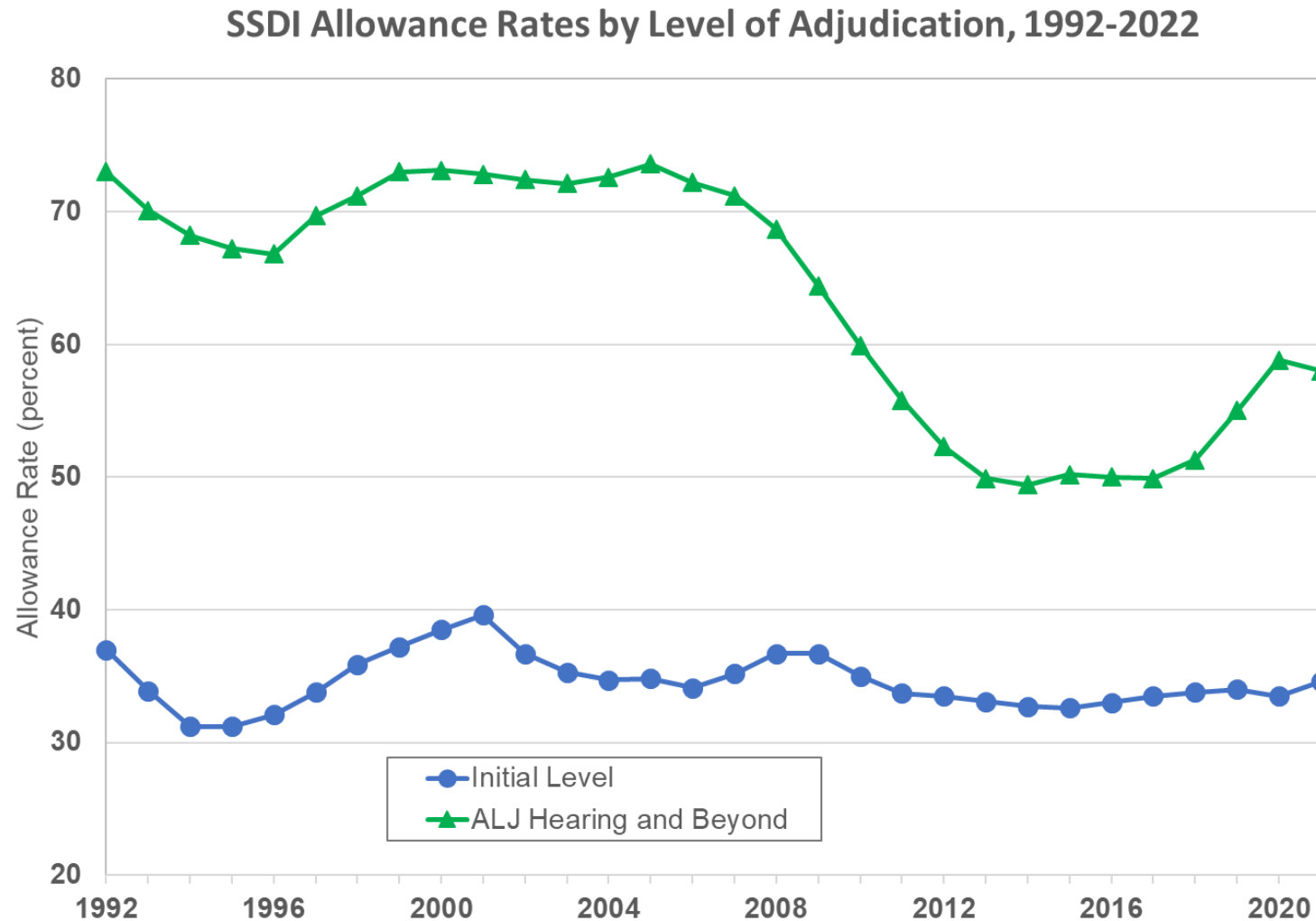
Recent SSDI caseload decline due to **trend break in new awards**

Total SSDI Beneficiaries, Number of Awards and Exits, 2001-2022
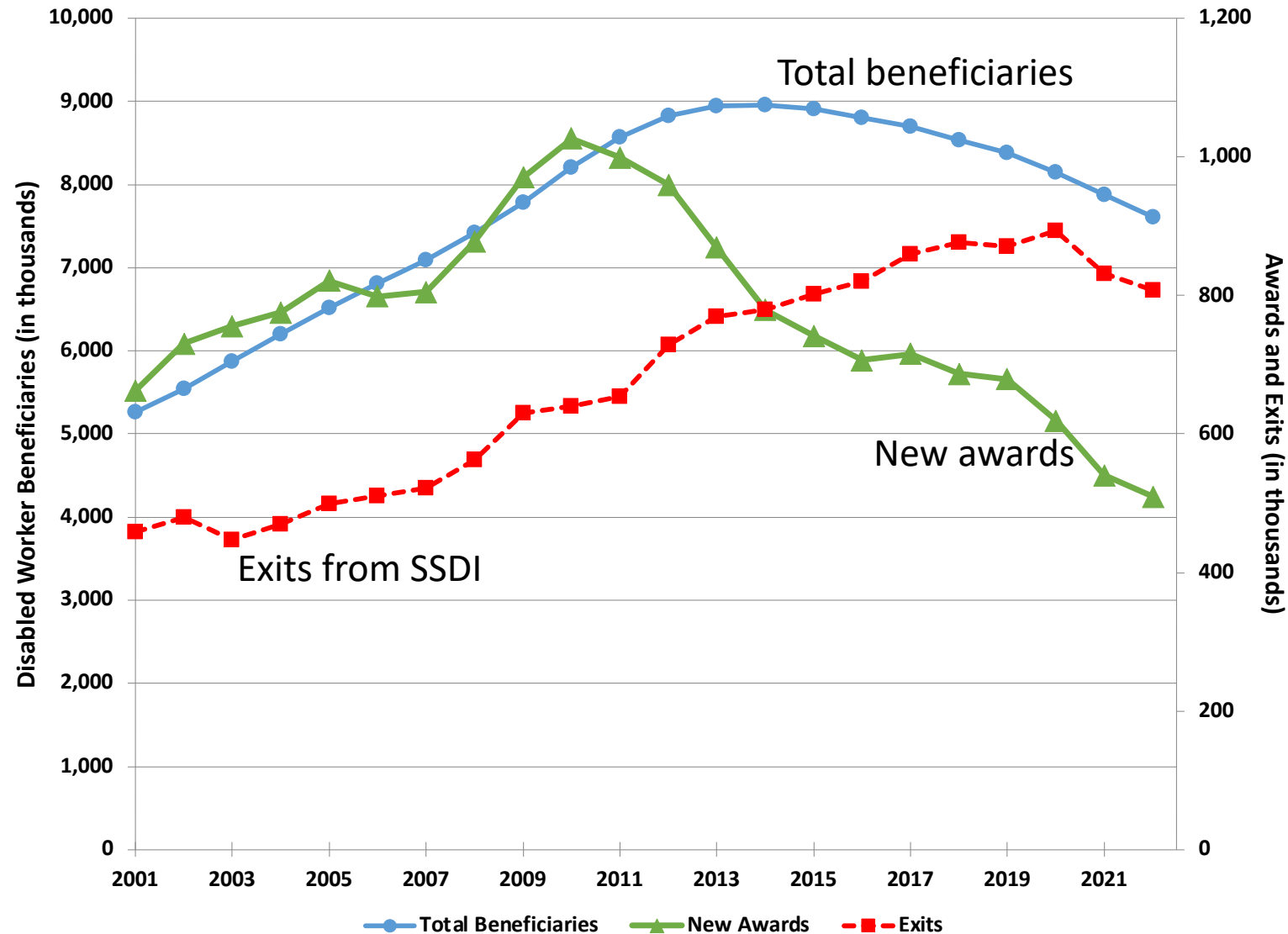
Changes coming from reduction in new awards begin to decline in 2010

Source: SSA, Annual Statistical Report on the Social Security Disability Insurance Program, 2021, Tables 1, 35, and 49

Decline in new awards due to **fewer allowances at appellate level**

SSDI Allowance Rates by Level of Adjudication, 1992-2022

Appellate Allowances
Peak of 74% in 2005; down to 50% in 2015; 32 percent decline

Initial Allowances
No change during this period (~32%) despite using the same eligibility standard

- Given these recent trends in SSDI, we focus our analysis on the **appeals process**

- In the appeals process, cases are handled by Administrative Law Judges (following denial at initial denial)

- Roughly **40% of awards** are made on appeal

Effects of Reforms on SSDI Targeting: Conceptual Approach

Targeting – Conceptual Approach

% of applicants

| All judges allow | Some do, some don't | No judges allow |
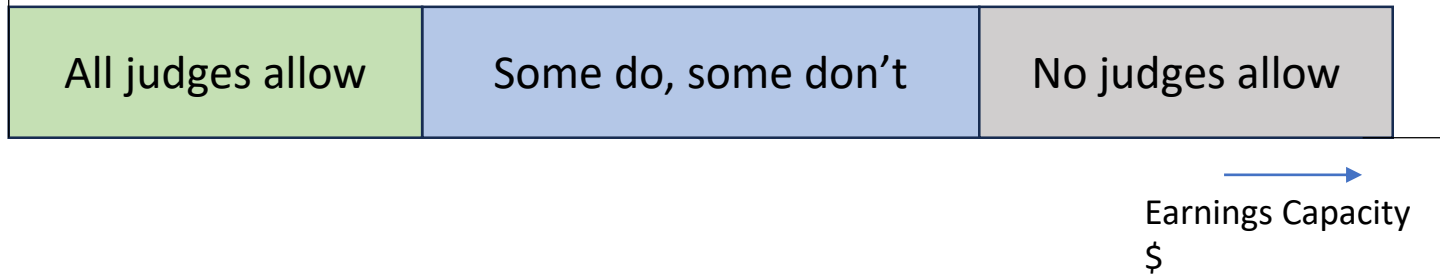
Earnings Capacity
$

*Changes in this distribution over time reveal shifts in program targeting*

SSDI rules:
- If earnings capacity > SGA → Deny
- If earnings capacity < SGA → Allow
- SGA $13,080/year in 2015; $18,600 in 2024

Imperfect information means some judges allow and others don't in range around SGA threshold (blue area)

Allowance rate = green area + portion of blue area allowed

Applying this framework to period of falling allowance rates

SGA?

% of applicants

| All judges allow | Some do, some don't | No judges allow |

Earnings Capacity $
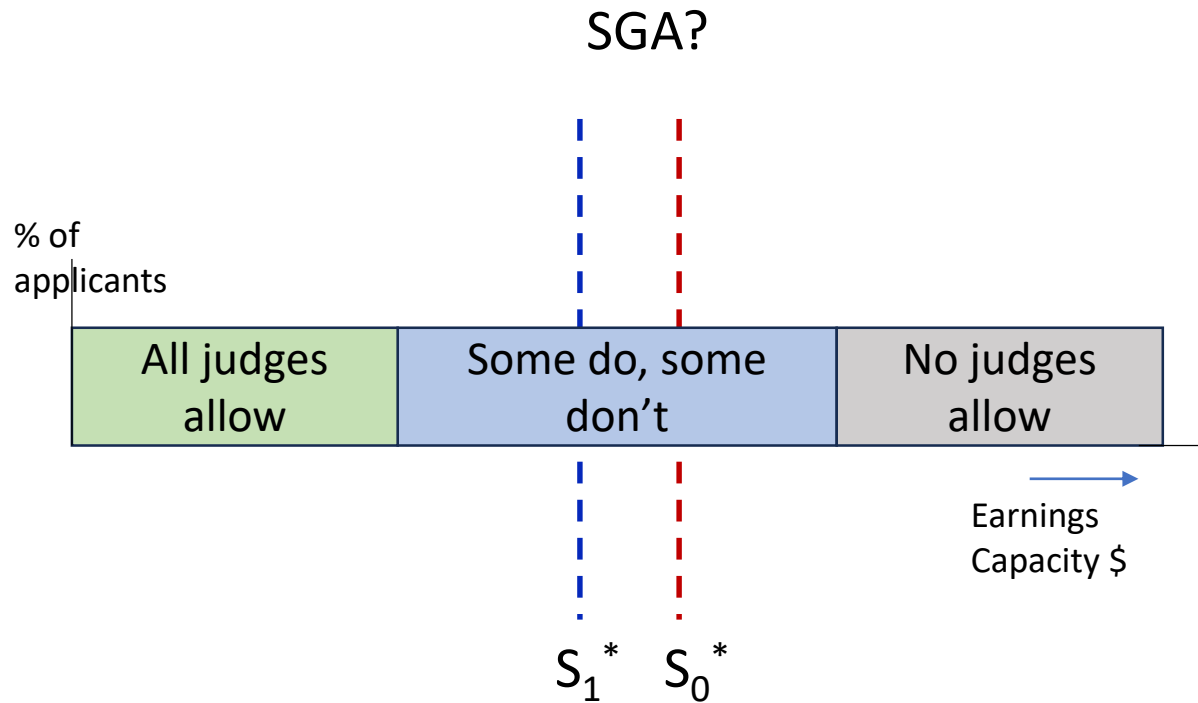
S_1^* S_0^*

A judge allows if their allowance threshold $S^* >$ claimant's earnings capacity
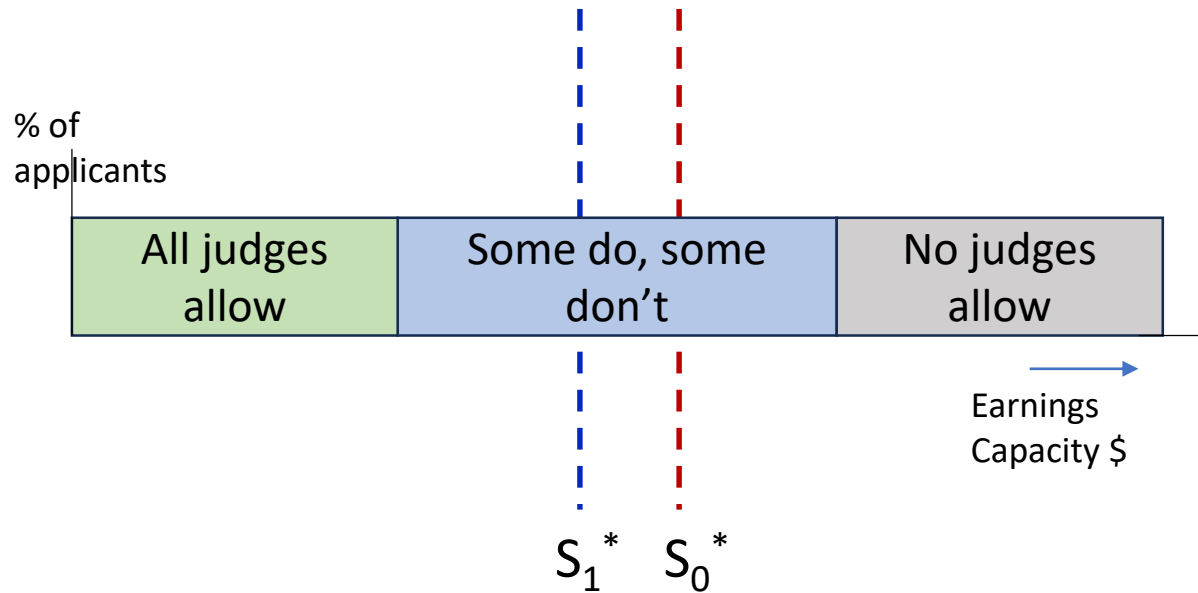
$S_0^* =$ allowance threshold before tightened eligibility
$S_1^* =$ allowance threshold after tightened eligibility

Declining allowance rates imply $S_1^* < S_0^*$, rejections extend lower into earnings capacity

Applying this framework to period of falling allowance rates

SGA?

% of applicants

All judges allow	Some do, some don't	No judges allow

Earnings Capacity $

S_1^* S_0^*

A judge allows if their allowance threshold $S^* >$ claimant's earnings capacity

Testable Implications // Targeting
A decrease in S^* should lead to a:
- **reduction in work capacity** of marginal claimant
- **smaller (negative) effect of allowance on earnings** because of lower work capacity.

Does decrease in S^* get us closer to or further from the the **target (SGA)**?

Data

Data

- SSA Case Processing and Management System (CPMS) – all SSDI decisions made at the appellate level between 2007 and 2015
  - Identifies decision maker (ALJ)
  - Exclude judges who make < 20 decisions in a year
  - Sample: 3,996,049 allowance decisions; 3,433 decision makers
- Supplement with information from Management Information Electronic Disability Folder, the Appointed Representative Data Base, the Modernized Claim System, the 831 files
- SSA Master Earnings File (from W-2 & 1099s), earnings before & after SSDI decision
  - Earnings in 2014 dollars and include those with zero earnings, top-coded at 99th percentile
  - 1979 – 2019 (we can estimate earnings 4 years after decision for observations)
- For mechanisms, we document and record dates of SSA administrative policy reforms affecting judicial corps; from Office of Appellate Operations

Empirical approach – Judge IV

Judge IV – Effect of SSDI on Labor Supply

$$Y_{idt} = \alpha + D_{j(i)}\beta + X_{idt}\eta + X^A_{j(i)}\lambda + \theta_{dt} + \epsilon_{ijdt}$$

- Y_{idt} = labor supply outcome (e.g., earnings 4 years after decision)

- $D_{j(i)}$ = 1 if i's case heard by judge j is allowed (0 if denied or dismissed)

- X_{idt} = covariates at the case and place level (include state-time unemployment rate)

- θ_{dt} = office by quarter fixed effects (conditional random assignment control)

- $X^A_{j(i)}$ = other conditional random assignment controls (judge: SAA, trainee; case: expedited conditions *[terminal, suicidal, dire need]*)

- Cluster standard errors on judge

Judge IV – Effect of SSDI on Labor Supply

- OLS estimate of β will be biased upward if denied applicants have less severe impairments

- In our 2SLS approach, we use an instrument: judge leniency

- $Z_{j(i)}$ = judge leniency (share allowed) for judge j, residualized for case characteristics, leave out mean

- Design hinges on random assignment of judges to cases; used in prior work (French and Song 2014)

Sample Means – individual claims

	All	Denied	Allowed
Share allowed	0.52	0	1
Demographics			
Female	0.52	0.51	0.52
Concurent claim	0.57	0.66	0.50
Noncitizen	0.00	0.00	0.00
in jail	0.00	0.00	0.00
Age 30-39	18.3%	23.1%	14.8%
Age 40-49	36.8%	38.7%	35.3%
Age 50-54	20.1%	15.9%	23.1%
Age 55-61	15.5%	10.0%	19.5%
Age 62+	1.5%	0.9%	1.9%
BMI	39	39	40
Pain indicated	0.93	0.91	0.94
Expedited - terminal	0.01	0.00	0.01
Expedited - suicidal	0.01	0.01	0.01
Expedited - dire need	0.03	0.02	0.03
Primary diagnosis			
Musculoskelatal	0.43	0.40	0.46
Respiratory	0.03	0.03	0.03
Cardiovascular	0.07	0.06	0.08
Endocrine	0.04	0.04	0.05
Neurological	0.07	0.06	0.07
Mental	0.19	0.23	0.16
Case characteristics (prior to appeal)			
Months between onset and application	23.0	29.6	18.1
Presence of second diagnosis	0.81	0.79	0.82
Initial case denial - technical	0.07	0.09	0.05
Initial case denial - diag not severe	0.21	0.25	0.18
Initial case denial - able to do prev job	0.31	0.27	0.34
Initial case denial - able to do job; grid	0.40	0.38	0.42
Unemployment rate	7.4	7.6	7.3
Observations	3,996,049	1,694,797	2,301,252

Allowed are: older, not mental impairment

Sample Means – by judge

	All years	2007	2011	2015
Number of decisions per judge per year	269	271	254	276
Allowance rate	0.617	0.660	0.655	0.466
Share SAA	0.150	0.008	0.244	0.006
Share trainee	0.167	0.009	0.189	0.137
Number	3433	1106	2010	1552

# Effects of SSDI Allowance on Future Labor Market Outcomes

Judge IV model

First stage – full sample

	(1)	(2)
Judge Allowance	0.957***	0.975***
	(0.002)	(0.003)
Observations	3,996,049	3,996,049
Office x Quarter	yes	yes
Case covariates	no	yes
R-squared	0.203	0.133
Mean allowance	0.576	0.576
1st stage F-stat.	285294	105638

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Local Linear First Stage Regression
And Distribution of Judge Propensities,
Net of Claimant Characteristics
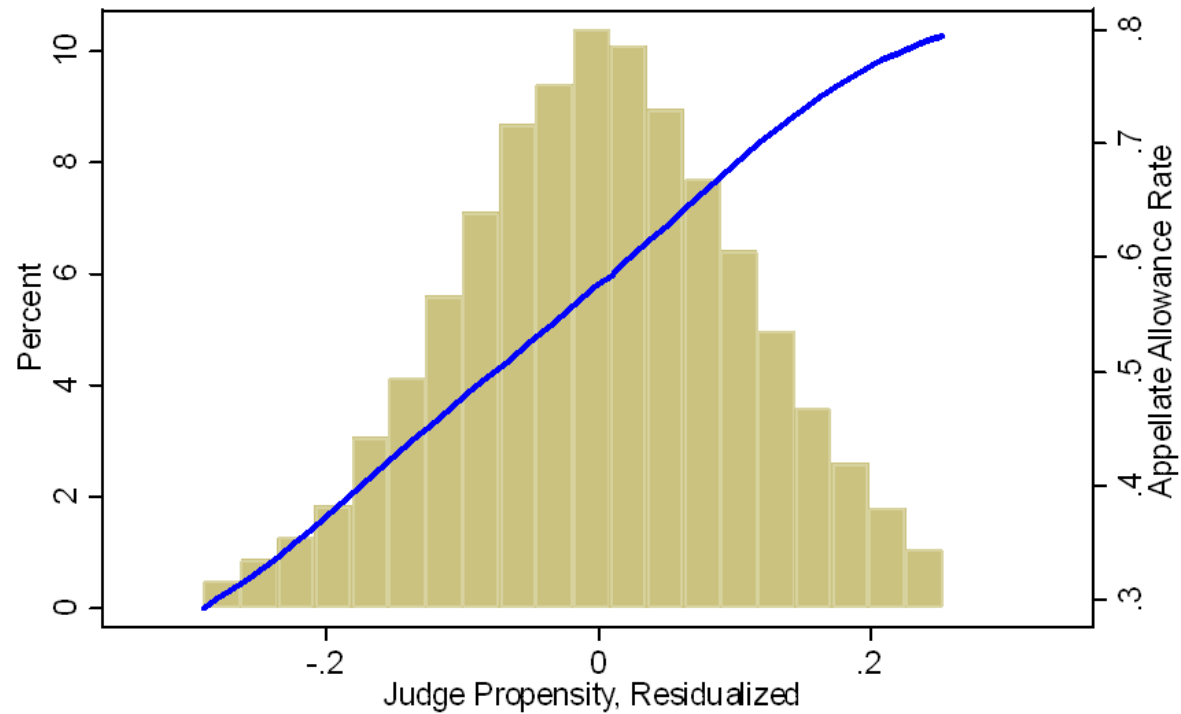
Local Linear Regression
Monotonicity looks good

Judge IV Estimates of SSDI Allowance on Labor Market Outcomes
4 years after decision (2023 $)

	(1) Earnings	(2) Earnings > $0	(3) Earnings > $1,000	(4) Earnings > SGA	(5) Number of employers	(6) SE Income	(7) SE income > $0
Appellate Allowance	-4,005***	-0.151***	-0.146***	-0.094***	-0.279***	-252***	-0.017***
	(66)	(0.002)	(0.002)	(0.002)	(0.005)	(15)	(0.001)
Observations	3,996,049	3,996,049	3,996,049	3,996,049	3,996,049	3,996,049	3,996,049
Office x quarter	yes	yes	yes	yes	yes	yes	yes
Case covariates	yes	yes	yes	yes	yes	yes	yes
R-squared	0.078	0.106	0.106	0.071	0.094	0.004	0.011
MeanY: complier not allowed	5,469	0.265	0.241	0.118	0.48	430	0.034
MeanY: 6-10 yr pre application	33,420	0.857	0.838	0.394	1.63	1249	0.062

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

- The untreated (not allowed) complier means show very low earnings capacity.
- The mean of the lab mkt vars 6-10 yrs prior to initial application reveals a fairly low earnings population.

All earnings are annual and in 2023 dollars include those with zero earnings (SGA $13,080/year in 2015, $17,540 in 2023).

Judge IV Estimates of SSDI Allowance, Robustness to Controls
4 years after decision (2023 $)

	Earnings		Earnings > $0		Earnings > SGA	
Appellate Allowance	-4,101***	-4,005***	-0.157***	-0.151***	-0.097***	-0.094***
	(69)	(66)	(0.003)	(0.002)	(0.002)	(0.002)
Observations	3,996,049	3,996,049	3,996,049	3,996,049	3,996,049	3,996,049
Office x quarter	yes	yes	yes	yes	yes	yes
Case covariates	no	yes	no	yes	no	yes
R-squared	0.049	0.078	0.063	0.106	0.048	0.071

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

- Qualitatively similar results with and without controls

Compliers

- The IV coefficient on allowance pertains to applicants who could have received a different allowance decision in the ALJ process had their case been assigned to a different judge

- 50.1% are compliers

- 29.3% are always takers (allowed under any ALJ)

- 20.6% are never takers (denied under any ALJ)

Calculated using the difference in the predicted allowance rate at the max and min of the instrument with trimming (Dobbie et al 2018).

% of applicants

| All judges allow | Some do, some don't | No judges allow |

Earnings Capacity $
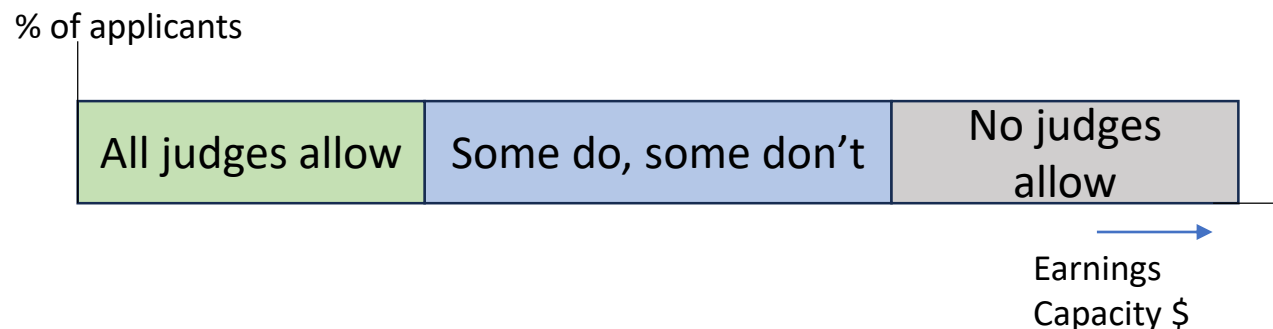
Threats to validity of approach – judge propensity IV

Conditions for LATE interpretation

- Relevance (already shown in F statistics)

- Monotonicity (already shown with local linear regression)

- Independence of instruments – explore with series of balance tests and event study

- Exclusion restriction – explore by estimating effect of judge IV on applications and appeals (do declines in ALJ allowance rates affect flow into initial applications or appeals?)

Balance Test 1

	(1) Allowance	(2) Judge IV Z
Demographics		
Female	0.007***	-0.000
	(0.001)	(0.000)
Concurent claim	-0.080***	-0.001***
	(0.001)	(0.000)
Noncitizen	0.016	0.005
	(0.011)	(0.003)
in jail	-0.189***	-0.001
	(0.006)	(0.002)
Age 30-39	0.050***	0.001**
	(0.001)	(0.000)
Age 40-49	0.114***	0.001***
	(0.002)	(0.000)
Age 50-54	0.211***	0.003***
	(0.002)	(0.001)
Age 55-61	0.259***	0.004***
	(0.003)	(0.001)
Age 62+	0.243***	0.004***
	(0.003)	(0.001)
BMI	0.001***	0.000
	(0.000)	(0.000)
Pain indicated	0.026***	0.000
	(0.001)	(0.000)
Primary diagnosis		
Musculoskelatal	0.009***	-0.001
	(0.003)	(0.001)
Respiratory	-0.015***	-0.000
	(0.005)	(0.001)
Cardiovascular	0.008**	0.001
	(0.004)	(0.001)
Endocrine	0.011***	0.001
	(0.004)	(0.001)
Neurological	0.039***	0.001
	(0.004)	(0.001)
Mental	-0.010**	0.001
	(0.004)	(0.001)

(continued)	(1) Allowance	(2) Judge IV Z
Case characteristics (prior to appeal)		
Time between onset and application	-0.002***	-0.000***
	(0.000)	(0.000)
Presence of second diagnosis	0.029***	0.001***
	(0.001)	(0.000)
Initial case denial - technical	-0.049***	0.001
	(0.002)	(0.001)
Initial case denial - earn > SGA	0.002	-0.004
	(0.024)	(0.005)
Initial case denial - diag not severe	-0.044***	0.000
	(0.002)	(0.001)
Initial case denial - able to do prev job	0.016***	0.001
	(0.002)	(0.001)
Initial case denial - able to do job; grid	0.044***	0.001
	(0.002)	(0.001)
Unemployment rate	0.002***	0.001**
	(0.001)	(0.000)
Observations	3,996,049	3,996,049
Office x Quarter	yes	yes
Other Assignment Variables	yes	yes
R-squared	0.156	0.011
MeanY	0.576	0
F - joint significance	1192	4

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

(Column 1) Controls are highly predictive of allowance

(Column 2) Controls do not predict the judge leniency instrument $Z_{j(i)}$ joint F statistic = 4

Balance Test 2

First Stage and Covariate Balance
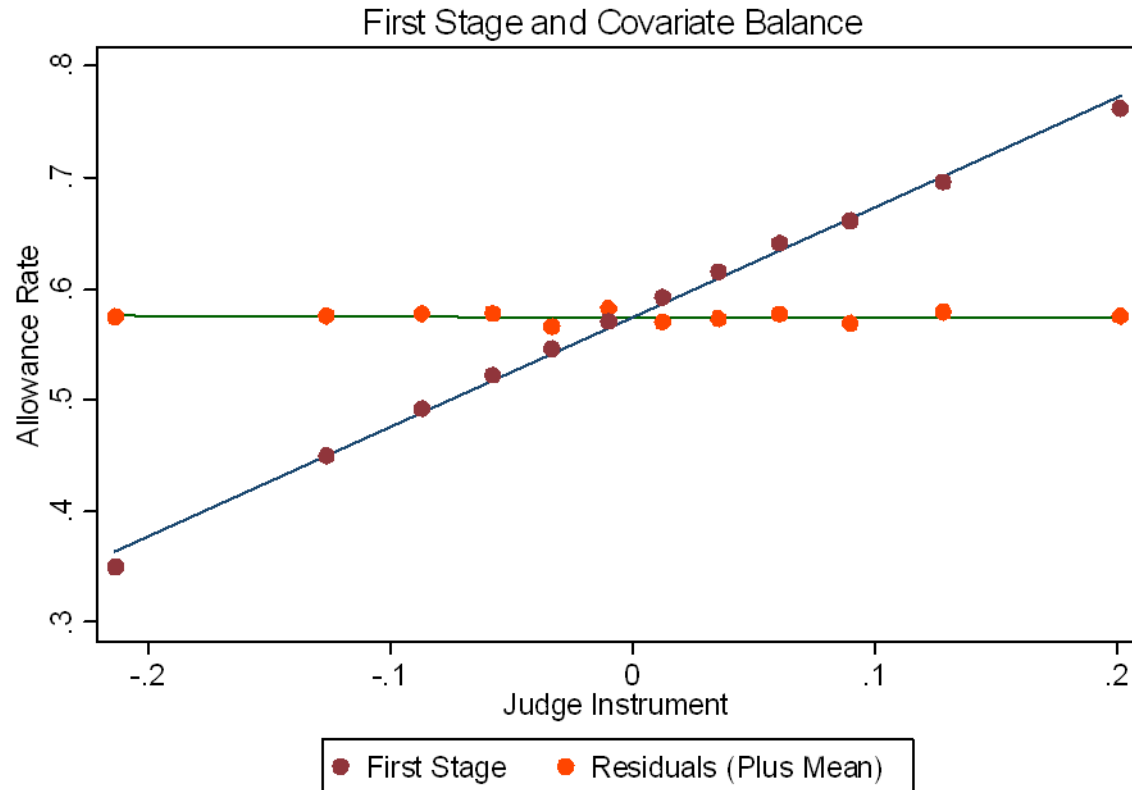
Alternative balance test focusing on unobservables

We compare the first stage residuals and the first stage predicted allowance rate against bins of the judge leniency instrument

Balance test 3 – Event Study

$$Y_{ijdl} = \alpha + \boxed{\sum_{k=-10,4} \delta_k \, Z_{j(i)} 1(l = t + k)} + \sum_{k=-10,4} \gamma_k \, 1(l = t + k) + X_{idt}\eta + X_{j(i)}^A\lambda + \theta_{dt} + \epsilon_{ijdt}$$

Event Study Coefficients Interacted with Judge Leniency

We plot an event study for earnings in the 10 years prior to and 4 years after ALJ decision.

We interact the event study coefficients with the allowance rate of the judge you are (ultimately) assigned to.

No relationship of judge Z to pre-trends, confirming random assignment. As expected being assigned to a more allowing judge does lead to larger effects on earnings (effectively a reduced form in event time).
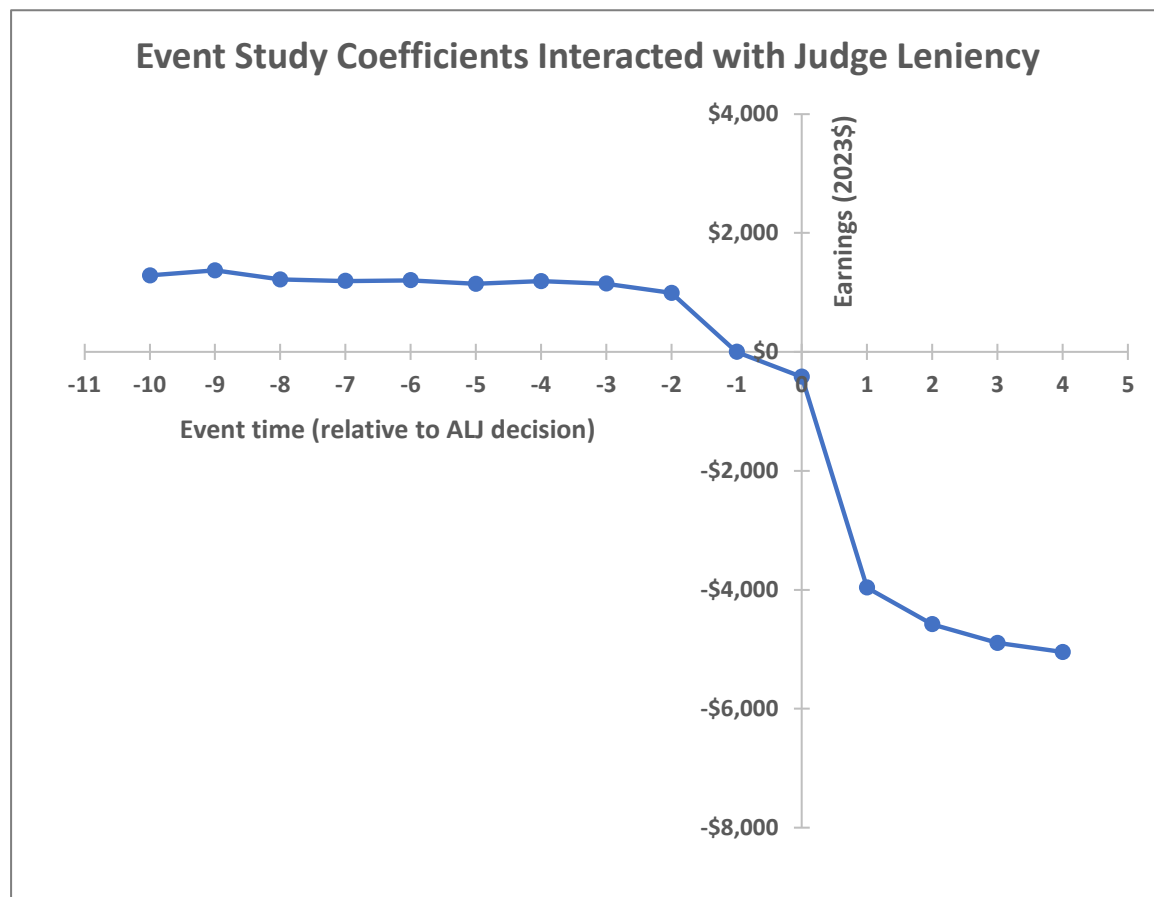
Do the instruments have feedback effects on flows into appeals? No.

Leads and Lags of Log(Appeals) on Judge IV Instrument

		judgeiv		Observations	R-squared
Lags	5 Years	0.487	(0.412)	2,192	0.852
	4 Years	0.345	(0.494)	2,740	0.849
	3 Years	0.210	(0.381)	3,288	0.845
	2 Years	0.191	(0.477)	3,836	0.839
	1 Year	-0.356	(0.402)	4,384	0.840
	1 Quarter	-0.162	(0.327)	4,795	0.846
Leads	1 Quarter	0.043	(0.421)	4,795	0.845
	1 Year	-0.181	(0.330)	4,384	0.859
	2 Years	-0.078	(0.277)	3,836	0.873
	3 Years	-0.158	(0.264)	3,288	0.884
	4 Years	-0.228	(0.225)	2,740	0.900
	5 Years	-0.065	(0.282)	2,192	0.904

Covariates include the same assignment variables and case characteristics Judge IV variables as the main specification. Weighted by state population. Robust standard errors in parentheses, clustered by office. *p<0.10   ** p<0.05   *** p<0.01
*Only offices with complete data included.

Judge IV instrument has no impact on flows into appeals

Data collapsed to office x quarter.

These estimates are from a regression of the log of office-quarter counts of DI appellate cases, with controls (case characteristics, local labor market vars, office-quarter FE).

Effects of SSDI Allowance on Labor Market Outcomes

Judge IV model

Program Targeting – Estimates across Time

Extending Judge IV Model to Examine Period of Decline in Allowance Rates

SSDI Allowance Rates by Level of Adjudication, 1992-2022
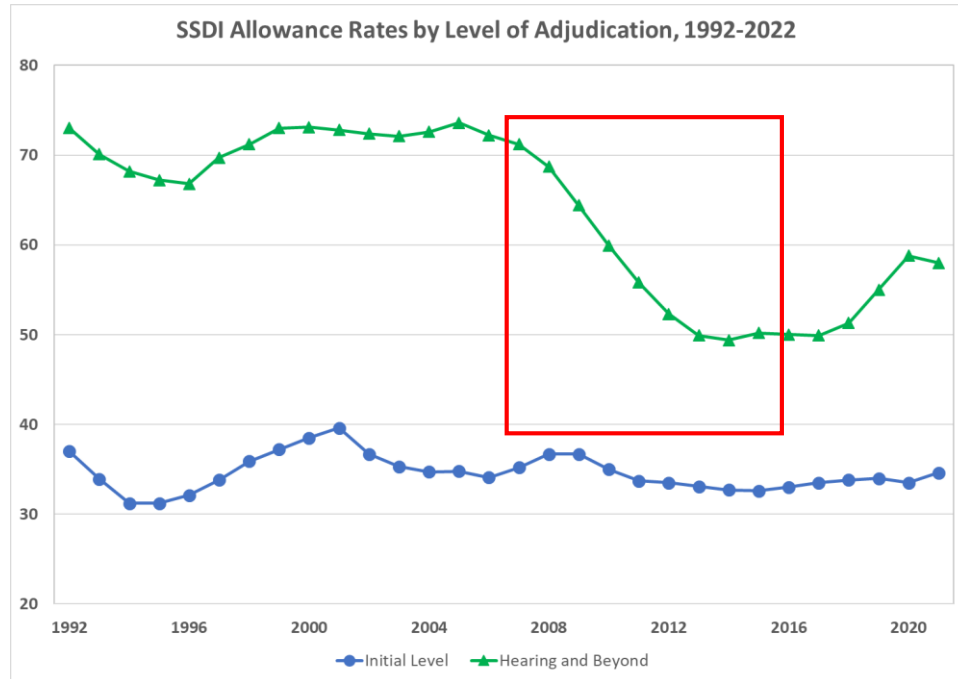
- We estimate the judge IV by year (3 year moving averages) for 2008-2014

- Recalculate judge leniency in each period

- Allows us to examine how the impacts on work capacity change and implications for program targeting as the program becomes more stringent

First Stage, by year

	2008	2009	2010	2011	2012	2013	2014
Judge Allowance	0.966***	0.970***	0.968***	0.963***	0.957***	0.951***	0.945***
	-0.003	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
Observations	1,101,100	1,281,426	1,422,616	1,503,218	1,507,668	1,446,204	1,349,120
Office x Quarter	yes	yes	yes	yes	yes	yes	yes
Case covariates	yes	yes	yes	yes	yes	yes	yes
R-squared	0.185	0.194	0.199	0.197	0.184	0.168	0.155
Mean allowance	0.678	0.659	0.626	0.583	0.538	0.501	0.476
1st stage F-stat.	126682	178940	233240	253915	234273	198984	157201

Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Judge IV Estimates of SSDI Allowance on Labor Market Outcomes
4 years after decision (2023 $), 3 year moving average

	2008	2009	2010	2011	2012	2013	2014
Earnings							
Allowance	-3,315***	-3,330***	-3,672***	-4,112***	-4,520***	-4,548***	-4,654***
	(114)	(100)	(93)	(92)	(96)	(108)	(112)
Mean, complier not allowed	4,315	4,372	4,884	5,519	6,109	6,313	6,654
Estimate / mean *100	-76.8%	-76.2%	-75.2%	-74.5%	-74.0%	-72.0%	-69.9%
Earnings > $0							
Allowance	-0.127***	-0.130***	-0.143***	-0.156***	-0.169***	-0.169***	-0.170***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Mean, complier not allowed	0.214	0.224	0.248	0.274	0.294	0.302	0.31
Estimate / mean *100	-59.3%	-58.0%	-57.7%	-56.9%	-57.5%	-56.0%	-54.8%
Earnings > SGA							
Allowance	-0.075***	-0.078***	-0.085***	-0.096***	-0.106***	-0.109***	-0.113***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Mean, complier not allowed	0.092	0.094	0.103	0.118	0.131	0.139	0.149
Estimate / mean *100	-82.0%	-83.0%	-82.5%	-81.4%	-80.9%	-78.4%	-75.8%
Observations	1,101,100	1,281,426	1,422,616	1,503,218	1,507,668	1,446,204	1,349,120
Office x Quarter	yes	yes	yes	yes	yes	yes	yes
Case covariates	yes	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

All earnings are annual and in 2023 dollars include those with zero earnings (SGA $13,080/year in 2015, $17,540 in 2023).

What do these results say about targeting in SSDI

Recall that for the full period, the compliers who are not allowed have very low work capacity

Untreated (denied) complier means (2023$)

Earnings 4 yrs post decision	Judge IV
Earnings	$5,469
Earnings > $0	0.265
Earnings > SGA	0.118

Testable Implications // Targeting
A decrease in S*
- should lead to a **reduction in work capacity** of marginal claimant
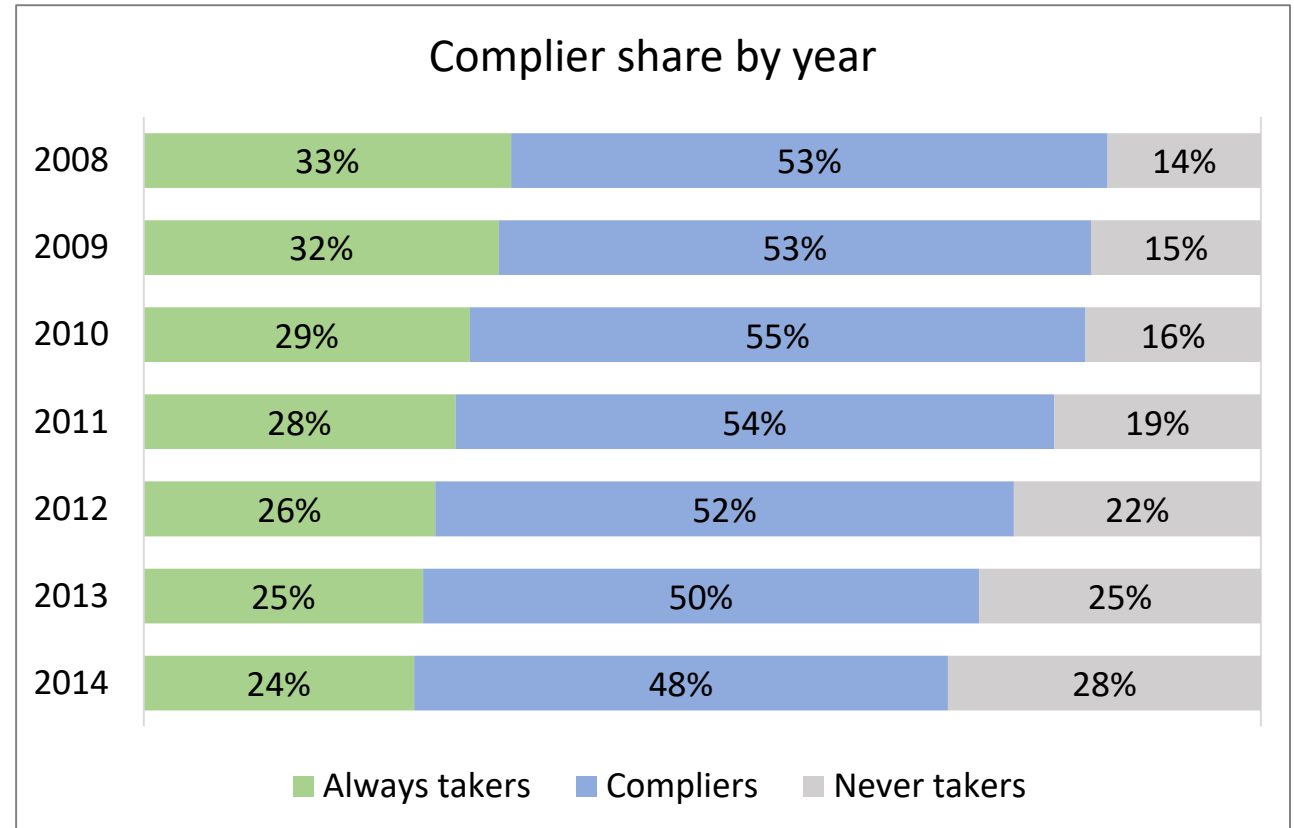- should lead to **smaller (negative) effects of allowance on earnings** because of lower work capacity.

Does decrease in S* get us closer to or further from the the **target (SGA)**?

Targeting: What happens to complier shares as DI eligibility tightens?
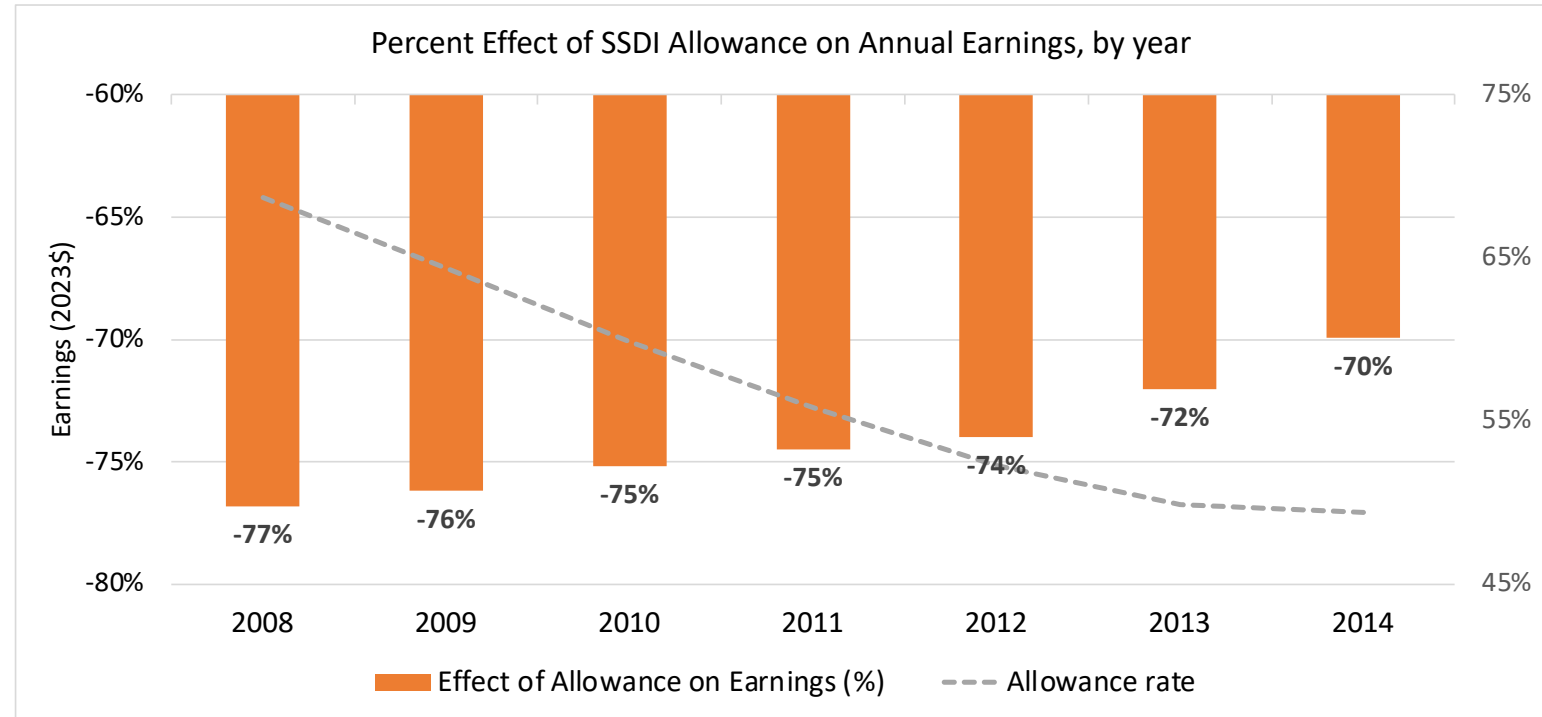
Complier share falls, fewer always takers and more never takers

The leftward shift in the eligibility margin → eligibility margin shifts down the work capacity distribution
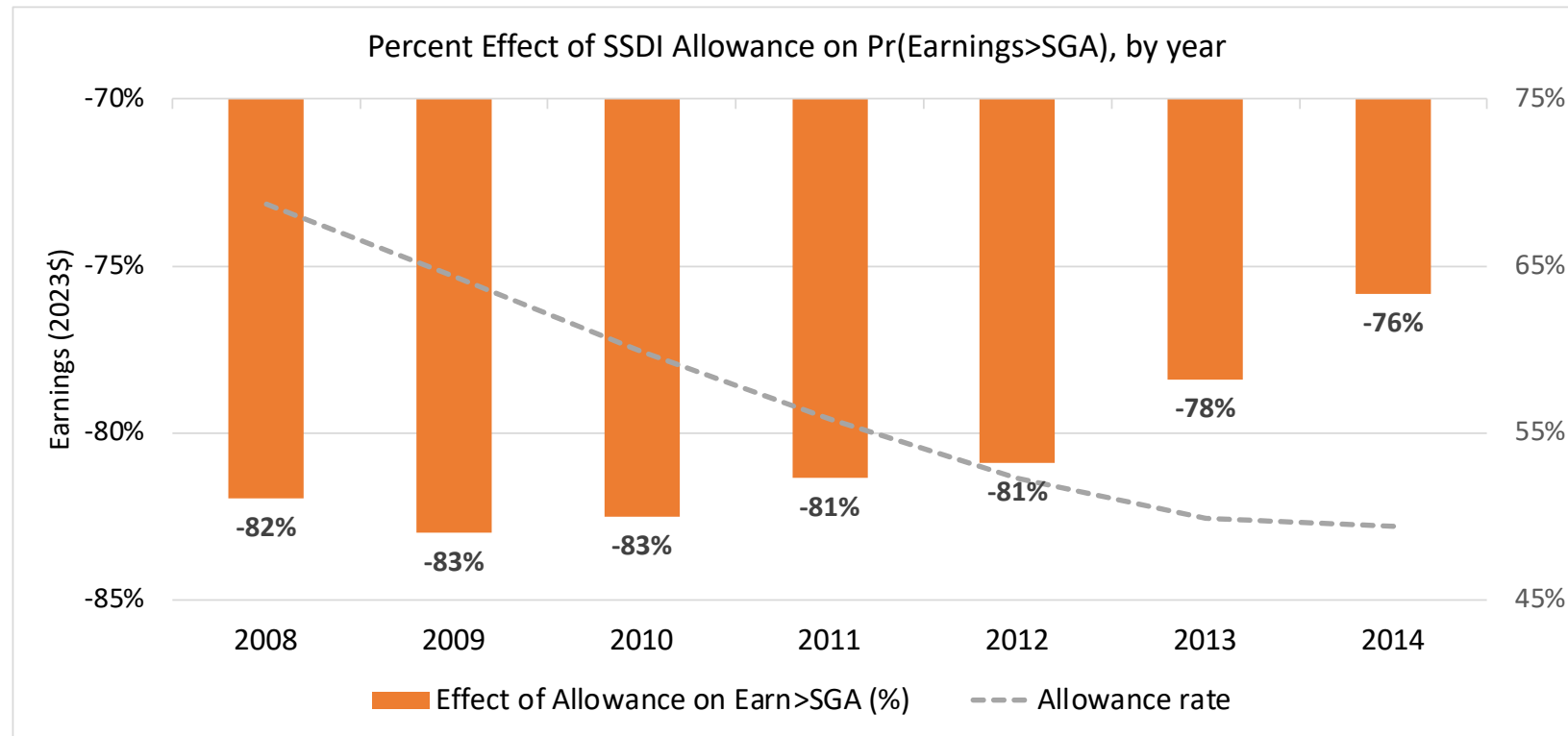
Complier share by year

Year	Always takers	Compliers	Never takers
2008	33%	53%	14%
2009	32%	53%	15%
2010	29%	55%	16%
2011	28%	54%	19%
2012	26%	52%	22%
2013	25%	50%	25%
2014	24%	48%	28%

■ Always takers ■ Compliers ■ Never takers

Targeting: What happens to treatment effects when eligibility tightens?

- <u>Prediction:</u> a decrease in S* should lead to smaller (negative) effects of allowance on earnings because marginal claimant has less work capacity.
- We find that the effect of DI allowance, relative to complier mean, declines over time (less work capacity at the margin)

Percent Effect of SSDI Allowance on Annual Earnings, by year

Earnings (2023$)

Year	Effect of Allowance on Earnings (%)
2008	-77%
2009	-76%
2010	-75%
2011	-75%
2012	-74%
2013	-72%
2014	-70%

- - - Allowance rate

Legend: ■ Effect of Allowance on Earnings (%)   - - - Allowance rate

Targeting: Does tightening of eligibility get us closer or further from SGA (as the elig standard)

- Effect of allowance on Pr(Earn>SGA) also declines (in abs value) by year
- Also, in all years, the compliers that are not allowed have very low work capacity >> only 11.6% earn above SGA
- This suggests that further restricting DI moves us further away from SGA

Percent Effect of SSDI Allowance on Pr(Earnings>SGA), by year

Legend: Effect of Allowance on Earn>SGA (%) ; Allowance rate

Bar values: 2008: -82%, 2009: -83%, 2010: -83%, 2011: -81%, 2012: -81%, 2013: -78%, 2014: -76%

Y-axis (left): Earnings (2023$), -70% to -85%; Y-axis (right): 45% to 75%

Magnitudes

- French & Song (2014) used judge propensity IV to estimate causal effects of SSDI allowance on labor supply 3 yrs post decision in 1990-1999.

- In our analysis we used judge propensity IV to estimate causal effects of SSDI allowance on labor supply 4 yrs post decision in 2007-2015.

- Our estimates are 42% lower than French and Song, consistent with the possibility that SSDI was less strict than in French and Song period

Mechanisms

What led to reductions in ALJ allowance rates?

SSDI Administrative Policy Reforms – Overview and approach

- Studies have investigated the role of demographic and economic forces affecting DI caseload growth (e.g., Autor & Duggan 2003; Liebman 2015; Maestas, Mullen & Strand 2021) . But there has been little attention to recent **administrative policy reforms**

- We document SSA reforms including (1) a generational turnover in the judicial corps and (2) the launch of training initiatives to improve policy compliance. Notably, these affect the appellate process, not the initial process (consistent with allowance rates falling at ALJ level).

- We build the first data set of SSDI reforms at the judge and office level including the **timing of receipt of policy compliant training for each administrative law judge**

- We examine how these policy reforms affect allowance rates (first stage)

- Taking advantage of the **random assignment of judges to cases**, we use an IV approach to estimate causal effects of SSDI allowance on subsequent employment using the SSA policy reforms as instruments

SSA Policy Reform (1): generational turnover in the judicial corps

- From 1999-2007, very few ALJs were hired (throughout fed. govt.)
  - 1999-2003: "ALJ register" (list of qualified candidates) was suspended due to lawsuit over veterans' preference in hiring
  - 2003-2007: lawsuit resolved but OPM designed new exam
  - **October 2007**: OPM established new register, new exam → hiring resumed

- Surge in hiring beginning in FY2008 (e.g. SSA hired 185 judges in FY2008 adding to 1,007 incumbent judges)

- Half of judge corp were eligible to retire by end of 2008; many did (GAO 2010)

- Additionally, Senior adjudicative attorney (SAA) program begins in Nov 2007 to help with backlog; they can only allow cases (evaluate case for evidence added after initial denial that can support an allowance) before they get to ALJ hearing

SSA Policy Reform (2): trainings to create more policy compliance

- Designed and subsequently rolled out new **policy compliant training** ("PCT") for all ALJs
    - ALJs hired after July 2010 received updated training at hire
    - Other ALJs received training in rotating fashion over summers of 2011-2013
- Also gave **focused reviews of judges** who were outliers, making errors – could have direct or network effects
- ALJ (non)compliance got national attention with Wall Street Journal May 2011 story on corrupt Appalachia attorney and ALJ getting rich off 100% allowance rates; subsequent Congressional hearings

THE WALL STREET JOURNAL.

Disability-Claim Judge Has Trouble Saying 'No'

Near-Perfect Approval Record; Social-Security Program Strained
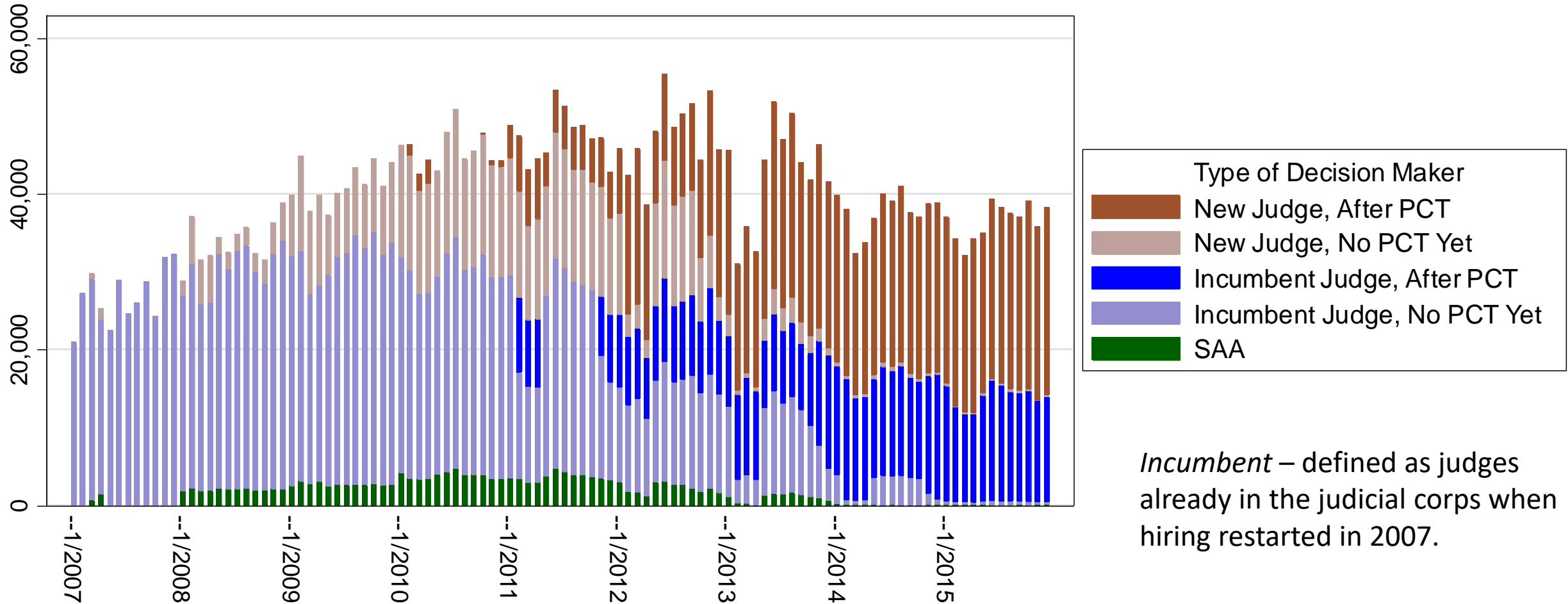
By Damian Paletta

Updated May 19, 2011

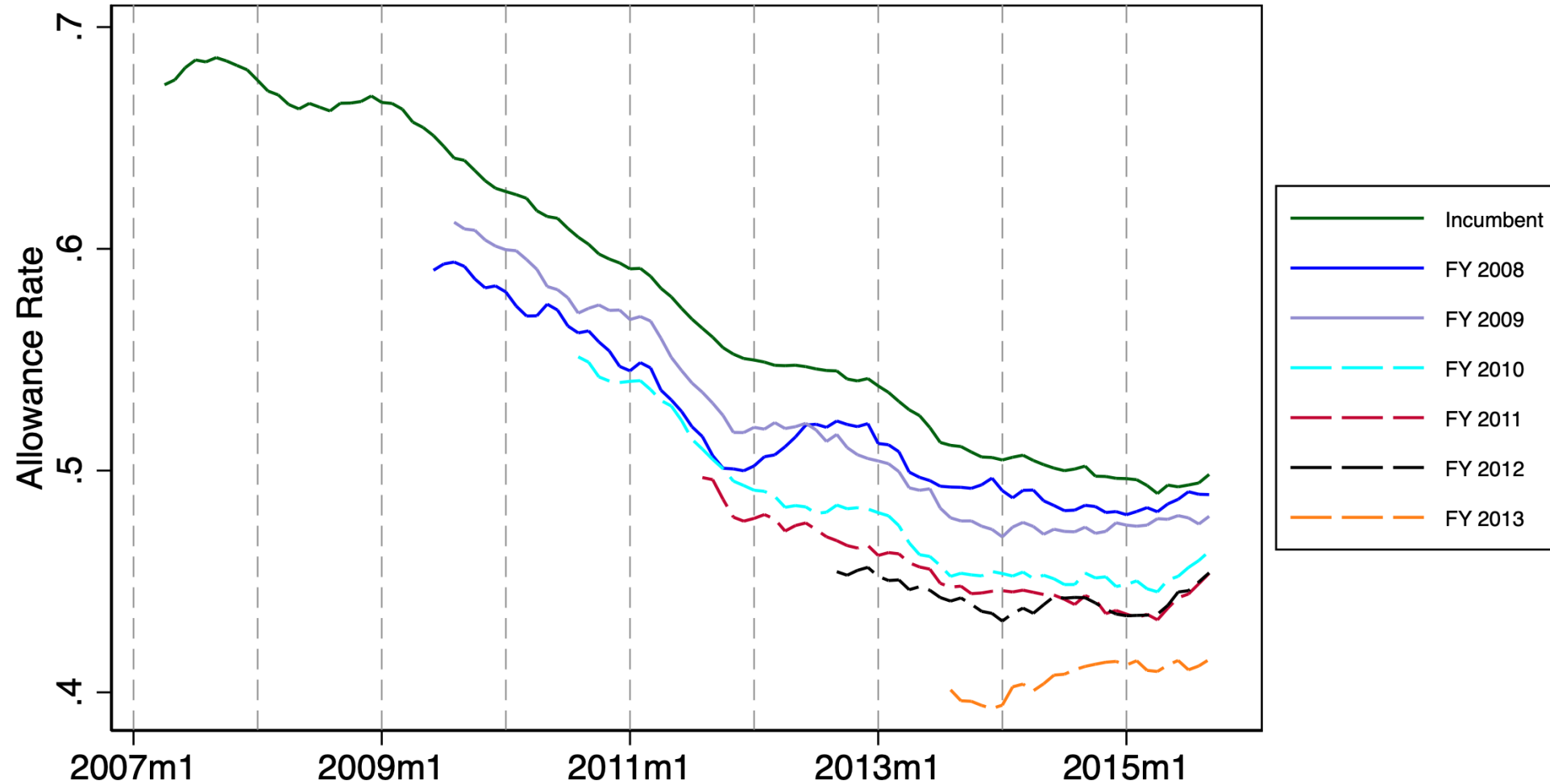# The changing landscape of SSDI appellate decision makers; generational turnover and PCT rollout

Number of Appellate Decisions by Type of Decision Maker and Policy Compliant Training

Type of Decision Maker
- New Judge, After PCT
- New Judge, No PCT Yet
- Incumbent Judge, After PCT
- Incumbent Judge, No PCT Yet
- SAA

Incumbent – defined as judges already in the judicial corps when hiring restarted in 2007.

Note: PCT=Policy Compliant Training.

Allowance Rate by Judge Cohort

Legend:
- Incumbent
- FY 2008
- FY 2009
- FY 2010
- FY 2011
- FY 2012
- FY 2013

Large differences across cohorts

Cascading lower allowance rates

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.

Research design for mechanisms – IV using judge-level policy reforms as instruments

Instruments

Judge Cohort (omitted = incumbent)
FY2008
FY2009
FY2010
FY2011
FY2012
FY2013
FY2014
FY2015

Post Policy Compliant Training
(FY2011+ hired with PCT)
Incumbent, post PCT
FY2008, post PCT
FY2009, post PCT
FY2010, post PCT
Post Own Focused Review

Assignment Variables

Office x Quarter fixed effects
SAA (=1 if SAA)
Trainee (=1 if ALJ in initial 9-mo trainee period)
Vars flagging expedited case

Reforms – at office level
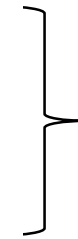
Office-month environment
Share with OTR decisions
Number SAAs / Number ALJs
Post focused review in office

First Stage: Effects of Policy Reforms on Appellate Allowances

Judge Cohort (omitted = incumbent)		
FY2008	-0.033**	(0.013)
FY2009	-0.038***	(0.010)
FY2010	-0.051***	(0.008)
FY2011	-0.071***	(0.008)
FY2012	-0.078***	(0.010)
FY2013	-0.099***	(0.011)
FY2014	-0.130***	(0.022)
FY2015	-0.139***	(0.012)
Post Policy Compliant Training *(FY2011-2015 hired with PCT)*		
Incumbent, post PCT	-0.022***	(0.007)
FY2008, post PCT	-0.0002	(0.011)
FY2009, post PCT	-0.012	(0.008)
FY2010, post PCT	-0.013*	(0.007)
Post Own Focused Review	-0.045**	(0.021)
Observations	3,996,049	
Office x Quarter	yes	
Case covariates	yes	
R-squared	0.133	
Mean allowance	0.576	
1st stage F-stat.	16.5	

FY2011-15 ALJ cohort effects include impacts of PCT

Notes: All models control for case characteristics, body system major class, state unemployment rate, unemployment rate by body system class, and hearing office by quarter FE. SE clustered on judge.

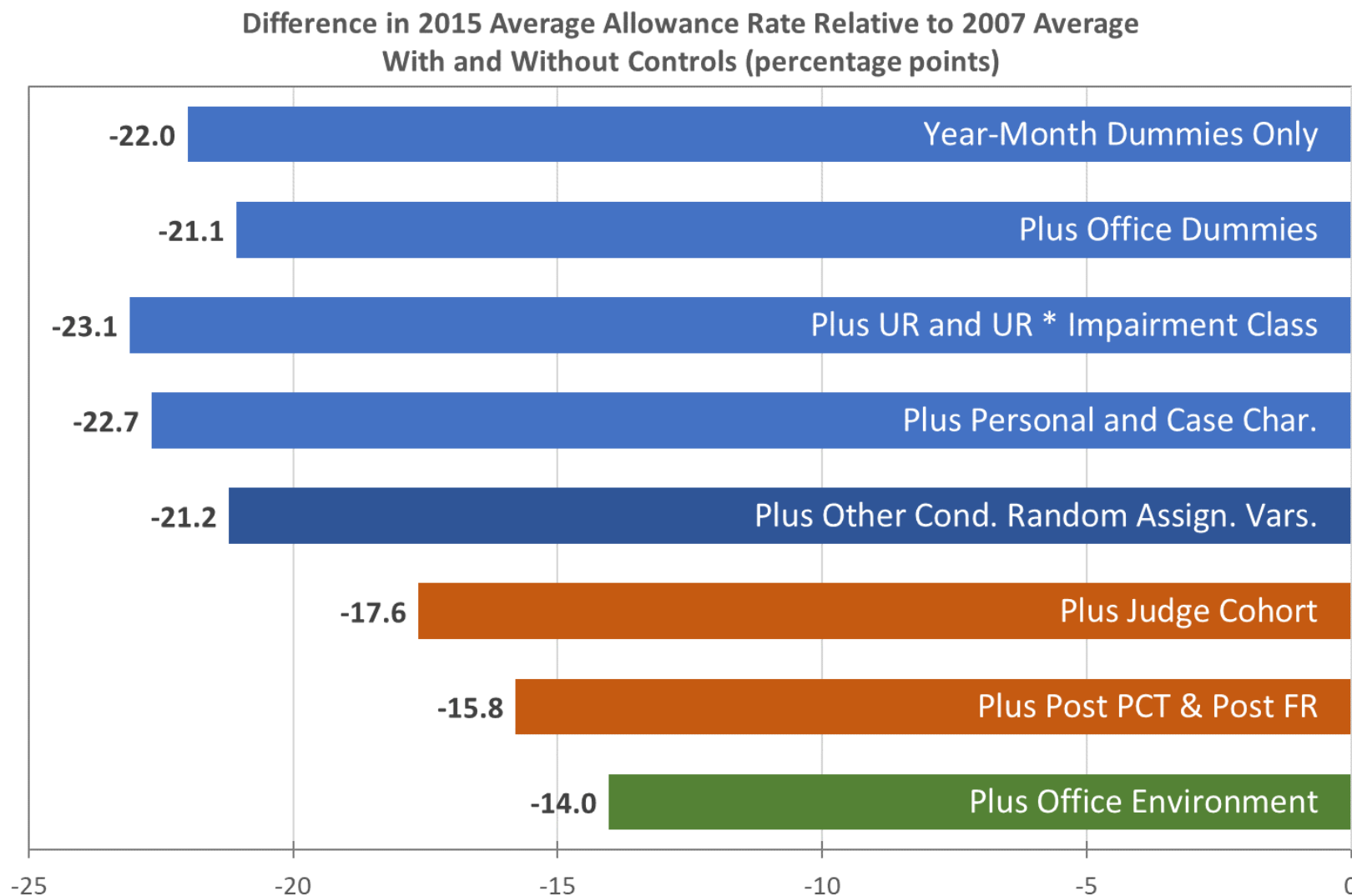# IV Estimates of Effect of SSDI Allowance on Annual Earnings, 4 years after decision *(policy reforms as instruments)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Earnings	Earnings > $0	Earnings > $1,000	Earnings > SGA	Number of employers	SE Income	SE income > $0
Appellate Allowance	-6,127***	-0.143***	-0.225***	-0.225***	-0.434***	-297***	-0.019***
	(349)	(0.01)	(0.01)	(0.01)	(0.03)	(88)	(0.00)
Observations	3,996,049	3,996,049	3,996,049	3,996,049	3,996,049	3,996,049	3,996,049
Office x quarter	yes	yes	yes	yes	yes	yes	yes
Case covariates	yes	yes	yes	yes	yes	yes	yes
R-squared	0.080	0.073	0.108	0.109	0.095	0.004	0.011
MeanY: complier not allowed	7,464	0.338	0.311	0.168	0.63	499	0.039
MeanY: 6-10 yr pre application	33,420	0.857	0.838	0.394	1.63	1249	0.062

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

All earnings are annual and in 2023 dollars include those with zero earnings.

Plausibility of SSA policy reforms as mechanism – three pieces of evidence

(1) Reforms account for 28-36% of the decline in allowance rates

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

Category	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

First stage estimates, adding policy variables

Our model explains 6.2-8.0pp of the 22 pp decline in allowance rates

Or 28-36% of the decline

Note this is a cumulative decline in allowance rates with the sequential addition of controls.

(2) Fully one-third of appellants had cases impacted by reforms (complier share)

- 33.7% are compliers
- 28.8% are always takers (allowed under any ALJ)
- 37.5% are never takers (denied under any ALJ)

We calculate the compliers share by generating a *single best instrument*. We residualize each instrument by regressing them on all first stage controls. Then we regress the allowance rate on the vector of residualized instruments; the prediction from this model is the single best instrument. Once we have this, we use the usual approach of taking the difference in the predicted allowance rate at the max and min of the instrument.

(3) SSA reforms screened out those with (relatively) higher work capacities

Untreated (denied) complier means (2023$)

Earnings 4 yrs post decision	Judge IV	Policy Reform IV
Earnings	$5,469	$7,464
Earnings > $0	0.265	0.338
Earnings > SGA	0.118	0.168

- Compare the untreated complier means across the **policy reform** IV and **judge** IV
- Reforms targeted people with relatively more work capacity on the margin
- Thus, reforms could plausibly be an explanation for the decline in work capacity

Conditions for LATE interpretation (not shown)

In terms of threats to interpreting these findings, we show that conditions of LATE interpretation are met:

- Relevance (shown in F statistics)
- Consistent with conditional random assignment, instruments satisfy **balance tests**
- The policy reforms **do not lead to a change in applications or appeals**; ruling out this as a confounder
- Monotonicity violation in upper range of instrument; can't estimate MTE

Conclusion

- We document recent declines in appellate allowance rates in SSDI

- Using a judge IV approach, we show that SSDI leads to reductions in labor supply yet the underlying work capacity of compliers is very low

- As eligibility in the program falls, the induced change in labor supply declines reflecting lower work capacity on the margin

- We uncover and document recent reforms to the appellate process that are possibly the most important (and least heralded) SSDI reforms of our time

- We show that turnover among judges and the implementation of policy-compliant training led to significant declines in the allowance rate, explaining 28-36 percent of the 22 percentage point decline in allowance rates between 2007 and 2015

- We conclude that the tightening of eligibility has moved the threshold further below SGA (the intended threshold in the program)

EXTRA SLIDES

MORE DETAIL ON SSA REFORMS

With the suspension of ALJ hiring, the appellate backlog increased:

Appellate Case backlog
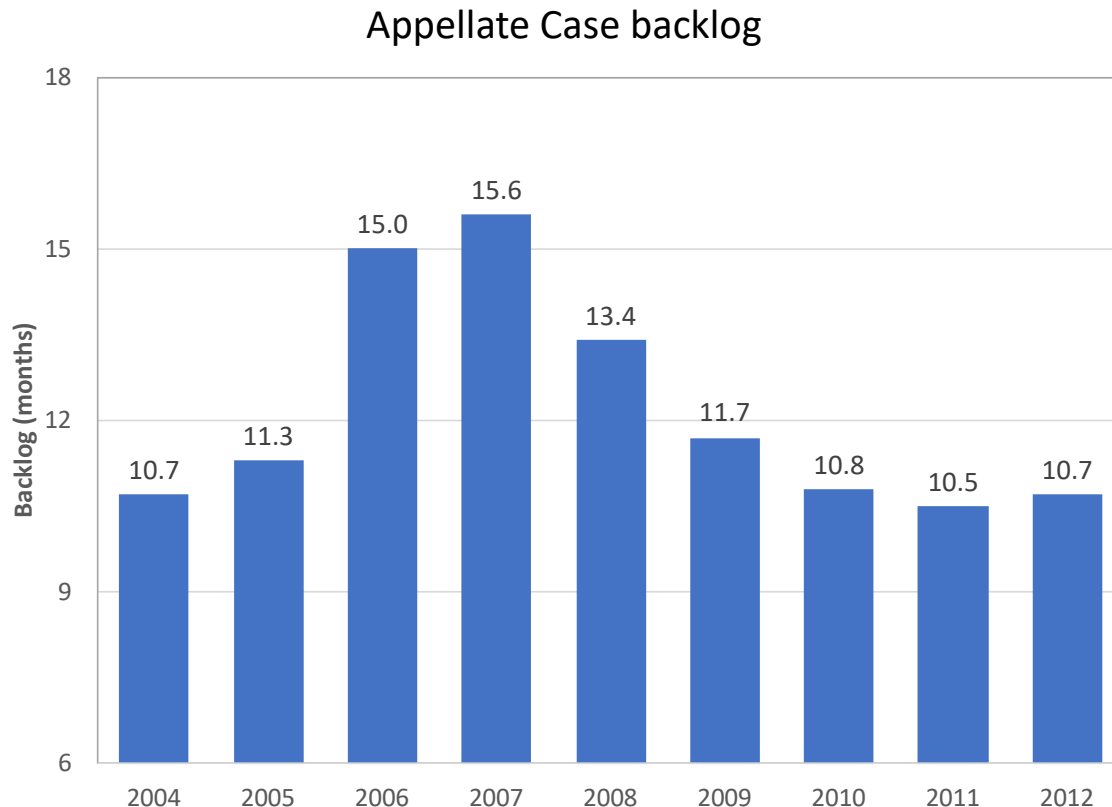
Backlog defined as number of pending hearings divided by the average hearings per month.

It captures the number of months it would take to eliminate the backlog if counterfactually there were no further applications.

Source: Annual Statistical Supplement to the Social Security Bulletin, various years

Efforts to make decisions more consistent

THE WALL STREET JOURNAL.

U.S.

Three Indicted for Alleged Social Security Fraud Scheme in Kentucky

Government disbursed benefits in excess of $600 million as result of scheme, indictment says

Attorney Eric Conn during a Senate committee hearing on Capitol Hill in 2013.
He has been indicted by a federal grand jury in Lexington, Ky., on charges
including conspiracy to commit mail fraud and wire fraud.

Source: 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.

- **May 2011** WSJ reported an ALJ with a 100% allowance rate, all from one attorney

- **July 2011** hearing by House Ways and Means

- Before this, SSA began using analytics to assess ALJ policy compliance (Ray & Lubbers 2015)

- Identified systematic decision errors, outlier judges, too many OTR decisions

Research Design & data for mechanisms – SSA Reforms

Causal model: effect on SSDI allowance on labor supply

$$Y_{idt} = \alpha + D_{j(i)}\beta + X_{idt}\eta + X_{j(i)}^{A}\lambda + \theta_{dt} + \epsilon_{ijdt}$$

- Same model – estimate impacts of SSDI allowance on future labor supply outcomes
- Use SSA policy reforms as instruments for allowance
- As with judge propensity IV, this is valid due to random assignment of judges to cases at appellate level.

SSA Policy Reform: other ways to promote policy compliance

- Introduced "How MI Doing" (nationwide) in 2011 – App for judges to compare their decision-making with others; included feedback mechanisms, training modules, and hyperlinks to detailed agency policy information

- Scaled-down the SAA program (reassigned SAAs to decision writing)

Data using Policy Reform

- Created judge level reform data base from Office of Appellate Operations
  - dates of hire, for new hires FY2008 and later
  - dates of training
  - dates of focused reviews of each judge (if received)
- Merge this with our SSA Case Processing and Management System (CPMS) data on all SSDI decisions made at the appellate level between 2007 and 2015; merged to Master Earnings File.

Summary statistics – judges and policy variables

	2007	2011	2015
Number DI Decisions per Judge, Mean	157	160	109
Allowance Rate	0.660	0.655	0.466
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
Number	1,106	2,010	1,552

MORE DETAIL ON ESTIMATES OF SSA REFORMS

Summary of First Stage

- Hiring of new judges and rollout of policy compliant changes led to large reductions in allowance rates

- No evidence of spillover effects of reforms on other judges (e.g., focused reviews)

- Robust to inclusion of labor market variables, case characteristics

IV Effects on Earnings

How does DI allowance affect earnings, using DI reforms as instruments?

IV Estimates of Allowance on Earnings 4 Years Later (2014$), By Subgroup

	Appellate Allowance		First Stage F Stat	Mean of Y compliers untreated	Mean of Y 6-10 yr pre appl	Number of Observations
Full Sample	-4,134***	(259)	15.4	$5,552	$23,348	4,121,375
Musculoskeletal	-4,267***	(354)	14.1	$5,795	$25,421	1,789,621
Mental	-4,576***	(423)	13.2	$6,004	$18,603	791,025
Cardiovascular	-3,247***	(708)	9.9	$4,416	$26,773	288,511
Neurological	-3,402***	(719)	11.5	$4,991	$24,536	274,202
Endocrine	-2,729***	(856)	7.3	$4,350	$23,299	179,147
Respiratory	-2,959***	(950)	8.2	$4,398	$22,405	135,202
Female	-3,528***	(263)	15.2	$4,898	$20,224	2,142,367
Male	-4,917***	(399)	13.9	$6,394	$26,731	1,979,008
Prior Earn Q1	-3,480***	(352)	15.4	$4,685	$3,166	1,024,901
Prior Earn Q2	-3,302***	(341)	14.4	$4,754	$12,269	1,032,158
Prior Earn Q3	-4,250***	(404)	13.7	$5,736	$24,288	1,032,158
Prior Earn Q4	-5,622***	(596)	10.7	$7,152	$52,488	1,014,405
Age 20-29	-4,586***	(724)	13.9	$7,123	$5,585	325,421
Age 30-29	-5,332***	(448)	16.5	$7,352	$17,752	757,276
Age 40-49	-4,044***	(335)	16.6	$5,134	$23,922	1,516,125
Age 50-54	-3,350***	(461)	10.5	$4,354	$27,331	825,402
Age 55+	-1,598***	(536)	4.7	$2,720	$31,771	696,677

Subgroup estimates

Larger effects for:
Men
Mental condition
Higher prior earnings
Younger workers

Reforms (= lower allowance rates) led to larger increases in earnings for these groups.

Very low mean earnings untreated compliers despite prior earnings (low work capacity)

Threats to validity of approach – SSA REFORMS IV

Balance Test

	Years of Education
Judge Cohort (omitted = incumbent)	
FY2008	0.0204
	(0.0312)
FY2009	0.0105
	(0.0196)
FY2010	-0.0137
	(0.0167)
FY2011	0.0276
	(0.0173)
FY2012	-0.000193
	(0.0184)
FY2013	0.0141
	(0.0195)
FY2014	-0.0264
	(0.0394)
FY2015	-0.0114
	(0.0245)
Post Policy Compliant Training	
Incumbent, post PCT	0.0213*
	(0.0124)
FY2008, post PCT	0.00728
	(0.0293)
FY2009, post PCT	-0.0125
	(0.0179)
FY2010, post PCT	0.0370**
	(0.0169)
Post Own Focused Review	0.0161
	(0.0276)
Observations	4,003,408
R-squared	0.048
Mean of the dependent variable	12.02

Notes: All models control for personal and case characteristics, body system major class, state unemployment rate, unemployment rate by body system class, and hearing office. SE clustered on judge.

Do the instruments predict individual "hold out" characteristics (not included in first stage)?
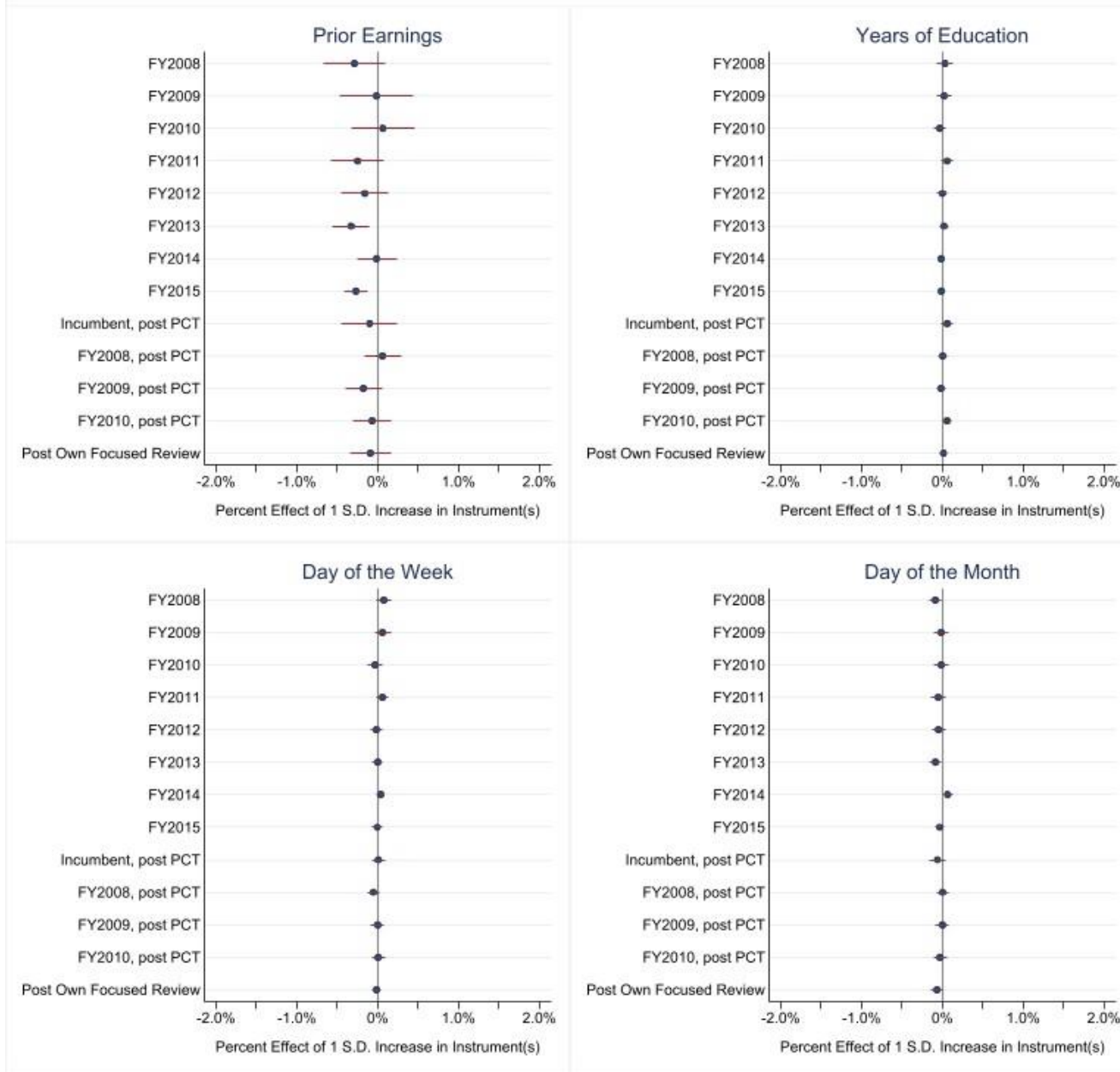
1 of the 13 is statistically significant

We estimate the same model as the first stage, but with hold out covariate as the outcome variable.

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

Balance Test

Notes: Prior earnings, average of years 6-10 prior to filing, 2014 dollars.

We repeat this balance test for other hold out variables

<u>The instruments have no impact (statistically or economically) on:</u>
Earnings prior to filing
Years of Education
Case filing variables

Important evidence for validity of instruments

To put the outcomes on a similar scale we simulate the effect of a 1 SD change in Z on the outcome, and express it as a percent of the baseline mean of the outcome.

Leads and Lags of Log of Claims Regressed on Instruments (and model controls)

	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 Revie	-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

Instruments have no impact (statistically or economically) on SSDI applications
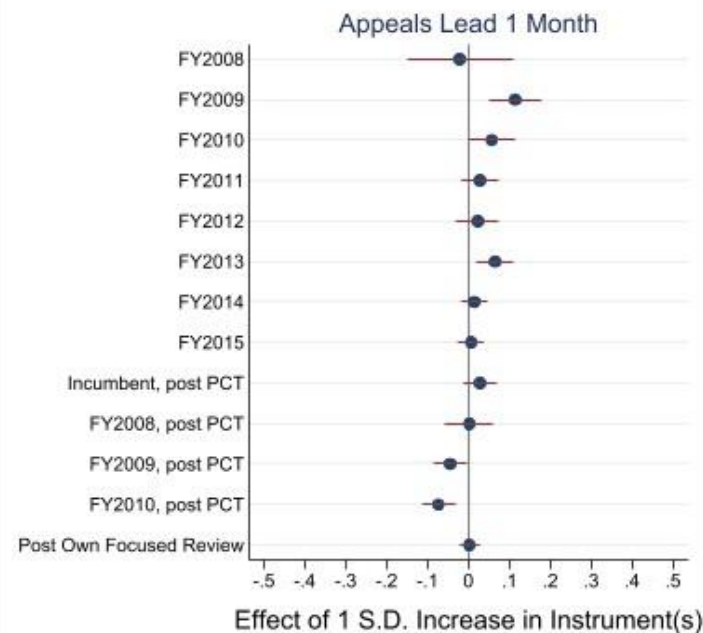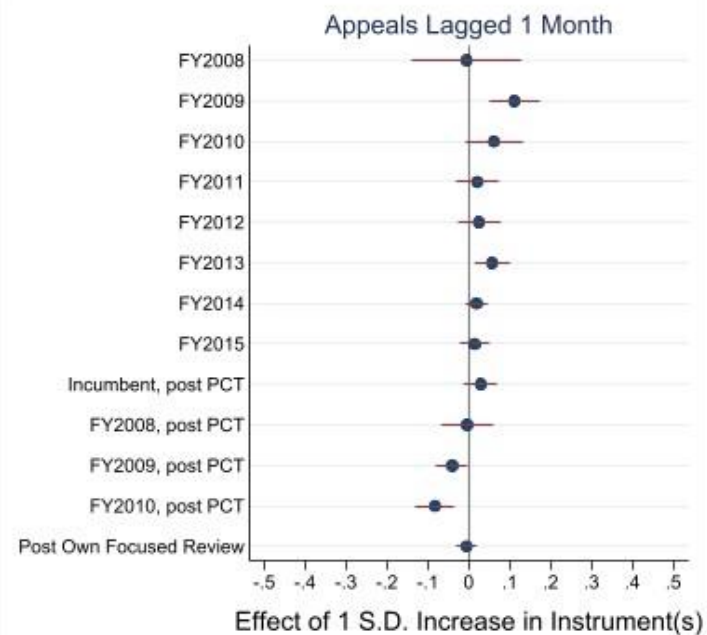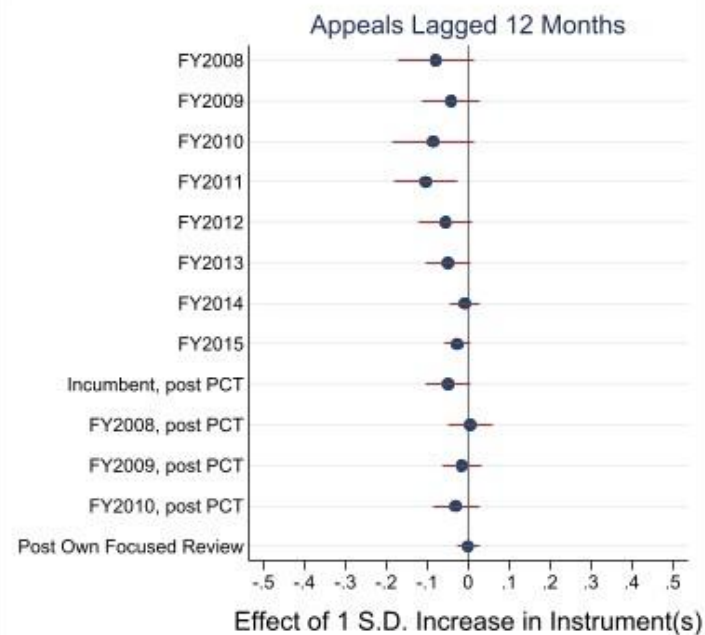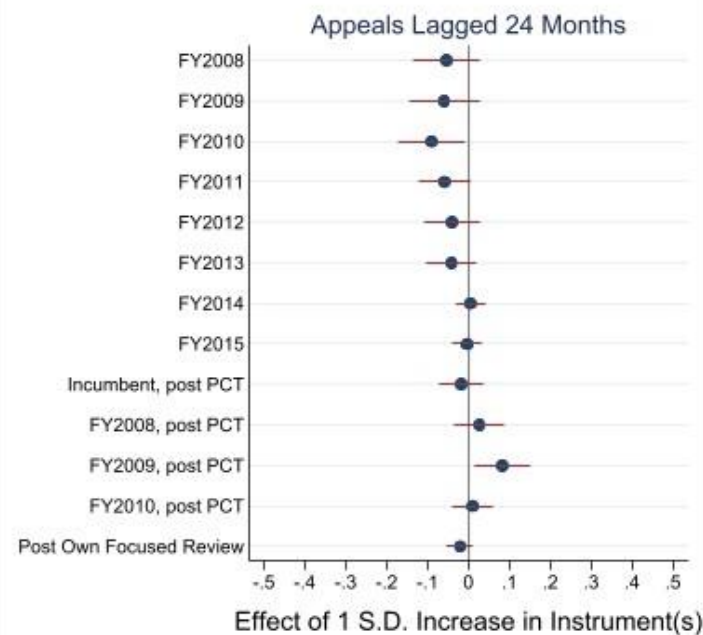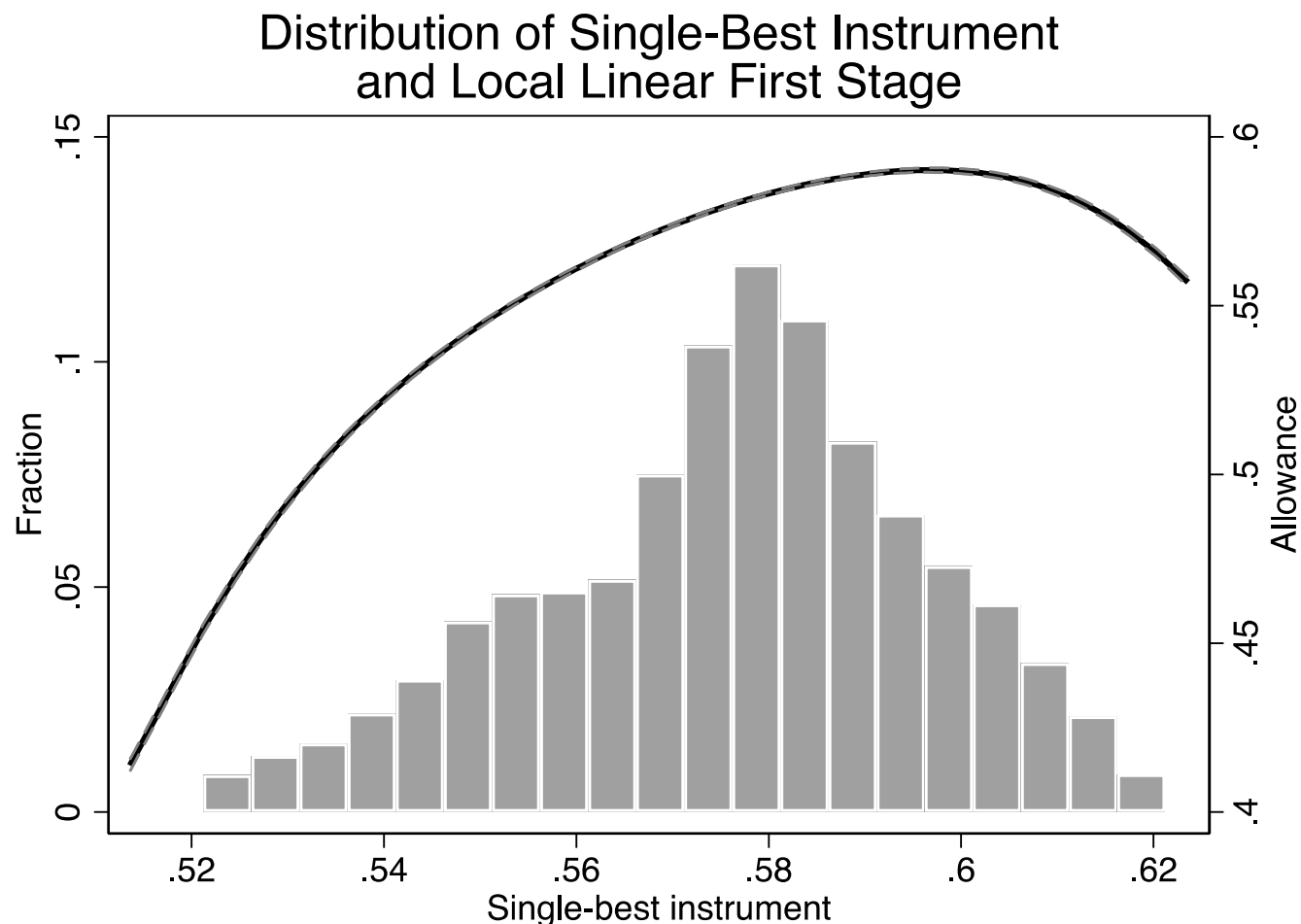
These estimates are from a regression of the log of office-month counts of DI applications, with controls (local labor market vars, means of individual + case char., office and month-yr FE).

Log Claims Regressed on Instruments

Claims Lagged 24 Months

FY2008	
FY2009	
FY2010	
FY2011	
FY2012	
FY2013	
FY2014	
FY2015	
Incumbent, post PCT	
FY2008, post PCT	
FY2009, post PCT	
FY2010, post PCT	
Post Own Focused Review	

Effect of 1 S.D. Increase in Instrument(s)

Claims Lagged 12 Months

Effect of 1 S.D. Increase in Instrument(s)

Claims Lagged 1 Month

Effect of 1 S.D. Increase in Instrument(s)

Claims Lead 1 Month

Effect of 1 S.D. Increase in Instrument(s)

Claims Lead 12 Months

Effect of 1 S.D. Increase in Instrument(s)

Claims Lead 24 Months

Effect of 1 S.D. Increase in Instrument(s)

Log Appeals Regressed on Instruments

Monotonicity

Distribution of Single-Best Instrument and Local Linear First Stage
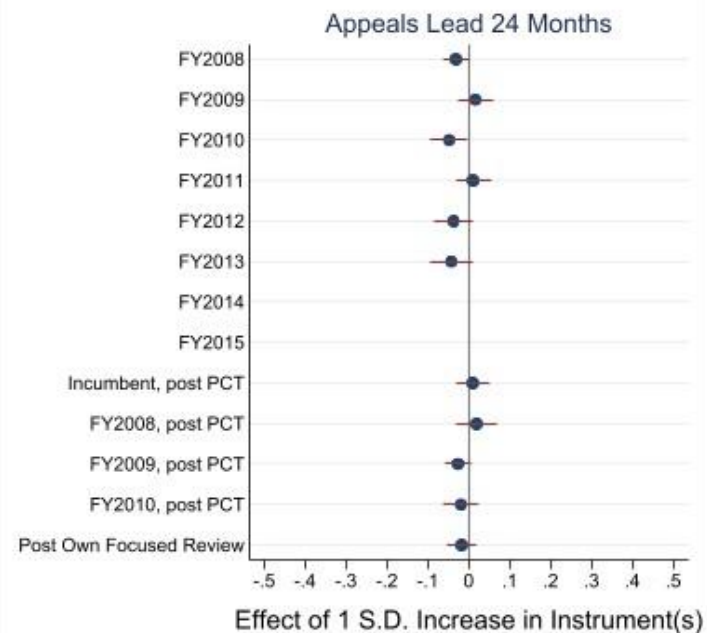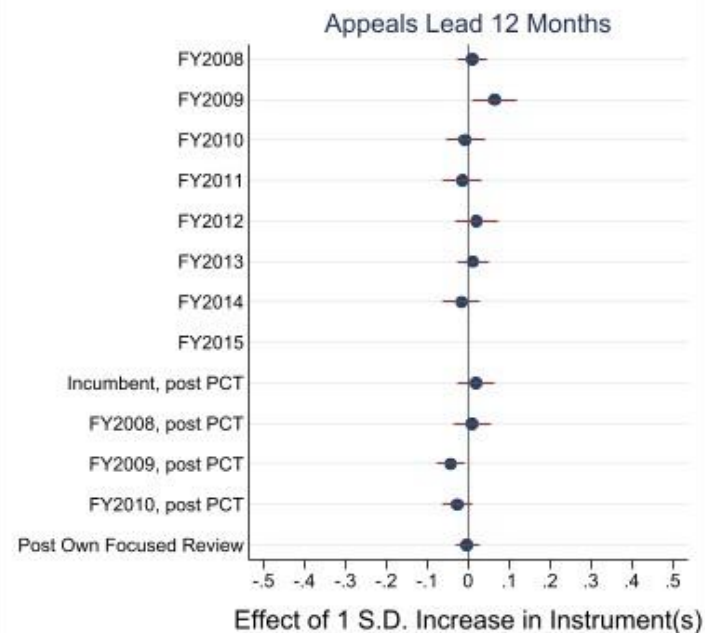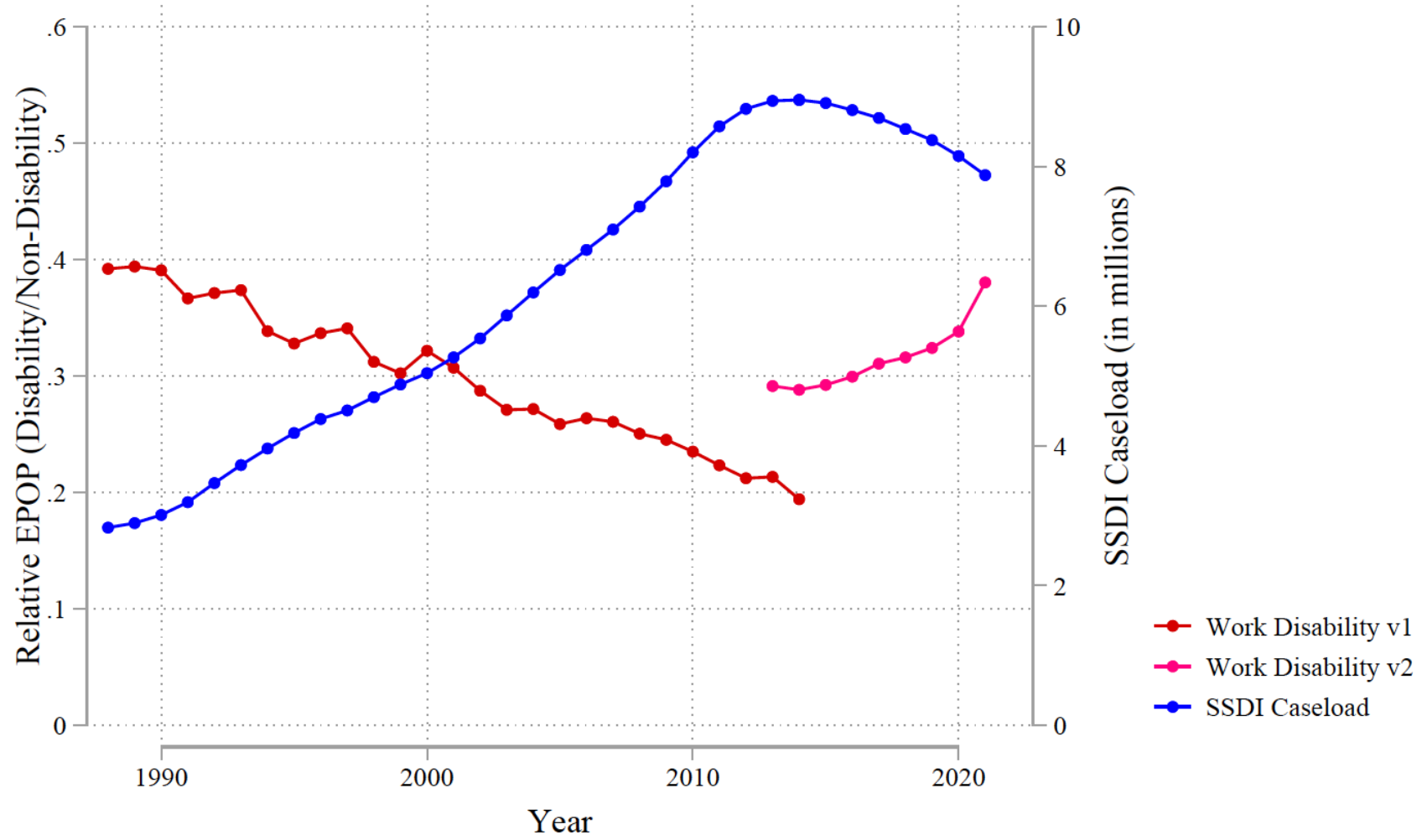

- Because multiple IVs, construct "single best" IV

- Monotonicity violation in upper range (when predicted allowance high)

Implications for Trends in Work among Persons with Disabilities

SSDI Caseload and Relative Work Disability/No Work Disability Employment

A rough back of envelope calculation suggests the effect is small, explaining about 0.25 percentage points out of the (approximately) 5 percentage point increase in employment among persons with disabilities during our time frame (about 4 percent).