Workplace Disruptions, Judge Caseloads, and Judge Decisions: Evidence from SSA Judicial Corps Retirements^{*}

Clayson Shumway Riley Wilson[†]

December 2, 2021

Click here for latest version

Abstract

We exploit judge retirements from the Social Security Disability Insurance judicial corps to document how remaining judges respond to workplace disruptions. When a peer judge retires, the remaining judges see a 5 percent increase in dispositions and decisions that lasts 6 months. Institutional features of the disability appeal process allow us to estimate what happens to judge decisions when caseloads increase, holding the composition of cases fixed. Increased caseloads are accompanied by a 1 percent decrease in the judges' share of favorable decisions, suggesting 16,600 claimants in-sample were not awarded disability insurance who would have been, absent the workplace disruption.

Keywords: Social Security Disability, Judges, Peer Effects, Caseloads JEL Codes: H55, J14, K23

^{*}Thanks to Manasi Deshpande, Emily Leslie, Yue Li, Nicole Maestas, Kyung Park, Cody Tuttle, and WEAI International participants for helpful comments and discussion. Declarations of interest: none.

[†]Corresponding Author: Brigham Young University, Department of Economics, 2146 West View Building, Provo UT, 84602. Email: riley_wilson@byu.edu.

1 Introduction

Do judges' decisions change following a disruption in the work environment? Does workload affect judge decision-making? Researchers have explored how individual and judge characteristics affect eventual case rulings, but changes in the judge's work setting might also impact their decisions. As such, two otherwise identical individuals could receive different rulings from the same judge.

The impact of judge caseloads is of particular interest, as it highlights the role resource constraints might play in judge decisions. Judges facing high caseloads might approach cases differently, leading to inconsistent judgements. The potential impact of heavy caseloads on judge behavior is not a new concept (Nardulli, 1979; Robel, 1990). However, as Yang (2016) shows, increased judge caseloads in the criminal justice system induce prosecutors to change the number and composition of cases brought before judges, making it difficult to identify the impact of judge caseloads on judge decisions. By examining Social Security Administration Disability Insurance (SSDI) appeals, where cases are quasi-randomly assigned to judges and typically in queue over one year in advance, we are able to isolate changes in judge decisionmaking when caseloads increase but the composition of cases does not change. In this setting, we can test if judges facing heavier workloads become more strict or lenient.

Using publicly available administrative Social Security Administration (SSA) data on monthly Administrative Law Judge (ALJ) cases and decisions from 2011 to 2019 in an event study framework, we examine how peer judge retirements affect caseloads and allowance rates (share of decisions favorable) in SSDI appeal cases of remaining judges in the same office. When a fellow judge retires, cases are reallocated across the remaining judges leading to a sharp, 5 percent increase in judge-level decisions that remains elevated for 6 months. Given this increased caseload and pressure to keep applicant wait times short, we explore impacts on judges' allowance rates. After the retirement, within-judge allowance rates fall by one percent, suggesting the workplace disruption impacted case outcomes. These estimates are robust to various controls and alternative approaches that address twoway fixed effect event study concerns (Goodman-Bacon, 2021; Sun and Abraham, 2020). Judges that experienced the largest increase in dispositions after the peer judge retirement also exhibit the largest reduction in allowance rates, consistent with increased caseloads affecting judge decisions. Female judges and less-experienced judges saw the largest caseload increase and exhibit larger declines in allowance rates, providing further suggestive evidence that judge decisions respond to workload changes. Alternative mechanisms, such as retiring judge peer effects, are less supported in the data.

Workplace changes that increase workload and reduce allowance rates have important implications for the scope of the disability insurance program. Back of the envelope calculations suggest a one percent average reduction in ALJ allowance rates from in-sample retirements led to 16,600 claimants not receiving as much as \$246 million in annual benefits when they would have been deemed eligible under different judge workplace circumstances. Given the absorbing nature of disability insurance, this has long-run consequences for both individuals and the program. These patterns suggest that caseload increases affect judgements and that policies that mitigate office-level disruptions would result in more consistent rulings.

2 Work Environment, Workload, and Judge Decision

Existing work shows that claimant characteristics and judges' own characteristics affect court decisions.¹ A growing literature suggests that judges' rulings are also influenced by time-varying environmental factors such as temperature, upsetting sports loses, media coverage, or judge peer composition (Danziger et al., 2011; Eren and Mocan, 2018, 2020; Heyes and Saberian, 2019; Lim et al., 2015). As such, two individuals with similar characteristics could

¹For example, Alesina and Ferrara (2014); Mustard (2001); Park (2017) and Rehavi and Starr (2014) document the role of defendant characteristics, while Depew et al. (2017) and Boyd et al. (2010) document the role of judge characteristics. Lim et al. (2016) find little evidence that judges' gender, ethnicity or political affiliation affect case outcomes.

receive different rulings from the same judge, because of the work environment. Unlike some of the previous examples, we are interested in understanding how policy-relevant features, such as judge workload and workforce dynamics, influence judge behavior, which could ultimately affect applicants' well-being and outcomes. From a policy perspective, understanding how factors like workload affect applicant outcomes can shed light on the importance of program structure, financing, and management.

Work across criminology, law, and economics find in other judge settings that caseloads affect case outcomes, the number of cases heard, and the likelihood of overruling appeals (Beenstock and Haitovksy, 2004; Engel and Weinshall, 2020; Huang, 2011; Iverson, 2018; Ulmer and Johnson, 2004). However, in many judicial settings, it is not possible to isolate the impact of caseloads on judge decisions because the number and composition of cases might endogenously respond. As Yang (2016) documents in the criminal justice system, heavy judge caseloads due to judge vacancies induce prosecutorial changes in the number and types of cases brought before judges. As such, any observed change in judge decision could reflect changes in the composition of cases in addition to judge responses to workload. Isolating the impact of workload on judge decisions is not possible if the composition of cases changes. We present the first work we are aware of that explores the impact of caseloads on judge decisions in a setting where the composition of cases is predetermined, and the only work that documents how work environment factors affect the SSA judicial corps, a body of approximately 1,500 judges making over 500,000 decisions a year.

SSA is consistently trying to reduce hearing backlogs and wait times. A judge retirement here might be particularly disruptive as it increases remaining judges' pending cases, potentially intensifying pressure to reduce the backlog. This added burden and stress could lead judges to make more errors and reject cases they would have otherwise allowed or allow cases they would have rejected. In other settings, negative environmental conditions lead judges to be more harsh, punitive, and less favorable to the defendant (Danziger et al., 2011; Eren and Mocan, 2018; Heyes and Saberian, 2019). However, judges could also become more lenient, by avoiding time-consuming denials that risk appeal from a higher level. Any net effect on allowance rates is an empirical question.

3 Social Security Disability Application and Appeal Process

SSA administers disability insurance (SSDI) and supplemental security insurance (SSI) to insure against the risk of becoming unable to work due to disability. SSDI is for workers with a sufficient work history and evidence of a permanent disability that inhibits them from participating in substantial gainful activity. SSI does not have the same work history requirement and is available to the elderly, blind, and low-income, low-asset disabled. About 2.5 million workers apply for SSDI, SSI, or both each year. Applications are submitted to local field offices, which verify non-medical requirements (e.g., work history, income limits), then send the application to a state-run Disability Determination Services (DDS) agency that evaluates medical criteria to either allow or deny the applicant. If a denied applicant appeals the decision, it goes to Reconsideration, where the review process is repeated by new field office and DDS examiners. Individuals denied at Reconsideration can appeal to present their case and any new evidence before an Administrative Law Judge (ALJ). In 2010, 19 percent of all applicants sought an appeal hearing and nearly one-third of all disability awardees were approved through the hearing process (Social Security Administration, 2019).²

A hearing request is assigned to one of 166 hearing offices based on the applicant's zip code of residence. Within a hearing office, cases are quasi-randomly assigned to one of the ALJs on a rotational basis, with the earliest request being sent to the judge that most recently completed a case.³ ALJs do not specialize in particular types of cases. If an ALJ retires, their pending cases are reassigned to the remaining judges so that cases are heard in a timely manner. If the ALJ denies the claim, the individual can appeal to the Social

²Approximately 51 percent of applicants that reach DDS review simultaneously apply for SSDI and SSI. Joint applicants are less likely to be allowed at all stages.

³There are a few exceptions to the rotational assignment. Critical cases (e.g., Terminal Illness and Veteran cases) and Court Remand Cases sent back from the Appeal Council are prioritized in the queue.

Security's Appeal Council, who will either deny, approve, or send the case back to an ALJ.

Across hearing offices and over time there are between 3 and 20 ALJs per office, with an average of 11 in our sample (see Figure A1 for the distribution of the number of judges at each retirement event). Between 2010 and 2019 there were approximately 1,505 judges at any point in time, averaging 38.3 dispositions and 31.2 decisions a month, with a favorable decision 56 percent of the time (Table 1).

A large literature shows that SSDI receipt reduces applicant labor supply (Autor and Duggan, 2003; Autor et al., 2016; French and Song, 2014; Gelber et al., 2017; Maestas et al., 2013; Moore, 2014), but also reduces mortality (Gelber et al., 2018) and financial distress (Deshpande et al., 2020). Work exploring ALJs in the SSA is limited to research exploiting judge fixed effects to estimate causal effects of SSDI receipt (French and Song, 2014) or exploiting regional variation in appeal wait times (Kearney et al., 2021).⁴

Similar to other judges' caseloads, SSA expects full-year judges to complete 500 dispositions a year, but no more than 700.⁵ During our analysis period, 54 percent of judges made over 500 dispositions while 14 percent made over 600, with rare cases when judges made over 1,000 dispositions in a year. The SSA appeal case backlong in long. System-wide pending cases increased from 700,000 in 2010 to over 1.1 million in 2016 before dropping off.⁶ This backlog leads to long appeal wait times of 14 months on average.

4 Data

Our data consist of monthly reports for each judge/hearing office combination which we scraped from the administrative SSA ALJ Disposition Data records.⁷ We observe each

 $^{^4\}mathrm{At}$ 2018 NBER Summer Institute, Nicole Maestas noted preliminary work on ALJ training, but this is not released.

⁵For example, appeal judges hear approximately 270 cases a year, while immigration judges hear approximately 895 (https://www.justice.gov/eoir/file/1198896/download).

⁶Pending cases dropped due to a drop in applications and appeals after the Great Recession as well as ALJ reforms aimed at reducing wait times.

 $^{^7\}mathrm{Some}$ judges split time between multiple hearing offices so they may have multiple observations for the same month.

judge's name and hearing office as well as the number of dispositions, decisions, allowances, denials, and partially- or fully-favorable allowances made during the month. Dispositions capture all cases the judge hears, including dismissals on technical grounds or remands from higher courts.⁸ Decisions only include cases where a final determination is made. Measures are reported cumulatively through the administrative year (starting in October), which we use to back-out monthly cases. Our data covers all judges that appear in the data between September 2010 (the first month ALJ disposition records were provided) and July 2019.

From these data we construct the monthly Hearing Office judge roster and identify retiring judges. We define a retirement as the month a judge leaves a hearing office and does not reappear in the data at any hearing office. Individuals that move to a different hearing office or experience an extended leave before reappearing in the disposition data are not counted as retirements. We do not exploit these office-to-office moves because the timing is more likely to be correlated with case characteristics that could impact peer judges' caseloads and decisions.⁹ We exclude retirement events before April 2011 and after August 2018 (to insure adequate pre- and post-observation) as well as retirements at national hearing centers, which lack a geographic catchment but service cases from hearing offices around the country. We are left with 753 retirement events across 158 different hearing offices. These retirements are fairly uniformly distributed over time (Figure A1). retiring judges' allowance rates are nearly identical to the those of all other judges during the same time period (Table 1).

ALJs that are dismissed by SSA or reprimanded and then leave would also be captured in these retirements. This is potentially problematic if judge dismissal is due to workplace conduct that influences peer judges' caseloads or decisions. By law, SSA cannot take disciplinary action that would infringe on a judge's ability to make independent decisions, and can only suspend or remove a judge after the Merit Systems Protection Board (MSPB) determines

⁸Dismissal at the ALJ level is due to technical reasons (e.g., failure to file appropriate documents or attend hearings) not the claimant's case merit. From 2007 to 2015, 13-18 percent of dispositions were dismissal (Government Accountability Office, 2017).

⁹For example, SSA might move a judge from one hearing office to another if the number of appeals is falling or rising, which might be accompanied by a change in case composition.

there is "good cause" (Office of the Inspector General, 2014). A judge's case decisions alone do not qualify as "good cause" (Office of the Inspector General, 2012). However, SSA can engage in targeted training if a judge's allowance rate is inconsistent with other judges in the office. In 2010 and 2011, SSA undertook a comprehensive review of ALJ decisions to standardize judge procedure, including training for judges with anomalous allowance rates for their hearing office. Disciplinary action against ALJs for any reason that result in retirement or removal is uncommon. Between 2007 and 2016, there were 98 reprimands,¹⁰ 34 proposed suspensions, and 16 proposed removals (Government Accountability Office, 2017). Eight of these judges were removed by the MSPB, 6 retired before the MSPB issued a decision, and 2 retired after a settlement. We collect all SSA case reports from the MSPB from 2010 to present and identify ALJs that had a case brought before the MSPB during our sample window. Results are similar if we exclude these judges.

The only individual information we observe about each judge is their name. We link each judge's first name to the SSA's record of the genderedness of names to probabilistically determine judge gender. Using the judge's work history, we also identify judges that are new to the SSA judicial corps (appear in the data less than one year before the focal retirement event). One major drawback of the data is the aggregate nature. Because we do not observe case-level measures, we cannot test whether case characteristics change after the retirement event or explore which types of cases judges might treat differently. As such, we must rely on the institutional quasi-randomization and can only provide limited evidence of the welfare effects of these retirement events.

We structure the data to match the event study framework. First, we identify the last month a judge appears in the SSA judicial corps and flag this as a retirement event at the hearing office-level. We then construct a panel with 6 months prior to the retirement, the month of the retirement, and 9 months after the retirement for each judge in the retiree's hearing office. We limit the panel to this size to avoid overlap between retirements in the

¹⁰A formal written warning that stays on the judge's record for one year.

same hearing office.¹¹ Only judges present the entire panel are included. We then stack the panels for each judge to construct our analysis sample. Judges present for multiple retirements will have multiple stacked balanced panels. There are 5,634 judges with full balanced panels, leading to a sample of 90,144 judge-by-month observations.

5 Empirical Strategy

As noted above, appeals are quasi-randomly assigned to ALJs, ALJs do not specialize in particular types of cases,¹² and the enormous backlog in appeal cases means cases take 14 months on average to process. As such, the number of cases an ALJ hears might change after a workplace disruption, like a peer retirement, but the composition of cases should remain the same. We can explore the impact of judge vacancies and caseloads on judge decisions, which typically cannot be separately identified since the composition of cases changes (Yang, 2016).

We estimate the impact of the retirement on cases and allowance rates (share of favorable decisions) using the stacked, balanced panel as follows

$$Y_{jot} = \sum_{\tau=-5}^{9} \beta_{\tau} (\tau \text{ months from retirement})_{ot} + \delta_j + \phi_o + \gamma_{yr} + \psi_{mo} + \varepsilon_{jot}$$
(1)

Our main outcomes of interest are the monthly dispositions, decisions, awards, and allowance rate for judge j in hearing office o in the year-month period t. The allowance rate is the number of awards divided by the total number of decisions.¹³ We examine both levels and natural log specifications. Our preferred specification is the natural log transformation as there is substantial variation across judges in monthly dispositions, decisions, and allowance

¹¹There are still some overlapping panels as multiple retirements might occur within a few months. The results are similar if we eliminate the overlap by excluding retirement events within 9 months of each other (Table A1).

¹²Because we only have aggregate judge-level data we cannot test this directly.

¹³We focus on allowance to decision rates throughout to match SSA definitions, but also to avoid potential compositional changes that could arise if appeals are pre-screened and dismissed for administrative error prior to being passed from the retiree to the peer judge.

rates and this allows us to estimate percent changes in outcomes.¹⁴ We include judge fixed effects to control for time-invariant judge characteristics. Hearing office fixed effects make this a within office comparison over time, and calendar year and month fixed effects control for potential secular and seasonal trends in outcomes. Our coefficients of interest are the β_{τ} , which trace out a judge's monthly outcomes six months prior to the peer judge's retirement and nine months after. The first month in the event panel (six months prior to the retirement) is treated as the omitted group, so all monthly effects are relative to that month. The pattern is similar if the month prior to the retirement is omitted (Figure A3). This framework allows us to explore pre-period trends as well as treatment effect dynamics for nine months. Standard errors are corrected for clustering at the hearing office-level. To avoid sample composition changes, we only include judges who appear in the hearing office for all 16 months of the panel. Judges that join the hearing office during the panel to replace the retiree are not included.¹⁵

As seen in Table 1, judges in the event study panel have similar disposition counts, decision counts, allowance rates, and gender composition prior to the retirement event as the full sample. After the retirement, these judges have slightly higher dispositions and decisions, and allowance rates that are one percentage point lower. In the six months leading up to retirement, retiring judges are also similar to the full sample of judge/month observations and to the event study panel of judges, suggesting they are not uniquely selected. They have similar allowance rates, but they do have about 6 fewer dispositions a month, 4 fewer decisions a month, and are less likely to be women.

Only event panels for judges around a retirement event are included, so unlike a differencein-differences event study there is no "control" group of judges or hearing offices that do not experience a retirement. The identifying assumption is that after accounting for judge-

¹⁴In the balanced panel there are 689 judge-month observations where judges have zero dispositions, even though they are still assigned to the hearing office. This is why the natural log outcome samples are slightly smaller. Results are similar if we use the inverse hyperbolic sine transformation (Figure A2).

¹⁵Average effects are similar if we relax this restriction and include any judge who holds at least one disposition both before and after the retirement event (Table A1).

specific, year-specific, and month-specific effects, dispositions, decisions, and allowance rates within a given hearing office would not have deviated if the retirement had not occurred. The six month pre-retirement period is crucial to understanding the plausibility of the identifying assumption. As seen throughout, pre-trends are mostly flat, with stark, discontinuous jumps in outcomes once the retirement occurs suggesting the estimated effects are not due to some underlying, secular trend. We also verify robustness to recently highlighted challenges associated with twoway fixed effect event study models with staggered treatment timing (Borusyak and Jaravel, 2020; Callaway and Sant'Anna, 2020; de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2020). Our estimates are insensitive to pre-trend tests (Borusyak and Jaravel, 2020), alternative event study estimators (Sun and Abraham, 2020), and event study models using a stacked panel with judges in hearing offices that will experience a retirement in the future as a control group (similar to Deshpande and Li (2019)), as suggested by Goodman-Bacon (2021).

We also estimate the following parameterized difference models:

$$Y_{jot} = \beta_0 Month of Colleague Retirement_{ot} + \beta_1 After Colleague Retirement_{ot} + \delta_j + \phi_o + \gamma_{yr} + \psi_{mo} + \varepsilon_{jot}$$
(2)

This allows us to identify the average effect of peer judge retirements on remaining judges' outcomes after the event has occurred. We do not observe when during the month a judge retires, only the last month they hold any dispositions. For this reason we include two separate time periods, the exact month of the retirement and the remaining nine months after the retirement. Because there are no pre-period indicators, β_0 and β_1 are the effects relative to the pre-period average, after controlling for judge, office, year, and month fixed effects. As exact retirement timing is unknown, β_0 is hard to interpret, so we are mostly interested in β_1 which is the average impact of a peer retirement on workload and allowance rates over the next nine months.

The impact of a peer-judge retirement could impact judges heterogeneously. In the appendix we report estimates that allow the effects in equation (2) to vary with the judge's characteristics, the retiree's characteristics, or other workplace conditions, but we find limited evidence that retiring judge characteristics affect the changes in allowance rates (see Appendix B for details on supplemental estimation and results).

6 Results

6.1 Change in the Number of Judges

We first document that a judge retirement leads to a change in the number of available judges in the hearing office. To do this we collapse the data to the hearing office level and count the number of judges that hold dispositions in the hearing office every month of the event panel, including the retiring judge and any new replacement judges. We then estimate

Number of
$$Judges_{ot} = \sum_{\tau=-6}^{9} \beta_{\tau} (\tau \text{ months from retirement})_{ot} + \phi_o + \gamma_{yr} + \psi_{mo} + \varepsilon_{ot}.$$
 (3)

As before, hearing office, year, and month fixed effects are included and standard errors are corrected for clustering at the hearing office-level.

The vector of β_{τ} coefficients from equation (3) is plotted in Figure 1 with 95 percent confidence intervals. The number of judges does not significantly vary during the six months prior to the retirement. The number of judges drops significantly by 0.84 judges the first month after the retirement, but slowly increases until reaching 0.5 judges lower nine months after the retirement.¹⁶ This would suggest that judges are immediately replaced in 16 percent of the retirement events, but that about half of the time they still have not been replaced 9 months after the retirement.¹⁷

At first it might seem surprising that retirees are not replaced more quickly, especially since these events are often predictable. SSA reports consistently emphasize that scarce resources (Office of the Inspector General, 2017) and insufficient funds (Social Security Ad-

¹⁶Because the judge leaves at some point during month zero, the number of judges does not drop until the first full month after retirement.

 $^{^{17}}$ If we extend the post sample to 18 months, most judges have been replaced (Figure A4).

ministration, 2016) have kept them from hiring enough judges to replace retirees. Applicants must also pass the ALJ examination to register as a qualified ALJ candidate. This registry is managed by another agency and only occasionally updated, leaving a limited pool of qualified candidates. As such, judge counts are slow to recover after a retirement.

6.2 Change in Remaining Judges Dispositions, Decisions, and Allowance Rates

In Panel A of Figure 2 we plot the event study coefficients from equation (1) for the natural log of dispositions, decisions, and awards. Dispositions are flat prior to the retirement event, then jump discontinuously by approximately 5 percent in the month of the retirement.¹⁸ Dispositions stay significantly higher for 6 months then drop slightly and become statistically insignificant as judges work through the re-assigned cases and replacement judges enter the office. Decisions follow a similar pattern, suggesting the remaining judges face a higher workload, making 5 percent more dispositions and decisions. The number of awards increases by a slightly smaller amount.

In Panel B of Figure 2 we plot the event study coefficients on the natural log of judge monthly allowance rates (awards/decisions). Allowance rates hold constant prior to the retirement then start falling around two months after the retirement. All of the post estimates are negative, but only three of the months are significant at the five percent level. If we pool observations into two month bins for power, the drop in the allowance rate is more precisely estimated (Figure A5). Notably, the decrease in allowance rates lags the increase in workload, which we would expect if it is the accumulation of backlogged cases that changes judge behavior.¹⁹ Estimates from a longer event study (9 months pre, 18 months post) show that both workload and allowance rate effects have fully faded by 18 months post retirement (Figure A6).

¹⁸This effect masks significant heterogeneity (Figure 3). Although judges experience a fair amount of month-to-month variation in workload, the retirement induces a simultaneous increase in caseloads across judges in the affected office.

¹⁹Office-level impacts on allowance rates are similar (-0.0075, s.e. 0.003 in the post period), suggesting the change is a treatment effect, and not simply a change in the composition of judge-case pairings.

In Table 2 we report the coefficients from equation (2). Dispositions and decisions increase by 4.5 and 4.9 percent respectively in the month of retirement and 2.6 and 3.3 percent over the next nine months after the retirement. At the mean, this corresponds to about one additional case a month. Average daily dispositions and decisions are around 1.6-1.8, suggesting a little more than half a day's additional disposition and decision work in a month. After the retirement, awards also increase, but by about one percent less than the increase in decisions. This results in allowance rates that are 1 percent lower on average. After a peer judge retires, remaining judges hear more cases, make more decisions, and award disability insurance to a smaller share of cases. Since cases are quasi-randomly assigned, this drop is not due to a compositional change in the judge's caseload.

6.3 Relationship Between Caseload and Allowance Rates

We further explore the role of caseload on allowance rates in Figure 3. For each judge in our analysis sample, we calculate average monthly dispositions and allowance rates during the 6 months before and 9 months after the retirement. In Panel A we use a bin-scatter plot with 20 equally size quantiles of the percent change in dispositions (current caseload) on the x-axis and bin-average percent changes in allowance rates on the y-axis. We restrict the sample to judges whose dispositions changed by less than 50 percent, to limit the influence of outliers. Consistent with Table 2, larger increases in dispositions, are associated with larger allowance rates reductions. The bivariate coefficient on percent change in dispositions is -0.032 (s.e. 0.016). In Panel B, we plot the relationship separately by whether the retiring judge had above or below median dispositions in the 6 months leading up to retirement. The negative effect is concentrated among events where the retiree had heavier caseloads (coefficient -0.067 (s.e. 0.026)), which we would expect if this is driven by increased strain from the retiree's additional cases.²⁰

²⁰Consistent with this pattern, if we estimate equation 2 by the retiree's average number of dispositions in the 6 months prior to retirement, in quartiles, the negative effect on allowance rates is monotonically increasing as retirees' caseloads become larger.

This pattern is consistent with existing work showing that negative environmental conditions make judges less favorable to defendants (Danziger et al., 2011; Eren and Mocan, 2018; Heyes and Saberian, 2019) and descriptive work relating caseloads to case outcomes. This works against anecdotal evidence that judges spend more time on denials (Office of the Inspector General, 2017), but other factors, like random quality audits and judge-specific allowance rate reports, discourage judges from simply lowering allowance standards. Unfortunately, the aggregate nature of the data does not allow us to fully tease out why judge allowance rates fall or which types of cases they become more likely to deny.²¹

6.4 Extensions

We explore both heterogeneity and robustness in detail in Appendix B, but highlight key results here.

Robustness Results are robust to various sample restrictions and controls (Table A1) and functional form (natural log versus levels or inverse hyperbolic sine) (Figure A2). Results are not sensitive to restricting the sample of judges or years to exclude retiring judges that were dismissed or targeted by the 2010-2011 ALJ reforms (Tables A1 and A2). Estimates are robust to pre-trend specification tests (Figure A7) (Borusyak and Jaravel, 2020) and using the interaction-weighted event study estimator proposed by Sun and Abraham (2020) (Figure A8). Comparing judges that experience a retirement to judges in hearing offices that experience a retirement exactly one year later as in Deshpande and Li (2019), to exclude negative weighted comparisons (Goodman-Bacon, 2021) and account for potential, underlying secular trends in caseloads and allowance rates, yield similar patterns (Table A3 and Figure A9).

Heterogeneity and Peer Effects As seen in Table A4, female judges experience a 67 percent larger increase in decisions after peer retirements (4.5 vs. 2.7 percent), and a marginally significant additional 0.9 percent decrease in allowance rates. Judges with less

²¹As seen in Table A8, this is not driven by increased dismissal for technical or administrative error.

tenure also experience higher caseloads and lower allowance rates after the retirement.²² These patterns are consistent with an increase in caseload negatively impacting allowance rates and suggest these groups might put more weight on correctly assessing denials relative to awards. Judges with above average allowance rates in their office see large declines in allowance rates after the peer retirement, consistent with mean reversion.

As an alternative to caseloads, peer effects could also affect remaining judge allowance rates. However, we see little evidence of retiring judge characteristics impacting the allowance rate response in the nine months following retirement. One exception is larger negative effects if the retiring judge is relatively new, which would plausibly be more unexpected than a typical retirement, with no replacement judge lined-up and more case-shifting (see Table A5).²³ We explore judge/retiree gender parity and office-level characteristics, but see limited differences (Tables A6 and A7).

7 Implications for the Disability Program

These patterns have potential implications for the size of the disability insurance program. During our analysis sample, judges make 32.4 decisions a month. A one percent lower allowance rate would suggest 0.32 fewer applicants are awarded disability insurance per judge per month. With 5,770 judges experiencing a peer judge retirement disruption, this would suggest that 16,618 additional claimants were denied disability insurance during our sample. These are claimants that would have plausibly been awarded disability insurance if the disruption had not occurred. Unfortunately, the aggregate data do not allow us to document who the marginal denied claimants are, limiting our ability to comment on the welfare effects of judge retirements. For example, judges might be denying more claimants with less severe disabilities (and some scope to work (French and Song, 2014)) or alternatively

²²Subgroups can experience differential increases in caseloads because pending cases are assigned to the judge that last completed a case. Judges that complete cases more quickly will be assigned more cases.

²³If we limit the sample to retirement events where the retiree's average caseload was no more than 10 percent lower 1-3 months before the retirement relative to 4-6 months before retirement (to avoid judges easing to retirement) we find larger caseload effects and a slightly larger decrease in allowance rates.

claimants with more complex cases. One margin we can explore is the fully- and partiallyfavorable allowance rates. Partially-favorable awards occur when the judge does not accepted the alleged onset date, implying the individual became sufficiently disabled to qualify at a later date. These might be cases where the evidence is less clear. As seen in Table A8, the percent change in fully- and partially-favorable allowance rates are nearly the same, suggesting the increased stringency is not so targeted. Regardless, retirement disruptions result in less benefit distribution. Average monthly benefits of \$1,233.70 for disabled workers would imply nearly \$246 million of forgone annual benefits for the 16,618 denied claimants. Even if these claimants are more negatively selected (on earnings) than the average awardee, even benefits at the 25th percentile (\$900) would still imply \$179.5 million of forgone annual benefits (Social Security Administration, 2019).

Higher caseloads could also impact claimant wait times. Claimants wait nearly 15 months on average for an ALJ hearing. Autor et al. (2017) find that delays in the SSDI determination process reduces labor force participation and annual earnings through skill depreciation, while SSA reports that annually nearly 7,800 people die waiting for an ALJ decision (Gross et al., 2018). Using the parameterized equivalent of equation (3), we estimate that a judge retirement increases time to hearing by an insignificant 0.04 months (Table A8). Given this small, insignificant increase in wait time, judge retirements likely have negligible impacts on skill depreciation or mortality of applicants. However, lower allowance rates at the ALJ-level could lead to increased applicant employment (French and Song, 2014).

8 Conclusion

We explore how peer judge retirements affect remaining judges' caseloads and allowance rates among SSA disability judges. Only about 50 percent of the retiring judges have been replaced by 9 months, leading to a 2.6-3.3 percent increase in remaining judges' workload. There is a corresponding one percent drop in the judges' monthly allowance rates, suggesting that some claimants that would have been awarded SSDI under different circumstances, were not because of the change in the work environment. The magnitude of these effects are about 7% of the typical within judge variation.

Judges with less experience exhibit the largest increase in workload and reduction in allowance rates. There is also suggestive evidence that female judges experience a relatively larger increase in caseloads and reduction in allowance rates. Results are consistent with the increase in caseload leading judges to change the way they judge and do not appear to be driven by peer effects from the retiring judge.

We estimate that the reduction in allowances associated with these workplace disruptions have resulted in approximately 16,600 applicants that would have been awarded disability being denied, missing out on hundreds of millions of dollars of annual benefits. With caselevel data we could better evaluate which types of claimants are affected and overall welfare effects. Importantly, these effects could potentially be mitigated or avoided through policy and action aimed at creating a more stable SSA judicial corps.

References

- Alesina, Alberto and Elaina La Ferrara, "A Test of Racial Bias in Capital Sentencing," American Economic Review, 2014, 104 (11), 3397–3433.
- Autor, David and Mark Duggan, "The Rise in the Disability Rolls and the Decline in Unemployment," *Quarterly Journal of Economics*, 2003, 118 (1), 157–206.
- _ , _ , Kyle Greenberg, and David Lyle, "The Impacts of Disability Benefits on Labor Supply: Evidence from the VA's Disability Compensation Program," American Economic Journal: Applied Economics, 2016, 8 (3), 31–68.
- _ , Nicole Maestas, Kathleen Mullen, and Alexander Strand, "Does Delay Cause Decay? The Effect of Administrative Decision Time on the Labor Force Participation Rate and Earnings of Disability Applicants," NBER Working Paper Series No. 20840, 2017.
- Beenstock, Michael and Yoel Haitovksy, "Does the Appointment of Judges Increase the Output of the Judiciary?," *International Review of Law and Economics*, 2004, 24 (3), 351–369.
- Borusyak, Kirill and Xavier Jaravel, "Revisting Event Study Designs," Working Paper, 2020.
- Boyd, Christina, Lee Epstein, and Andrew Martin, "Untangling the Causal Effects of Sex on Judging," American Journal of Political Science, 2010, 54 (2), 389–411.
- Callaway, Brantly and Pedro Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 2020, *In Press.*
- Danziger, Shai, Jonathan Levav, and Liora Avnaim-Pesso, "Extraneous factors in judicial decisions," *PNAS*, 2011, 108 (17), 6889–6892.
- de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," American Economic Review, 2020, 10 (9), 2964–96.
- **Depew, Briggs, Ozkan Eren, and Naci Mocan**, "Judges, Juveniles, and In-Group Bias," *Journal of Law and Economics*, 2017, 60 (2), 209–239.
- **Deshpande, Manasi and Yue Li**, "Who Is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Economic Policy*, 2019, 11 (4), 213–248.
- _ , Tal Gross, and Yalun Su, "Disability and Distress: The Effect of Disability Programs on Financial Outcomes," *American Economic Journal: Applied Economics*, 2020, *forthcoming.*

- Engel, Christoph and Keren Weinshall, "Manna from Heaven for Judges: Judges' Reaction to a Quasi-Random Reduction in Caseload," *Journal of Empirical Legal Studies*, 2020, 17 (4), 722–751.
- Eren, Ozkan and Naci Mocan, "Emotional Judges and Unlucky Juveniles," American Economic Journal: Applied Economics, 2018, 10 (3), 171–205.
- and _ , "Judge Peer Effects in the Courthouse," NBER Working Paper Series No. 27713, 2020.
- Frandsen, Brigham, Lars Lefgren, and Emily Leslie, "Judging Judge Fixed Effects," NBER Working Paper Series No. 25528, 2019.
- French, Eric and Jae Song, "The Effect of Disability Insurance Receipt on Labor Supply," American Economics Journal: Applied Economics, 2014, 6 (2), 291–337.
- Gelber, Alexander, Timothy Moore, and Alexander Strand, "The Effect of Disability Insurance Payments on Beneficiaries's Earnings," *American Economic Journal: Applied Economics*, 2017, 9 (3), 229–261.
- _ , _ , and _ , "Disability Insurance Income Saves Lives," Working paper, 2018.
- Goodman-Bacon, Andrew, "Difference-in-differences with variation in treatment timing," Journal of Econometrics, 2021, Forthcoming.
- **Government Accountability Office**, "Additional Measures and Evaluation Needed to Enhance Accuracy and Consistency of Hearings Decisions," 2017. https://www.gao.gov/assets/690/688824.pdf.
- Gross, Stephen, Jacqueline Walsh, and Bert Kestenbaum, "Probability of Death While Pending an Administrative Law Judge Determination," 2018. Note 159.
- Heyes, Anthony and Soodeh Saberian, "Temperature and Decisions: Evidence from 207,000 Court Cases," American Economic Journal: Applied Economics, 2019, 11 (2), 238–265.
- Huang, Bert I., "Lightened Scrutiny," Harvard Law Review, 2011, 124 (5), 1109–1152.
- Iverson, Benjamin, "Get in Line: Chapter 11 Restructuring in Crowded Bankruptcy Courts," *Management Science*, 2018, 64 (11), 5370–5394.
- Kearney, Melissa S., Brendan M. Price, and Riley Wilson, "Disability Insurance in the Great Recession: Ease of Access, Program Enrollment, and Local Hysteresis," AEA Papers and Proceedings, 2021, 111, 486–490.
- Kling, Jeffrey, "Incarceration length, employment, and earnings," American Economics Review, 2006, 96 (3), 863–876.

- Lim, Claire S. H., Bernardo Silveira, and James M. Snyder, "Do Judges' Characteristics Matter? Ethnicity, Gender, and Partisanship in Texas State Trial Courts," *American Law and Economics Review*, 2016, 18 (2), 302–357.
- _, James M. Snyder, and David Stromberg, "The Judge, the Politician, and the Press: Newspaper Coverage and Criminal Sentencing across Electoral Systems," American Economics Journal: Applied Economics, 2015, 7 (4), 103–135.
- Maestas, Nicole, Kathleen Mullen, and Alexander Strand, "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment of Estimate Causal Effects of SSDI Receipt," *American Economic Review*, 2013, 103 (5), 1797–1829.
- Moore, Timothy, "The employment effects of terminating disability benefits," Journal of Public Economics, 2014, 124, 30–43.
- Mueller-Smith, Michael, "The Criminal and Labor Market Impacts of Incarceration," Working Paper, 2015.
- Mustard, David, "Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts," *The Journal of Law and Economics*, 2001, 44 (1), 285–314.
- Nardulli, Peter F., "The Caseload Controversy and the Study of Criminal Courts," *Journal of Criminal Law and Criminology*, 1979, 70 (1), 89–101.
- Office of the Inspector General, "Congressional Response Report: The Social Security Administration's Review of Administrative Law Judges' Decisions," 2012. Report A-07-12-21234.
- __, "Congressional Response Report: Administrative Law Judges with Both High Dispositions and High Allowance Rates," 2014. Report A-12-14-24092.
- _ , "Factors Related to Decreased Administrative Law Judge Productivity," 2017. Report A-12-18-50289.
- Park, Kyung H., "Do Judges Have Tastes for Discrimination? Evidence from Criminal Courts," *Reivew of Economics and Statistics*, 2017, 99 (5), 810–823.
- Rehavi, M. Marti and Sonja Starr, "Racial Disparity in Federal Criminal Sentences," Journal of Political Economy, 2014, 122 (6), 1320–1354.
- Robel, Lauren K., "Caseload and Judging: Judicial Adaptations to Caseloads," Brigham Young University Law Review, 1990, 1, 3–65.
- **Social Security Administration**, "Leading the Hearings and Appeals Process into the Future," 2016. CARES Plan (January 13, 2016).
- _ , "Annual Statistical Report on the Social Security Disability Insurance Program, 2018," 2019. SSA Publication No. 13-11826.

- Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2020, *In Press.*
- Ulmer, Jeffery T. and Brian Johnson, "Sentencing in Context: A Multilevel Analysis," Criminology, 2004, 42 (1), 137–178.
- Yang, Crystal S., "Resource Constraints and the Criminal Justice System: Evidence from Judicial Vacanies," *American Economic Journal: Economic Policy*, 2016, 8 (4), 289–332.

Tables and Figures



Figure 1: Impact of Judge Retirement on Office-level Number of Judges

Notes: Observation at the Hearing Office by month level. Each point is an event study coefficient obtained from estimating equation (3), where the outcome is the number of judges hearing cases at the Hearing Office-level. Month zero is the last month the judge is observed in the data, meaning they left at some point during that month. The sixth month prior to the retirement is the omitted group. Regression includes Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.





Notes: Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (1), where the outcome is the natural log of the number of dispositions, decisions, awards, or the allowance rate. Only judges who appear all 6 months before the retirement event, and all 9 months after are included. Month zero is the last month the judge is observed in the data, meaning they left at some point during that month. The sixth month prior to the retirement is the omitted group. Regression includes judge, Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.



Figure 3: Potential Mechanisms: Relationship between Caseload and Allowance Rates Changes After a Peer Retirement

Notes: Each point represents the average percent change in allowance rates from the pre- to post-retirement period by the average percent change in pre- to post-retirement dispositions in 20 equal-sized quantile bines. The sample is restricted to judges that experiences changes in dispositions less than 50 percent. In the left panel, all judges are included. The bivariate regression coefficient on dispositions is -0.032 (s.e. 0.016). In the right panel, judges in retirement events where the retiree's average pre-retirement dispositions were above and below median are plotted separately. The bivariate regression coefficient on dispositions is -0.067 (s.e. 0.026) for above median judges and 0.004 (s.e. 0.024) for below median judges.

		E	vent Study Analysis	Sample	
	Full Sample (1)	Full Sample (2)	Before Retirement (3)	After Retirement (4)	Retirees (5)
Dispositions	38.28	40.17	39.97	40.29	34.34
Decisions	31.17	32.43	32.29	32.52	28.61
Awards	17.44	18.32	18.41	18.27	16.80
Fully Favorable	15.25	16.06	16.21	15.98	14.80
Allowance Rate	0.56	0.56	0.57	0.56	0.58
Female First Name	0.34	0.34	0.34	0.34	0.25
Observation Year	2014.75	2014.80	2014.39	2015.05	2014.34
Monthly Observations	158,065	90,144	33,804	56,340	4,488

Table 1: Judge Monthly Summary Statistics

Notes: Observation at the judge by month level. Judges at National Hearing Centers are excluded. Mean values for the entire sample of available data (September 2010 - July 2019) available in column (1). The sample in columns (2) through (4) is restricted to judge by month observations in the event study window of a retirement event and only includes judges that are observed in all months of the event study window. As some retirement events happen within a short period of time, some judge by month observations will be duplicated in the analysis sample. The sample in column (5) is restricted to monthly observations during the "pre-period" of the event study window of the judges that retire.

		Natur	al Log of					
	Dispositions (1)	Decisions (2)	Awards (3)	Allowance Rate (4)	Dispositions (5)	Decisions (6)	Awards (7)	Allowance Rate (8)
Month of Colleague Retirement	0.045^{***} (0.008) 0.026^{***}	0.049^{***} (0.008) 0.033^{***}	0.045^{***} (0.010) 0.025^{***}	-0.004 (0.005) -0.009***	1.799^{***} (0.311) 0.888^{***}	1.501^{***} (0.272) 0.882^{***}	0.782^{***} (0.181) 0.305^{***}	-0.003 (0.002) -0.005***
The concugate from one of	(0.005)	(0.005)	(0.005)	(0.003)	(0.149)	(0.133)	(0.078)	(0.002)
Dependent Mean (in Levels) Observations	$40.48 \\ 89,455$	32.77 89,205	$18.61 \\ 88,753$	$0.56 \\ 88,753$	$40.17 \\ 90,144$	$32.43 \\ 90,144$	$18.32 \\ 90,144$	$0.56 \\ 89,205$

Table 2: Average Impact of Colleague Retirement on Judge-level Caseload and Allowance Rate

Notes: Observation at the judge by month level. Sample restricted to the event study balanced panel and only includes judges who appear all 6 months before the retirement event, and all 9 months after are included. The sample size is not uniform because some judge-month observations equal zero, making the outcome undefined. Estimates are robust to estimation in levels or using the inverse hyperbolic sine (Figure A2). Judge, Hearing Office, year, and month fixed effects are included. Standard errors corrected for clustering at the Hearing Office level. p<0.01 ***, p<0.05 **, p<0.1 *.





Figure A1: Retirement Event Timing and Number of Judges the Month Before a Retirement Event



Figure A2: Event Study Effects Untransformed (in Levels) or Using the Inverse Hyperbolic Sine Transformation

Notes: Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (1), where the outcome is the number of dispositions, decisions, awards, or the allowance rate in levels (in the top panel) or using the inverse hyperbolic sine transformation (IHS) (in the bottom panel). The IHS transformation is similar to the natural log, but is defined at zero. Only judges who appear all 6 months before the retirement event, and all 9 months after are included. Month zero is the last month the judge is observed in the data, meaning they left at some point during that month. The sixth month prior to the retirement is the omitted group. Regression includes judge, Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.



Figure A3: Impact of Judge Retirement on Judge-level Caseload and Allowance Rates, Alternative Omitted Period

Notes: Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (1), where the outcome is the natural log of the number of dispositions, decisions, awards, or the allowance rate. Only judges who appear all 6 months before the retirement event, and all 9 months after are included. Month zero is the last month the judge is observed in the data, meaning they left at some point during that month. The month prior to the retirement is the omitted group. Regression includes judge, Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.



Figure A4: Impact of Judge Retirement on Office-level Number of Judges, Extended Post Period

Notes: Observation at the Hearing Office by month level. Each point is an event study coefficient obtained from estimating an equation similar to (3), but with 18 post period months, where the outcome is the number of judges hearing cases at the Hearing Office-level. Month zero is the last month the judge is observed in the data, meaning they left at some point during that month. The sixth month prior to the retirement is the omitted group. Regression includes Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.



Figure A5: Impact of Judge Retirement on Judge-level Allowance Rate, 2 Month Bins

Notes: Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (1), where the outcome is the natural log of the monthy allowance rate. Only judges who appear all 6 months before the retirement event, and all 9 months after are included. Month zero is the last month the judge is observed in the data, meaning they left at some point during that month. The sixth month prior to the retirement is the omitted group. Regression includes judge, Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.



Figure A6: Impact of Judge Retirement on Judge-level Caseload and Allowance Rates, Extended Sample Window

32

Notes: Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (1), where the outcome is the natural log of the number of dispositions, decisions, awards, or the allowance rate. Only judges who appear all 9 months before the retirement event, and 18 months after are included, resulting a smaller sample than our baseline estimates. Month zero is the last month the judge is observed in the data, meaning they left at some point during that month. The ninth month prior to the retirement is the omitted group. To avoid overlapping panels, retirement events within nine months of each other are excluded. Regression includes judge, Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.



Figure A7: Pre-Trend Specification Checks

Notes: These tests builds on (Borusyak and Jaravel, 2020) to explore robustness of effects to the pre-trend. Observation at the judge by month level. In the top graph, each point is an event study coefficient obtained from estimating an equation similar to equation (1), but only the pre-period indicators are included and the indicators for both six and one months prior are excluded. In the bottom graphs, each point is an event study coefficient obtained from estimating equation (1), where the outcome is the natural log of the number of dispositions, decisions, awards, or allowance rates but all of the pre-period estimates are omitted so that dynamic effects can be identified. Only judges who appear all 6 months before the retirement event, and all 9 months after are included. Month zero is the last month the retiring judge is observed in the data, meaning they left at some point during that month. Regression includes judge, Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.



Figure A8: Staggered Event Study Corrections: Impact of Judge Retirement on Judge-level Caseload and Allowance Rates

Notes: This replicates the results in Figure 2, but uses the interaction-weighted estimator proposed by Sun and Abraham (2020) to correct for improper weighting in event study analysis with staggered treatment timing. The last retirement event is treated as the control unit. Observation at the judge by month level. Only judges who appear all 6 months before the retirement event, and all 9 months after are included. Month zero is the last month the judge is observed in the data, meaning they left at some point during that month. The sixth month prior to the retirement is the omitted group. Regression includes judge, Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.



30

Figure A9: Impact of Judge Retirement on Judge-level Caseload, Stacked Event Panel Including Future Retirement Events as Controls

Notes: Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (4), where the outcome is the natural log of the number of dispositions, decisions, or awards. Only judges who appear all 6 months before the retirement event, and all 9 months after are included. Month zero is the last month the judge is observed in the data, meaning they left at some point during that month. The sixth month prior to the retirement is the omitted group. Regression includes judge, Hearing Office, year, and month fixed effects. Standard errors are corrected for clustering at the Hearing Office-level. 95 percent confidence intervals are provided.

Table A1: Robustness: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rates

		Natur	al Log of					
	Dispositions	Decisions	Awards	Allowance Rate	Dispositions	Decisions	Awards	Allowance Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A.		I	nclude Lab	or Market, Indust	ry, and Demog	raphic Cont	trols	
Month of Colleague Retirement	0.044^{***}	0.048^{***}	0.038^{***}	-0.010*	1.730***	1.459***	0.660^{***}	-0.006**
	(0.010)	(0.010)	(0.011)	(0.005)	(0.365)	(0.319)	(0.207)	(0.002)
After Colleague Retirement	0.025***	0.032***	0.023***	-0.011***	0.871***	0.881***	0.279***	-0.006***
	(0.005)	(0.006)	(0.005)	(0.003)	(0.163)	(0.145)	(0.081)	(0.002)
Dependent Mean (in Levels)	40.36	32.81	18 57	0.56	40.08	32.48	18.28	0.56
Observations	68 651	68 437	68 076	68.076	69 134	69 134	69 134	68 437
D 1D	00,001	00,101	00,010	T L L N C L	U : C /	00,101	00,101	00,101
Panel B. Month of Colleague Betirement	0.044***	0.046***	0.044***	Include National	1 606***	ers 1 415***	0 749***	0.003
Month of Colleague Retrement	(0.008)	(0.040	(0.010)	(0.005)	(0.312)	(0.275)	(0.142	(0.002)
After Colleague Retirement	0.024***	0.031***	0.024***	-0.008***	0.808***	0.809***	0.287***	-0.005***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.149)	(0.138)	(0.075)	(0.002)
Dependent Mean (in Levels)	40.55	32.82	18.54	0.56	40.24	32.47	18.25	0.56
Observations	93,838	93,571	93,099	93,099	94,576	94,576	94,576	93,571
Panel C.			Include E	veryone with Any	Pre- and Post-	Observatio	ons	
Month of Colleague Retirement	0.028***	0.033***	0.031***	-0.003	1.188^{***}	1.029^{***}	0.514^{***}	-0.003
	(0.009)	(0.009)	(0.010)	(0.005)	(0.301)	(0.265)	(0.168)	(0.002)
After Colleague Retirement	0.025***	0.034^{***}	0.030***	-0.013***	0.913***	0.957^{***}	0.314^{***}	-0.008***
	(0.005)	(0.006)	(0.005)	(0.003)	(0.148)	(0.139)	(0.079)	(0.001)
Dependent Mean (in Levels)	39.47	32.02	18 15	0.56	38 78	31.32	17.60	0.56
Observations	103.862	103.379	102.498	102.498	105.714	105.714	105.714	103.379
Panel D	,	,	,	Non Overlapping	Front Windo			,
Month of Colleague Betirement.	0.050***	0.055***	0.058***	0 001	2.091***	us 1 768***	1 039***	-0.001
Month of Concegae Rectionent	(0.011)	(0.012)	(0.013)	(0.006)	(0.448)	(0.377)	(0.248)	(0.003)
After Colleague Retirement	0.034***	0.042***	0.032***	-0.012***	1.284***	1.201***	0.439***	-0.006***
	(0.005)	(0.006)	(0.006)	(0.003)	(0.174)	(0.162)	(0.097)	(0.002)
Dependent Mean (in Levels)	40.92	33.16	18.67	0.56	40.63	32.85	18.41	0.56
Observations	51,930	51,809	51,576	51,576	52,304	52,304	52,304	51,809
Panel E.		Exc	lude Event	s Where Retiree D	isciplined or F	eviewed by	MSPB	
Month of Colleague Retirement	0.046***	0.049***	0.045***	-0.005	1.789***	1.485***	0.767***	-0.004
After Colleague Petirement	(0.008)	(0.008)	(0.010)	(0.005)	(0.304)	(0.200)	(0.177)	(0.002)
Arter Coneague Retrement	(0.020	(0.005)	(0.025)	-0.009	(0.147)	(0.132)	(0.079)	-0.005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.111)	(0.102)	(0.010)	(0.002)
Dependent Mean (in Levels)	40.45	32.75	18.60	0.57	40.14	32.40	18.31	0.56
Observations	88,848	88,599	88,147	88,147	89,536	89,536	89,536	88,599
Panel F.			Hearii	ng Office and Year	-by-Month Fix	ed Effects		
Month of Colleague Retirement	0.023***	0.027***	0.030***	0.002	0.940***	0.827***	0.503^{***}	-0.000
	(0.007)	(0.007)	(0.008)	(0.005)	(0.230)	(0.205)	(0.139)	(0.002)
After Colleague Retirement	0.021^{***}	0.028^{***}	0.023^{***}	-0.006*	0.751^{***}	0.755^{***}	0.307^{***}	-0.003**
	(0.005)	(0.005)	(0.005)	(0.003)	(0.149)	(0.131)	(0.080)	(0.002)
	40,40	00.77	10.01	0.50	10.17	20.42	10.90	0.50
Observations	40.48	32.77 80.205	18.01	0.50	40.17	32.43	18.32	06.00
Observations	69,455	69,200		00,155	90,144	90,144	90,144	69,200
Panel G.	0.005***	0.095***	Ret	irement Event and	1 Month Fixed	Effects	0 5 49***	0.004
Month of Colleague Retirement	0.035***	0.035***	0.032***	-0.004	1.423***	1.109^{+++} (0.267)	(0.175)	-0.004
After Colleague Retirement	0.008	0.000)	-0.003)	-0.012***	0.207	0.207)	-0.206**	-0.002/
The concague rectionent	(0.006)	(0.006)	(0.001)	(0.003)	(0.175)	(0.152)	(0.089)	(0.002)
	()	()	()	()	()	(=)	()	()
Dependent Mean (in Levels)	40.48	32.77	18.61	0.56	40.17	32.43	18.32	0.56
Observations	89,455	89,205	88,753	88,753	90,144	90,144	90,144	89,205

Notes: Observation at the judge by month level. Sample restricted to the event study balanced panel and only includes judges who appear all 6 months before the retirement event, and all 9 months after are included. Judge, Hearing Office, year, and month fixed effects are included. Standard errors corrected for clustering at the Hearing Office level. Panel A includes judges at the four National Hearing Centers which take excess cases from other Hearing Offices. Panel B includes the Hearing Office-level unemployment rate, labor force participation rate, industry composition shares, race shares, percent of the population under 20, and percent of the population over 65. These measures are aggregated up from county-level measures based on Hearing Office assignment. Panel C includes any judge that appears in both the pre- and post-period, and is not balanced. Panel D only includes retirement events that are at least nine months from any other retirement event, so that there are no overlapping event study windows. Panel E excludes events where there is any record that the retiring judge underwent review through the MSPB. Panel F includes hearing office and month-by-year fixed effects. Panel G includes month and retirement event-level fixed effects. p<0.01***, p<0.05 **, p<0.1 *.

Table A2: Heterogeneous Impact of Colleague Retirement on Judge-level Caseload and Allowance Rates by Time Period

		Natur	al Log of					
	Dispositions (1)	Decisions (2)	Awards (3)	Allowance Rate (4)	Dispositions (5)	Decisions (6)	Awards (7)	Allowance Rate (8)
				Retirement	in 2011-2014			
Month of Colleague Retirement	0.040***	0.045***	0.046***	0.003	1.722***	1.535***	0.909***	0.000
	(0.011)	(0.011)	(0.013)	(0.006)	(0.476)	(0.408)	(0.280)	(0.003)
After Colleague Retirement	0.010	0.015**	0.018**	0.001	0.357	0.403*	0.163	-0.001
	(0.006)	(0.006)	(0.008)	(0.004)	(0.235)	(0.210)	(0.151)	(0.002)
Dependent Mean (in Levels)	42.89	35.29	20.25	0.57	42.57	34.97	19.98	0.56
Observations	41,273	41,208	41,029	41,029	$41,\!584$	$41,\!584$	$41,\!584$	41,208
				Retirement	in 2015-2018			
Month of Colleague Retirement	0.049***	0.051***	0.044***	-0.009	1.819***	1.454***	0.640***	-0.006*
	(0.011)	(0.012)	(0.014)	(0.008)	(0.369)	(0.324)	(0.212)	(0.004)
After Colleague Retirement	0.033***	0.043***	0.026***	-0.018***	1.215***	1.207***	0.345***	-0.009***
0	(0.007)	(0.008)	(0.008)	(0.005)	(0.206)	(0.195)	(0.124)	(0.002)
Dependent Mean (in Levels)	38.42	30.61	17.20	0.56	38.12	30.26	16.91	0.56
Observations	48,182	47,997	47,724	47,724	48,560	48,560	48,560	47,997

Notes: Observation at the judge by month level. Sample restricted to the event study balanced panel and only includes judges who appear all 6 months before the retirement event, and all 9 months after are included. Judge, Hearing Office, year, and month fixed effects are included. Judge's gender is proxied by the genderedness of their first name. Names that are over 50 percent female in the SSA birth name records are labeled women. Judge's race is proxied by the racial composition of their surname using US Census data that reports what fraction of the population with a given last name is White, Black, or Hispanic. Standard errors corrected for clustering at the Hearing Office level. p<0.01 ***, p<0.05 **, p<0.1 *.

		Natur	al Log of					
	Dispositions	Decisions	Awards	Allowance Rate	Dispositions	Decisions	Awards	Allowance Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Stacked Baselin	ne Specificatio	1		
Treated	-0.014**	-0.018***	-0.012*	0.005	-0.366**	-0.396**	-0.110	0.003^{*}
	(0.006)	(0.006)	(0.006)	(0.003)	(0.184)	(0.158)	(0.105)	(0.002)
Treated*Month of Colleague Retirement	0.017	0.022^{*}	0.022	-0.002	0.663^{*}	0.636^{*}	0.324	-0.002
	(0.011)	(0.012)	(0.014)	(0.008)	(0.398)	(0.351)	(0.234)	(0.004)
Treated*After Colleague Retirement	0.034^{***}	0.041^{***}	0.028^{***}	-0.011**	0.966^{***}	0.942^{***}	0.311**	-0.006**
	(0.007)	(0.008)	(0.008)	(0.005)	(0.234)	(0.208)	(0.132)	(0.002)
Dependent Mean (in Levels)	39.96	32.46	18.58	0.57	39.69	32.16	18.31	0.57
Observations	839,737	$837,\!625$	$833,\!178$	$833,\!178$	$845,\!440$	$845,\!440$	$845,\!440$	$837,\!625$
		Stac	ked Baselir	ne Specification wi	th Year and M	lonth Fixed	Effects	
Treated	-0.018***	-0.021***	-0.015**	0.004	-0.497^{***}	-0.494***	-0.170*	0.002
	(0.005)	(0.005)	(0.006)	(0.003)	(0.167)	(0.145)	(0.097)	(0.001)
Treated*Month of Colleague Retirement	0.017	0.024**	0.021	-0.004	0.614	0.660*	0.300	-0.003
	(0.011)	(0.011)	(0.013)	(0.007)	(0.389)	(0.343)	(0.228)	(0.003)
Treated*After Colleague Retirement	0.041***	0.046***	0.035***	-0.009**	1.204***	1.114***	0.421***	-0.005**
	(0.007)	(0.007)	(0.007)	(0.004)	(0.204)	(0.184)	(0.116)	(0.002)
Dependent Mean (in Levels)	39.96	32.46	18.58	0.57	39.69	32.16	18.31	0.57
Observations	839,737	837,625	833,178	833,178	845,440	845,440	845,440	837,625

Table A3: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rates, Retirement-Event Stacked Panel

Notes: Observation at the judge by month level. Sample includes judges from each retirement event, as well as retirement events exactly one year later, but during the same time as the focal retirement, similar to Deshpande and Li (2019). Sample restricted to the event study balanced panel and only includes judges who appear all 6 months before the retirement event, and all 9 months after are included. Judge, event, Hearing Office, and month fixed effects are included. Standard errors corrected for clustering at the Hearing Office level. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A4:	Impact of	Colleague	Retirement	on Judge-level	Caseload	and Allowance	Rates by	Judge	Characteristics
	1	0		0				0	

$\begin{tabular}{ c c c c c c c c c c c c c c c c c c c$	10owance Rate (8) -0.003 (0.002) -0.004*** (0.002) 0.000 (0.004) -0.004*
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	(8) -0.003 (0.002) -0.004** (0.002) 0.000 (0.004) -0.004* (0.002)
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c} -0.003 \\ (0.002) \\ -0.004^{**} \\ (0.002) \\ 0.000 \\ (0.004) \\ -0.004^{*} \\ (0.002) \end{array}$
Month of Colleague Retirement 0.032^{+10} 0.042^{+10} 0.042^{+10} 0.001 1.44^{+10} 1.425^{+10} 0.744^{+10} After Colleague Retirement (0.009) (0.010) (0.005) (0.363) (0.314) (0.213^{*}) Month of Colleague Retirement 0.021^{**} 0.022^{**} -0.006^{*} 0.651^{***} 0.665^{***} 0.213^{**} Month of Colleague Retirement 0.021^{*} 0.019 0.001 -0.010 0.192 0.232 0.114 *Female (0.011) (0.012) (0.013) (0.009) (0.363) (0.316) (0.206)	-0.003 (0.002) -0.004** (0.002) 0.000 (0.004) -0.004*
After Colleague Retirement (0.009) (0.010) (0.005) (0.005) (0.005) (0.010) (0.005) (0.005) (0.010) (0.005) (0.005) (0.005) (0.005) (0.005) (0.005) (0.005) (0.005) (0.005) (0.005) (0.005) (0.005) (0.005) (0.005) (0.004) (0.158) (0.142) (0.090) Month of Colleague Retirement 0.021^* 0.019 0.001 -0.010 0.192 0.232 0.114 *Female (0.011) (0.012) (0.013) (0.009) (0.363) (0.316) (0.206)	(0.002) - 0.004^{**} (0.002) 0.000 (0.004) - 0.004^{*}
After Colleague Retirement 0.521 0.524 0.522 -0.600 0.531 0.605 0.731 Month of Colleague Retirement 0.021^* 0.019 0.001 -0.010 0.192 0.232 0.114 *Female (0.011) (0.012) (0.013) (0.009) (0.363) (0.316) (0.206)	(0.004) (0.002) (0.000) (0.004) -0.004^*
Month of Colleague Retirement (0.003) (0.003) (0.003) (0.004) (0.133) (0.142) (0.030) *Female (0.011) (0.012) (0.013) (0.009) (0.363) (0.316) (0.206)	(0.002) 0.000 (0.004) -0.004^* (0.002)
Female (0.011) (0.013) (0.009) (0.363) (0.316) (0.206)	(0.004) -0.004
(0.012) (0.012) (0.003) (0.003) (0.003) (0.003) (0.003) (0.200)	-0.004*
0.019 0.019 0.019 0.011 0.0008 0.719888 0.021888 0.02788	-0.004
After Concegue Retirement 0.015 0.016° 0.011 -0.009° 0.122° 0.517° 0.217° *Energie (0.000) (0.000) (0.000) (0.000) (0.122^{\circ}) (0.124^{\circ}) (0.124^{\circ})	(() () () ())
(0.009) (0.009) (0.009) (0.005) (0.257) (0.199) (0.154)	(0.002)
Dependent Mean (in Levels) 40.49 32.78 18.62 0.57 40.18 32.44 18.33	0.56
Observations 89,313 89,063 88,611 88,611 90,000 90,000 90,000	89,063
Month of Colleague Retirement 0.036*** 0.040*** 0.037*** -0.004 1.508*** 1.244*** 0.622***	-0.003
(0.008) (0.008) (0.009) (0.005) (0.285) (0.253) (0.162)	(0.002)
After Colleague Retirement 0.018*** 0.026*** 0.021*** -0.006* 0.724*** 0.747*** 0.287***	-0.004**
(0.005) (0.005) (0.005) (0.003) (0.146) (0.131) (0.081)	(0.002)
Month of Colleague Retirement 0.139^{***} 0.132^{***} 0.102^{***} -0.021^{**} 4.093^{***} 3.544^{***} 1.897^{***}	-0.011*
* < 1 Year as Judge (0.030) (0.029) (0.032) (0.012) (1.128) (0.975) (0.691)	(0.006)
After Colleague Retirement 0.144*** 0.139*** 0.073*** -0.063*** 3.471*** 2.783*** 0.286	-0.033***
*< 1 Year as Judge (0.026) (0.026) (0.027) (0.010) (0.799) (0.696) (0.451)	(0.005)
	(0.000)
Dependent Mean (in Levels) 40.48 32.77 18.61 0.56 40.17 32.43 18.32	0.56
Observations 89,455 89,205 88,753 88,753 90,144 90,144 90,144	89,205
Month of Colleague Retirement 0.045*** 0.048*** 0.047*** -0.001 1.767*** 1.466*** 0.807***	-0.002
(0.008) (0.008) (0.009) (0.005) (0.310) (0.272) (0.175)	(0.002)
After Colleague Retirement 0.026*** 0.033*** 0.028*** -0.007** 0.883*** 0.878*** 0.345***	-0.004***
(0.005) (0.005) (0.005) (0.003) (0.148) (0.133) (0.077)	(0.001)
Month of Colleague Retirement 0.029 0.050 -0.197*** -0.272*** 0.658 1.221 -2.059**	-0.132***
*Judge Pre- Award Rate Minus Office Award Rate (0.046) (0.050) (0.050) (0.033) (1.523) (1.327) (0.820)	(0.014)
After Colleague Retirement -0.027 -0.045 -0.327*** -0.325*** -1.561* -1.516** -5.457***	-0.182***
*Judge Pre- Award Rate Minus Office Award Rate (0.031) (0.033) (0.035) (0.024) (0.868) (0.729) (0.578)	(0.012)
	. /
Dependent Mean (in Levels) 40.48 32.78 18.61 0.56 40.18 32.44 18.32	0.56
Observations 89,446 89,196 88,744 90,128 90,128 90,128	89,196

Notes: Observation at the judge by month level. Sample restricted to the event study balanced panel and only includes judges who appear all 6 months before the retirement event, and all 9 months after are included. Judge, Hearing Office, year, and month fixed effects are included. Judge's gender is proxied by the genderedness of their first name. Names that are over 50 percent female in the SSA birth name records are labeled women. Standard errors corrected for clustering at the Hearing Office level. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A	5:	Impact of	Ċ	Colleague	Retirement	on .	Judge-l	evel	Caseload	and	Allowance	Rates b	v Retiree's	s C	haracteristics
				()			()						/		

		Natu	ral Log of					
	Dispositions (1)	Decisions (2)	Awards (3)	Allowance Rate (4)	Dispositions (5)	Decisions (6)	Awards (7)	Allowance Rate (8)
Month of Colleague Retirement	0.055***	0.058***	0.059***	0.000	2.118***	1.799***	1.007***	-0.002
After Colleague Retirement	(0.010) 0.027^{***} (0.005)	0.034***	0.026***	-0.009***	(0.331) 0.970*** (0.164)	(0.335) 0.949^{***} (0.151)	(0.219) 0.355^{***} (0.094)	-0.005***
Month of Colleague Retirement *Retiree Female	-0.039^{**} (0.019)	-0.037^{*} (0.021)	-0.057^{***} (0.021)	-0.018* (0.011)	(0.101) -1.270* (0.657)	-1.189** (0.571)	-0.894^{**} (0.354)	-0.005
After Colleague Retirement *Retiree Female	-0.004 (0.010)	-0.003 (0.012)	-0.002 (0.013)	-0.001 (0.006)	-0.320 (0.321)	-0.265 (0.291)	-0.198 (0.186)	0.000 (0.003)
Dependent Mean (in Levels) Observations	40.48 89,455	32.77 89,205	18.61 88,753	0.56 88,753	40.17 90,144	32.43 90,144	18.32 90,144	0.56 89,205
Month of Colleague Retirement	0.042^{***} (0.008)	0.044^{***} (0.009)	0.043*** (0.010)	-0.002 (0.005)	1.666^{***} (0.293)	1.365*** (0.260)	0.708*** (0.166)	-0.003
After Colleague Retirement	0.027*** (0.005)	0.035*** (0.005)	0.029*** (0.005)	-0.008** (0.003)	0.988*** (0.159)	0.979*** (0.142)	0.378*** (0.084)	-0.005*** (0.002)
Month of Colleague Retirement *Retiree < 1 Year as Judge	(0.032) (0.029)	(0.047) (0.030)	(0.020) (0.029)	-0.025* (0.013)	1.327 (1.092)	1.333 (0.979)	0.592 (0.651)	-0.011 (0.007)
After Colleague Retirement *Retiree < 1 Year as Judge	-0.002 (0.014)	-0.013 (0.015)	-0.037** (0.017)	-0.024*** (0.009)	-0.820* (0.492)	-0.853* (0.469)	-1.012*** (0.340)	-0.015*** (0.005)
Dependent Mean (in Levels) Observations	$40.48 \\ 89,455$	32.77 89,205	18.61 88,753	$0.56 \\ 88,753$	$40.17 \\ 90,144$	$32.43 \\ 90,144$	$18.32 \\ 90,144$	$0.56 \\ 89,205$
Month of Colleague Retirement	0.046***	0.049***	0.045***	-0.004	1.823***	1.519***	0.793***	-0.003
After Colleague Retirement	(0.000) 0.025^{***} (0.005)	(0.000) 0.033^{***} (0.005)	(0.010) 0.025^{***} (0.005)	-0.009*** (0.003)	(0.310) 0.881^{***} (0.148)	(0.274) 0.878^{***} (0.134)	(0.102) 0.304^{***} (0.078)	-0.005*** (0.002)
Month of Colleague Retirement *Retiree Pre- Award Rate Minus Office Award Rate	-0.014 (0.051)	(0.025) (0.058)	(0.013) (0.060)	0.034 (0.034)	-2.505 (1.999)	-1.866 (1.835)	(1.029) (1.083)	0.008 (0.016)
After Colleague Retirement *Retiree Pre- Award Rate Minus Office Award Rate	(0.052) (0.034)	0.052 (0.039)	(0.042) (0.038)	-0.011 (0.021)	(1.399) (1.153)	0.986 (1.072)	(0.437) (0.592)	-0.006 (0.010)
Dependent Mean (in Levels) Observations	40.47 89,375	32.77 89,125	$18.61 \\ 88,673$	0.56 88,673	40.16 90,064	$32.43 \\ 90,064$	$18.32 \\ 90,064$	$\begin{array}{c} 0.56\\ 89,125\end{array}$

Notes: Observation at the judge by month level. Sample restricted to the event study balanced panel and only includes judges who appear all 6 months before the retirement event, and all 9 months after are included. Judge, Hearing Office, year, and month fixed effects are included. Judge's gender is proxied by the genderedness of their first name. Names that are over 50 percent female in the SSA birth name records are labeled women. Standard errors corrected for clustering at the Hearing Office level. p<0.01 ***, p<0.05 **, p<0.1 *.

		Natu	al Log of					
	Dispositions (1)	Decisions (2)	Awards (3)	Allowance Rate (4)	Dispositions (5)	Decisions (6)	Awards (7)	Allowance Rate (8)
Month of Colleague Retirement	0.051^{***}	0.055^{***}	0.061^{***}	0.002	2.095^{***}	1.792^{***}	0.991^{***}	-0.003
After Colleague Retirement	(0.010) 0.027^{***} (0.006)	(0.011) 0.034^{***} (0.006)	(0.012) 0.028^{***} (0.007)	-0.006 (0.004)	(0.420) 0.889^{***} (0.181)	(0.373) 0.873^{***} (0.164)	(0.243) 0.328^{***} (0.106)	-0.004^{**} (0.002)
Month of Collegue Retirement *Retiree Male, Female Judge	0.015 (0.013)	(0.000) (0.011) (0.014)	-0.008 (0.016)	-0.009 (0.010)	(0.101) (0.010) (0.437)	(0.101) -0.024 (0.375)	-0.048 (0.246)	-0.001 (0.005)
After Colleague Retirement *Retiree Male, Female Judge	0.012 (0.010)	0.015 (0.009)	0.006 (0.010)	-0.011^{*} (0.005)	0.592^{**} (0.266)	0.522^{**} (0.227)	0.253^{*} (0.150)	-0.004 (0.002)
Month of Colleague Retirement *Retiree Female, Female Judge	-0.002 (0.024)	0.001 (0.026)	-0.024 (0.026)	-0.018 (0.015)	-0.585 (0.762)	-0.366 (0.644)	-0.241 (0.414)	0.001 (0.007)
After Colleague Retirement *Retiree Female, Female Judge	-0.006 (0.017)	0.002 (0.019)	0.002 (0.018)	-0.002 (0.008)	-0.050 (0.470)	0.052 (0.439)	-0.048 (0.247)	-0.001 (0.004)
Month of Colleague Retirement *Retiree Female, Male Judge	-0.046^{**} (0.019)	-0.048^{**} (0.021)	-0.069^{***} (0.023)	-0.014 (0.012)	-1.637^{**} (0.703)	-1.642^{***} (0.624)	-1.115^{***} (0.403)	-0.004 (0.006)
After Colleague Retirement *Retiree Female, Male Judge	-0.008 (0.014)	-0.010 (0.015)	-0.013 (0.015)	-0.003 (0.006)	-0.647 (0.409)	-0.599^{*} (0.360)	-0.318 (0.226)	(0.000) (0.003)
Dependent Mean (in Levels) Observations	$40.5 \\ 94,905$	$32.8 \\ 94,628$	$18.5 \\ 94,141$	$\begin{array}{c} 0.6\\94,\!141\end{array}$	40.2 95,656	$32.4 \\ 95,656$	$18.2 \\ 95,656$	$\begin{array}{c} 0.6\\94,\!628\end{array}$

Table A6: Impact of Colleague Retirement on Caseload and Allowance Rates by Gender Parity

Notes: Observation at the judge by month level. Sample restricted to the event study balanced panel and only includes judges who appear all 6 months before the retirement event, and all 9 months after are included. Judge, Hearing Office, year, and month fixed effects are included. Judge's gender is proxied by the genderedness of their first name. Names that are over 50 percent female in the SSA birth name records are labeled women. Standard errors corrected for clustering at the Hearing Office level. p<0.01 ***, p<0.05 **, p<0.1 *.

Table A7: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rates by Hearing Office Characteristics

		Natur	al Log of					
	Dispositions	Decisions	Awards	Allowance Rate	Dispositions	Decisions	Awards	Allowance Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Month of Colleggue Batirement	0.044***	0.047***	0.043***	0.004	1 765***	1 465***	0.740***	0.003
Month of Concague Recircinent	(0.008)	(0.008)	(0.010)	(0.005)	(0.315)	(0.273)	(0.181)	(0.000)
After Colleague Retirement	0.021***	0.027***	0.018***	-0.011***	0.774***	0.737***	0.183*	-0.005***
Inter concegue recencinent	(0.006)	(0.006)	(0.006)	(0.003)	(0.174)	(0.151)	(0.104)	(0.002)
Month of Colleague Retirement	0.001	0.002	0.004	0.003	0.065	0.063	0.097	0.002
*De-Meaned Unemployment Rate	(0.004)	(0.004)	(0.004)	(0.002)	(0.138)	(0.122)	(0.089)	(0.001)
After Colleague Retirement	-0.005**	-0.007***	-0.003	0.003***	-0.281***	-0.259***	-0.110**	0.001**
*De-Meaned Unemployment Rate	(0.002)	(0.002)	(0.003)	(0.001)	(0.075)	(0.070)	(0.050)	(0.001)
1.5	()	()	()	()	()	()	()	()
Dependent Mean (in Levels)	40.47	32.76	18.57	0.56	40.16	32.42	18.28	0.56
Observations	88,166	87,917	87,471	87,471	88.848	88.848	88,848	87.917
Month of Colloague Batirement	0.044***	0.047***	0.044***	0.004	1 761***	1 468***	0.769***	0.003
Month of Coneague Retirement	(0.008)	(0.008)	(0.010)	(0.004)	(0.308)	(0.271)	(0.180)	(0.002)
After Colleague Betirement	0.024***	0.031***	0.024***	-0.009***	0.833***	0.834***	0.276***	-0.005***
Anter Concagae Activement	(0.005)	(0.001	(0.005)	(0.003)	(0.155)	(0.140)	(0.082)	(0.009)
Month of Colleague Batiroment	0.000	0.000)	0.000)	0.000)	0.004	0.003	0.001	0.000
*Do Moanod Avorage Monthly Dispositions	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)	(0.002)	(0.002)	(0.000)
After Colleague Betirement	0.000***	0.000	0.000	0.000)	0.005***	0.004***	0.002)	0.000)
*De-Meaned Average Monthly Dispositions	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.004	(0.001)	(0.000)
De Meaner Average Montally Dispositions	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Dependent Mean (in Levels)	40.48	32.77	18.61	0.56	40.17	32.43	18.32	0.56
Observations	89,455	89,205	88,753	88,753	90,144	90.144	90,144	89,205
		· · · · ·	,	· · · · · ·	,	· · · · ·	· · · · ·	
Month of Colleague Betirement	0.045***	0.048***	0.045***	-0.004	1 789***	1 494***	0 780***	-0.003
Month of Concegue Rectionente	(0.008)	(0.008)	(0.010)	(0.005)	(0.306)	(0.270)	(0.179)	(0.002)
After Colleague Betirement	0.025***	0.032***	0.025***	-0.009***	0.865***	0.867***	0.300***	-0.005***
Inter concegue recencinent	(0.005)	(0.005)	(0.005)	(0.003)	(0.147)	(0.132)	(0.078)	(0.002)
Month of Colleague Retirement	-0.002	-0.001	-0.004	-0.003*	-0.155	-0.108	-0.071	-0.001*
*De-Meaned Number of Judges at Retirement	(0.003)	(0.003)	(0.003)	(0.002)	(0.096)	(0.083)	(0.057)	(0.001)
After Colleague Retirement	-0.001	-0.002	-0.003	-0.001	-0.023	-0.033	-0.015	-0.001
*De-Meaned Number of Judges at Retirement	(0.002)	(0.002)	(0.002)	(0.001)	(0.049)	(0.041)	(0.026)	(0.001)
_ =	(0.002)	(0.002)	(0.00-)	(0.001)	(0.010)	(0.011)	(0.020)	(0.001)
Dependent Mean (in Levels)	40.48	32.77	18.61	0.56	40.17	32.43	18.32	0.56
Observations	89,455	89.205	88,753	88,753	90,144	90.144	90,144	89,205
	,	,		,	1	1	1	,

Notes: Observation at the judge by month level. Sample restricted to the event study balanced panel and only includes judges who appear all 6 months before the retirement event, and all 9 months after are included. Judge, Hearing Office, year, and month fixed effects are included. Judge's gender is proxied by the genderedness of their first name. Names that are over 50 percent female in the SSA birth name records are labeled women. Judge's race is proxied by the racial composition of their surname using US Census data that reports what fraction of the population with a given last name is White, Black, or Hispanic. Standard errors corrected for clustering at the Hearing Office level. p<0.01 ***, p<0.05 **, p<0.1 *.

			Natural Log of			
	Decisions to Dispositions	Awards to Dispositions	Fully Favorable	Partially Favorable	Months Wait	Share Cases
	Rate	Rate	Rate	Rate	for Hearing	by Video
	(1)	(2)	(3)	(4)	(5)	(6)
Month of Colleague Betirement	0.001	-0.003	-0.006	-0.021**	-0.005*	0.008
Month of Concague Recification	(0.002)	(0.006)	(0.006)	(0.009)	(0.003)	(0.031)
After Colleague Betirement	0.007***	-0.003	-0.010***	-0.014**	0.003	0.011
	(0.001)	(0.003)	(0.004)	(0.006)	(0.004)	(0.032)
Dependent Mean (in Levels)	0.80	0.45	0.49	0.09	14.67	22.20
Observations	89,205	88,753	88,417	69,668	11,856	10,994
	Decisions to Dispositions	Awards to Dispositions	Fully Favorable	Partially Favorable	Months Wait	Share Cases
	Rate	Rate	Rate	Rate	for Hearing	by Video
Month of Colleague Retirement	-0.000	-0.002	-0.002	-0.001	-0.064*	0.256
	(0.002)	(0.002)	(0.002)	(0.001)	(0.037)	(0.330)
After Colleague Retirement	0.005***	-0.002	-0.005***	-0.000	0.039	-0.051
	(0.001)	(0.001)	(0.002)	(0.001)	(0.053)	(0.299)
Dependent Mean (in Levels)	0.80	0.45	0.49	0.07	14.67	20.59
Observations	89,455	89,455	89,205	89,205	11,856	11,855

Table A8: Impact of Colleague Retirement on Other Judge- and Office-level Outcomes

Notes: Observation at the judge by month level for columns 1-4. Each point is an event study coefficient obtained from estimating equation (4). Observation at the Hearing Office by month level for columns 5-6. Each point is an event study coefficient obtained from estimating equation (3), where the outcome is the average office level wait time for appeals and the share of cases held over video. Cases for all judges (including the retiring judge) are included in the measure. Sample restricted to the event study balanced panel and only includes judges who appear all 6 months before the retirement event, and all 9 months after are included. Judge, Hearing Office, year, and month fixed effects are included. Standard errors corrected for clustering at the Hearing Office level. Partially favorable decisions occur when the ALJ awards disability, but adjusts the eligibility criteria. For example, they might set a later disability onset date than the date alleged by the claimant. p<0.01 ***, p<0.05 **, p<0.1 *.

Online Appendix B: Robustness and Heterogeneity Analysis

We adjust our baseline model to explore both heterogeneity and robustness. We highlight some of these results in the main text, but outline them here in full detail.

Robustness

We first probe the sensitivity and robustness of our estimates. It is possible that local economic conditions are changing in ways that lead the judges to reach different verdicts. In Panel A of Table A1 we control for labor market conditions and demographic characteristics in the hearing office catchment area. To do this we link county demographics and labor market measures from the Quarterly Census of Employment and Wages and population measures from the Surveillance, Epidemiology, and End Results Program to hearing offices using county-to-zip code and zip code-to-hearing office crosswalks. Effects are similar, with a 3.2 percent increase in decisions and 1.1 percent decrease in allowance rates after the peer retirement. Estimates are insensitive to changes in the sample of judges we include (Table A1). We estimate a significant increase in dispositions and decisions (between 2) and 4 percent) and a significant decrease in allowance rates (between 0.8 and 1.3 percent) when we include retirement events from National Hearing Centers, include any judge with at least one pre- and post- event observation in the hearing office (remove the balanced panel restriction), or limit the sample to retirement events that do not overlap. It is possible that we are mislabeling dismissed judges as retirees. This would be a concern if a peer judge dismissal affects other judges' behavior. The results are robust to excluding retirement events where the retiree has a known review by the MSPB board, meaning the retirement might have been non-voluntary.²⁴ The estimates are also robust to various sets of fixed effects to control for office specific shocks or trends. Year-by-month or retirement event fixed effects vield similar results.

Estimates are robust to transformations of the outcome. If we estimate the trends in levels, or use a inverse hyperbolic sine transformation rather than the natural log transformation (so that outcomes are defined when equal to zero), we estimate a similar 2.3 percent increase in dispositions, 3 percent increase in decisions, and a 0.6 percent decrease in allowance rates (see Figure A2). Results are also robust to extending the number of pre and post periods. Increasing the number of periods increases the amount of overlap between event windows, potentially confounding effects, but also changes the sample composition due to the balanced panel restriction. If we include 9 pre periods and 18 post periods and exclude retirement events within 9 months of each other, we see similarly flat pre-trends, followed by a spike in caseloads after the retirement that lasts 6 months, with some significant effects even out to 15 months. This could help explain the lower allowance rates observed 8-9 months after the retirement. Allowance rate estimates are similar but less precise, returning to previous levels by 10 months out.

Because SSA undertook various ALJ reforms aimed at standardizing outcomes across

²⁴Estimates for the disciplined judge retirements are less precise due to the small sample size, but the point estimates on dispositions and decisions are large and positive, and there is a significant 7 percent drop in allowance rates of the remaining judges.

judges beginning in 2010 and 2011, we also examine how the effects of a peer judge retirement differ over time. As part of this reform, judges with exceptionally high or low allowance rates received additional training. Historical allowance rates of retiring judges were similar across the panel, but did tend to be higher prior to 2015. We estimate effects separately for retirement events from 2011-2014 and 2015-2018 and observe increases in decisions between 2011 and 2014, but larger increases in dispositions and decisions and large, significant 1.8 percent reductions in allowance rates between 2015 and 2018, suggesting this pattern is not driven by the 2010-2011 ALJ reforms (Table A2).

There are several factors that might contribute to the heterogeneous effects on allowance rates. First, the increase in workload over the next nine months is significantly smaller before 2014 relative to 2015 and later. This is consistent with workload affecting allowance rates. Second (and perhaps related to the first point), limited funding and dwindling candidate pools from the ALJ registry, has led to increasingly understaffed Hearing Offices and longer backlogs and processing times that peak in 2016. As such, a judge retirement in the second half of the sample might have been more disruptive to the office environment than it was in earlier years. Third, SSA introduced the Compassionate and Responsive Service (CARES) Plan in 2016 that aimed to reduce the number of pending hearings and average hearing wait times. This increased focus on reducing backlogs could amplify the stress associated with taking on a retiring judge's workload. Finally, as part of the CARES Plan, SSA worked with the Office of Personnel Management (OPM) to refresh the qualified ALJ registry and hire new ALJs. SSA reports that they have hired over 600 ALJs since 2016. As we see in Table A4, the allowance rate effects are largest among less experienced ALJs, so this shift in the composition of judge experience could lead to the heterogeneous effects in Table A2.

We also test sensitivity to new concerns about twoway fixed effects models (Callaway and Sant'Anna, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2020). Our estimates are robust to concerns that it is impossible to separately identify a linear pre-trend from dynamic treatment effects when both judge and period fixed effects are included (Borusyak and Jaravel, 2020). If we estimate a model similar to equation (1), but only include preperiod months and omit both the first and last month indicators, we see the pre-trends are flat, with no significant trend (top panel, Figure A7). Only one point (log awards in month t-2) is significantly different from zero. Building on this, if we impose flat trends during the pre-period, allowing us to separately identify dynamic treatment effects, we estimate a similar pattern of effects, with a significant 5 percent increase in dispositions and decisions for six months, and a significant 1 percent decline in allowance rates (Figure A7).

Our results are similarly robust to issues of negative weighting that arise when using twoway fixed effects difference-in-differences and event study models with staggered treatment timing. We implement the Sun and Abraham (2020) interaction-weighted estimator and find patterns that are nearly identical (Figure A8). This is perhaps not surprising, our main coefficients essentially report average outcomes each month after removing the judge by office specific mean and year and month level differences.

As another way to account for potential, underlying secular trends in caseloads and allowance rates and eliminate negative weights we estimate an alternative model, similar to Deshpande and Li (2019), as suggested by Goodman-Bacon (2021). We adopt this approach, as the type of hearing offices that experience a retirement might be different than those that do not. First, we create a dataset for each retirement event. In each dataset, judges in the hearing office that is experiencing the focal retirement event are labeled as treated, while judges in hearing offices that experience a retirement event exactly one year later are labeled as control. We then stack each of these datasets to examine outcomes in the treatment versus control (future treated) hearing offices.²⁵ With in a given retirement event "experiment" the period-by-period month and year for the treatment and control units will be the same, allowing us to difference out secular trends. We then compare monthly judgelevel disposition, decisions, and allowance rate trends in the focal retirement event, relative to monthly judge-level outcomes during the same time period in the hearing office where there will be a retirement exactly one year later as follows

$$Y_{jot} = \sum_{\tau=-5}^{9} \beta_{\tau} Treated_{ot} * (\tau \text{ months from retirement})_{ot} + \sum_{\tau=-5}^{9} \alpha_{\tau} (\tau \text{ months from retirement})_{ot} + \nu Treated_{ot} + \delta_j + \theta_e + \phi_o + \varepsilon_{jot}$$

$$(4)$$

The outcomes are the same as above, but now the β_{τ} trace out the change in dispositions, decisions, awards, and allowance rates for judges in the treated hearing office, relative to judges that will experience a retirement in one year. We include the fixed effects for each dataset (θ_e) to make this a comparison between the treatment and control from each dataset during the same time period, as well as judge, office, and panel period fixed effects. If hearing offices that experience a retirement within a short period of time exhibit similar unobservable trends, this can capture any spurious secular trend, not captured in equation (1).²⁶

As seen in Figure A9 judges in the treated hearing office experience an increase in dispositions and decisions after the retirement. This increase is sustained for three months, and there is a significant difference in the t-5 month, but the overall pattern is consistent. There is also a drop in the allowance rate starting around month two, that is even more pronounced than in the baseline model. Results from the corresponding parameterized difference in differences model are also similar (Table A3). The retirement leads to a 4.3 percent increase in dispositions, a 5 percent increase in decisions, and a 1.1 percent reduction in allowance rates. If we further include in the bottom panel month and year fixed effects (which can only be separately identified because we have treated and untreated units in each panel) the effects are similar in magnitude and significant.

Heterogeneity

We also explore heterogeneity by remaining judge characteristics, retiring judge characteristics, and limited characteristics of the hearing office, but see few patterns of heterogeneity. To do this, we estimate a variant of equation 2, as follows

 $Y_{jot} = \beta_0 Month of Colleague Retirement_{ot} + \beta_1 After Colleague Retirement_{ot}$ $+ \beta_3 Month of Colleague Retirement_{ot} * Char_{jo} + \beta_4 After Colleague Retirement_{ot} (5)$ $* Char_{jo} + \beta_5 Char_{jo} + \delta_j + \phi_o + \gamma_{yr} + \psi_{mo} + \varepsilon_{jot}$

²⁵As such, retirement events will occur multiple times, once as a treated unit, and potentially multiple times as a control.

 $^{^{26}}$ We pick control events that are exactly one year away to keep the sample size managable. Results are similar if we increase the range of potential controls to include more events.

Where $Char_{jo}$ is a judge, retiree, or office specific characteristic. The β_3 and β_4 coefficients will allow us to test if impacts vary by the judge's characteristics (such as gender, historical allowance rate, or tenure), the retiree's characteristics, or office-level characteristics (such as office caseload or local labor market conditions). For time invariant characteristics, such as gender, the direct effect of $Char_{jo}$ will be absorbed by the fixed effects.²⁷

By Judge's Characteristics. We explore how effects vary by the judge's characteristics to test if certain groups are more responsive in Table A4. Both male and female judges experience an increase in dispositions and decisions and a decrease in allowance rates after the colleague's retirement. However, female judges experience a larger increase in decisions (4.5 vs 2.7 percent) and a larger reduction in allowance rates after the retirement (1.5 vs 0.6 percent). The effect on allowance rates for women is two and a half times as large as the effect for men, but the difference is only significant at the ten percent level. There are no differential impacts by the judge's race. Judges who are relatively new to the SSA judicial corps (less than one year) are significantly impacted. They experience a large increase in caseload and decisions after the retirement (15.7 and 15.8 percent) and an additional 6.3 percent reduction in allowance rates.²⁸ This might simply be driven by the fact that new judges make fewer decisions, so there is more volatility in allowance rates for new judges. However, the pattern is consistent with an increase in caseloads leading to lower allowance rates.²⁹

Judges that have above average allowance rates in their office are also differentially impacted. After a colleague's retirement, a judge with a 10 percent higher allowance rate reduces dispositions and cases by about one percent, and significantly reduces awards and allowance rates by about three percent. This pattern of impacts on allowance rates is consistent with mean reversion.

By Retiree's Characteristics. As an alternative to caseloads, peer effects could also affect remaining judge allowance rates. However, we see little evidence of retiring judge characteristics impacting the allowance rate response in the nine months following retirement. We explore how effects vary by the retiring judge's characteristics in Table A5 to understand if peer effects play a role. There is a marginally significant differential reduction in dispositions and allowance rates in the month of retirement if the retiring judge is female. However, there is no differential effects after the retirement. If the retiring judge is non-white, there is no differential changes in dispositions and decisions or awards. If the retiring judge has less than one year experience in the judicial corps, there is no significant change in dispositions or decisions, but an additional 2.4 percent reduction in the remaining judges' allowance rates. These retirements would plausibly be more unexpected than a typical retirement, meaning

²⁷We have tried linking judge surnames to the Census Bureau's Census Surname file to probabilistically determine the judges' race, but are concerned about error in this racial matching. There are no significant patterns by our race measure.

 $^{^{28}}$ Judges that have been with the SSA for 6-12 months exhibit a similar 12 percent increase in dispositions and decisions and a 5.6 percent reduction in allowance rates, suggesting this is not driven by judges that might have been preemptively hired to replace the retiree.

²⁹These patterns have potential implications for judge fixed effects research designs that use judge leniency to instrument for case outcomes. Workplace disruption that affect judge allowance rates heterogeneously might temporarily re-ordered judge leniency, violating monotonicity (Kling, 2006). Researchers might need to rely on weaker assumptions such as average monotonicity (Frandsen et al., 2019) or estimate interacted heterogeneous effect models (Mueller-Smith, 2015) to estimate local average treatment effects.

there is likely not a replacement judge lined-up and more case-shifting (see Table A5).³⁰ There are no differential impacts by the retiring judge's allowance rate prior to retirement, suggesting this is not driven by peer judge strictness or leniency.

These patterns do not appear to be consistent with the retiring judge's behavior influencing the peer judges. We explore potential peer effects further in Tables A6. While there are no clear patterns when exploiting the detailed race heterogeneity, there are some potential patterns by judge and retiree gender parity. Remaining female judges make marginally more decisions and have lower allowance rates after a male judge retires. Remaining male judges have fewer dispositions, decisions, and awards after a female judge retires.

By Hearing Office Characteristics. Finally, in Table A7 we explore how effects vary by the few hearing office characteristics available to see if the retirement event has an interactive effect with the environment or setting. Applicants whose disability is not on the pre-defined listing of impairments, can still be eligible if they meet vocational criteria. These criteria measure the applicants ability to find work or re-train. As such, economic conditions at the time of review can influence a judge's determination. Judges in hearing office catchment areas with a higher unemployment rate (worse labor market opportunities for applicants) have relatively fewer dispositions and decisions after the retirement and have slightly higher allowance rates. Judges in hearing offices that hear more dispositions on average see a slightly smaller increase in caseloads, but no differential impact on allowance rates. The number of judges pre-retirement does not significantly affect the caseload and allowance rate response. This is because there is significant variation across retiring judges in caseloads, conditional on office size. If instead we explore differences by the retiree's caseload size, the effect on allowance rates is more negative for retiree's with higher caseloads (Figure 3). The impact of a peer retirement does not appear to significantly vary with the local setting.

Other Potential Outcomes. We explore several other outcomes observable in the data. Decisions as a share of dispositions increase slightly, suggesting fewer dismissals, consistent with redistributed cases already being screened for technical errors. But, awards as a share of total dispositions do not decline, suggesting there is not an increase in cases that would have been awarded disability instead being dismissed. The falling allowance rates seem to be driven by both a fall in fully-favorable cases and partially favorable awards (where the judge does not accepted the alleged onset date), although the result is stronger for fully-favorable decisions. (Table A8). We also explore office level measures of appeal wait time and remote video cases. Judge retirements do not significantly affect the average wait time of appeal cases or the share of office level cases held by video.

 $^{^{30}}$ If we limit the sample to retirement events where the retiree's average caseload was no more than 10 percent lower 1-3 months before the retirement relative to 4-6 months before retirement (to avoid predictable transitioning cases) we find larger caseload effects and a slightly larger decrease in allowance rates.