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

Clayson Shumway   Riley Wilson†

June 21, 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.

†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 when there is a disruption in the work environment? Does a judge's workload affect their decision-making? Considerable research has explored how an individual's or a judge's characteristics affect eventual ruling and case outcomes,[1] but changes in the judge's work setting might also impact judge's 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, judicial vacancies in the criminal justice system not only increase caseloads, but change the composition of cases judges hear, 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 decision-making 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 2018 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

[1] For 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.

impacts on judges' allowance rates. After the retirement, within judge allowance rates fall by one percent, suggesting the workplace disruption impacted case outcomes. Similar patterns are observed when we control for local conditions in the hearing office catchment area or when we compare patterns in hearing offices with a retirement to patterns in hearing offices that have a retirement one year later, to control for secular trends, similar to Deshpande and Li (2019). 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. Subgroups that experience the largest increase in cases and decisions (female judges and judges with less experience) also exhibit significant 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.

These patterns of heterogeneity also have potential implications for the interpretation of empirical strategies that rely on cross-judge differences in leniency (i.e., judge fixed effects strategies). If environmental factors (such as the retirement of a peer judge) have heterogeneous impacts on remaining judges, the monotonicity assumptions required to generate interpretable, local average treatment effects might not hold. This underscores existing work showing that defendant and case characteristics can invalidate the monotonicity assumption for criminal cases (Frandsen et al., 2019; Mueller-Smith, 2015), suggesting that judge environmental factors should also be considered.

Workplace changes that increase caseloads and reduce allowance rates have important implications for the size and nature 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 of results suggest that caseload increases impact judgements and that policies that increase SSA judicial corps stability would result in more consistent case outcomes.

2 Conceptual Relationship Between Environment, Workload, and Decision Outcomes

A court decision depends on the claimant's characteristics and case details (δ_i) and the judge's own characteristics and preferences (θ_j). However, characteristics of the judge's work environment that potentially vary over time (α_{ot}) could also affect case outcomes. In particular, we will consider judge decisions on appeals of disability insurance applicants. The individual will be awarded disability if the combination of these characteristics puts the case decision-rule past some threshold γ, as follows

$$\textit{Disability Awarded}_{ijot} = \begin{cases} 1 & F(\delta_i, \theta_j, \alpha_{ot}) \geq \gamma \\ 0 & \textit{otherwise} \end{cases} \tag{1}$$

Most work has focused on the role of individual or case characteristics (δ_i) or judge characteristics (θ_j). A growing literature suggests that time-varying environmental factors such as temperature, upsetting sports loses, media coverage, or judge peer composition can also influence judges' rulings (Eren and Mocan, 2018, 2020; Heyes and Saberian, 2019; Lim et al., 2015). If the decision rule is elastic with respect to α_{ot}, it is possible two individuals with similar characteristics (δ_i) could receive different outcomes from the same judge. Unlike some of the previous work, we are interested in understanding how policy relevant features such as judge workload and workforce dynamics influence judge behavior, and ultimately applicants' well-being and long-run outcomes. From a policy perspective, understanding how environmental factors like workload affect applicant outcomes can shed light on the importance of program structure, financing, and management.

Work spread across criminology, law, and economics finds that for criminal and civil judges, caseloads affect case outcomes, the number of cases heard, and the likelihood of overruling civil appeals (Beenstock and Haitovksy, 2004; Engel and Weinshall, 2020; Huang, 2011; Ulmer and Johnson, 2004). There is also evidence court caseloads impact the composition of Chapter 11 bankruptcy court filings (Iverson, 2018). However, in many judicial settings, it is not possible to isolate the impact of workplace changes (a component of α_{ot}) on judge decisions because the number and composition of cases (a component of δ_i) might endogenously respond. As Yang (2016) documents in the criminal justice system, judge vacancies lead to heavy caseloads, which induce prosecutorial changes in the number and types of cases brought before judges. As such, any observed change in judge decision could be due to a change in the composition of cases in addition to a direct effect of the judge's workload. Isolating the impact of workload on judge decisions is not possible if the composition of cases changes. We present the only work of which we are aware 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 judges in the SSA judicial corps, a body of approximately 1,500 judges that make over half a million decisions a year.

3 The 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 available to workers with a sufficient work history and evidence of a permanent disability that inhibits the individual from participating in substantial gainful activity. SSI does not have the same work history requirement and is available to the elderly, blind, and disabled with limited work history, income, and resources. Between 2.1 and 2.8 million workers apply for SSDI, SSI, or both each year. Applications are submitted to a local field office, which verify non-medical requirements (e.g., work history, income limits) are met, then send the application to a state-run Disability Determination Services (DDS) agency that evaluates medical criteria and either allows or denies 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. If the individual is denied again, they can ask for a hearing before an Administrative Law Judge (ALJ). At this point the claimant (or their representation) will present their case and any new evidence to the ALJ. In 2010, 19 percent of all applicants went through the appeal process. With an appeal award rate of approximately 60 percent, nearly one-third of all disability awardees are approved through the appeals process (Social Security Administration, 2019).[2]

A hearing request will be assigned to one of approximately 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 having highest priority in the queue.[3] Relevant to our current exercise, if an ALJ retires, their pending cases will be reassigned to the remaining judges so that cases will be heard in a timely manner. If the ALJ denies the claim, the individual can ask that it be reviewed by the Social Security's Appeal Council, who will either deny, approve, or send the case back to an ALJ. Across hearing offices and over time the number of ALJs ranges from 3 to 20 per office. Between 2010 and 2019 there were approximately 1,505 judges at any point in time. These judges average 38.3 dispositions and 31.2 decisions a month, with a favorable decision 56 percent of the time (Table 1). ALJs are highly specialized, they must be a licensed attorney with at least seven years of experience to be eligible (Walker, 2016).

A large literature shows the SSDI application process and SSDI receipt severely reduces labor supply of applicants (Autor and Duggan, 2003; Autor et al., 2016, 2017; 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

[2]Workers can simultaneously apply for SSDI and SSI. Approximately 51 percent of applicants that reached the DDS review were applying for both SSDI and SSI. Joint applicants are less likely to be allowed at all stages.

[3]There 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 will be prioritized in the queue.

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]

SSA ALJ caseloads are similar to other judges in the judicial system.[5] SSA expects full-year judges to complete 500 dispositions a year, but no more than 700. However, there is substantial variation across judges. During our analysis period, 54 percent of judges held over 500 dispositions and 14 percent made over 600 dispositions. There are even rare situations where judges made over 1,000 dispositions in a year. SSA judges also typically face a long backlog of pending cases. From September 2010 to December 2016, the number of system-wide pending cases increased from 700,000 to over 1.1 million before starting to drop off.[6] This backlog leads to long appeal wait times of 14 months on average. We are interested in documenting how workplace changes that affect judge caseloads impact judge decisions.

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] For each judge we observe their name and hearing office as well as the number of dispositions, decisions, allowances, denials, and partially or fully favorable allowances they have made during the month. Dispositions are reported cumulatively through the administrative year (starting in October), so we difference cumulative totals from month-to-month to back-out the monthly number of 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.

[4]At 2018 NBER Summer Institute, Nicole Maestas noted preliminary work on ALJ training reforms in 2011, but this is not released.

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

[6]Pending 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]Some judges are assigned to multiple hearing offices at the same time so they may have multiple observations for the same month.

The monthly judge/hearing office data allows us to construct the set of available judges at each hearing office in every month and to identify retiring judges. When a judge leaves a hearing office and does not reappear in the data at any hearing office, we label that event as a retirement. Individuals that move to a different hearing office or that experience an extended leave before reappearing in the disposition data are not counted as retirements. We exclude retirement events that occur before April 2011 and after August 2018 (to insure adequate pre- and post-observation) as well as retirements at national hearing centers, which do not have a geographic catchment but rather 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, although there is a slight upward trend (Figure A1). The allowance rates of retiring judges are nearly identical to the those of all other judges during the same time period (Table 1).

ALJs that are dismissed by the SSA or reprimanded and then leave voluntarily would also be captured in these retirements. This is potentially problematic if judge dismissal is due to workplace conduct that could impact 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 there is "good cause" (Office of the Inspector General, 2014). A judge's decisions alone do not qualify as "good cause" (Office of the Inspector General, 2012). However, SSA can engage in targeted counsel and training if a judge's allowance rate is inconsistent with other judges in the office. In fact, in 2010 and 2011, SSA undertook a comprehensive review of ALJ decisions to standardize judge procedure. Part of this review included training for judges with allowance rates that were too high or too low relative to other judges in the hearing office. Disciplinary action against ALJs that result in retirement or removal is uncommon. Between 2007 and 2016, there were 98 reprimands,[8] 34 proposed suspensions, and 16 proposed removals (Government Accountability Office, 2017).

[8]A reprimand is a formal written warning that stays on the judge's record for one year.

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. Our results are similar if we exclude judges with known dismissals or reprimands.

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 the judge's gender. We link each judge's surname to the Census Bureau's Census Surnames file to probabilistically determine the judge's race or ethnicity (Black, Hispanic, or White). 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). We do not observe case-level measures, so we cannot explore how the characteristics of cases change after the retirement event.

In some specifications we also control for labor market conditions 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 (QCEW) and population measures from the Surveillance, Epidemiology, and End Results Program (SEER) to hearing offices using county-to-zip code and zip code-to-hearing office crosswalks.

5 Empirical Strategy

As noted above, appeals are quasi-randomly assigned to ALJs and ALJs do not specialize in particular types of cases. Also, given the enormous backlog in appeal cases (over 1.1 million in 2016) across hearing offices, cases take 14 months on average to be processed. Given these two characteristics, 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. As such, we are able to explore the impact of judge vacancies and caseloads on judge decisions, which cannot be separately identified when examining vacancies in the criminal or civil court system (Yang, 2016). We use an event study framework to explore the impact of peer judge retirement on remaining judges' cases, decisions, and allowance rates (share of favorable decisions). 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 an event study panel for each of the judges in the retiree's hearing office. We estimate the impact of the retirement on cases and allowance rates as follows

$$Y_{jot} = \sum_{\tau=-5}^{9} \beta_\tau (\tau \textit{ months from retirement})_{ot} + \delta_j + \phi_o + \gamma_{yr} + \psi_{mo} + \varepsilon_{jot} \tag{2}$$

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. We estimate models measuring outcomes in both levels and natural log specifications. Our preferred specification is the natural log transformation as there is substantial variation across judges in the monthly number of dispositions, decisions, and allowance rates and the natural log specification allows us to estimate percent changes in outcomes. Our coefficients of interest are the β_τ coefficients which trace out a judge's monthly outcomes six months prior to the peer judge's retirement and nine months after. Some hearing offices experience multiple retirement events. To limit overlap between retirement event panels within the same hearing office we limit the retirement event panel to six months prior to the month of retirement, the retirement month of retirement, and nine months after the retirement.[9] 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. This framework allows us to explore trends in the pre-period as well as treatment effect dynamics for nine months. 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 year and month fixed effects control for potential secular and

[9] Results are consistent if we construct longer panels that overlap more. With our baseline specification there is still some overlapping panels. In Table A1 we show that the results are similar if these overlapping panels are excluded.

seasonal trends in caseloads and allowance rates. Standard errors are corrected for clustering at the hearing office-level. To avoid compositional changes, we limit the sample to judges who appear in the hearing office for all 16 months of the panel. Judges that join the hearing office to replace the retiree are not included so we can estimate the impact on the pre-existing set of judges.[10]

As seen in Table 1, judges in the event study panel have similar disposition counts, decision counts, allowance rates, racial composition, 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.

This specification only includes event panels for judges around a retirement event, so unlike a difference-in-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-specific, year-specific, and month-specific effects, dispositions, decisions, and allowance rates within a given hearing office would not have deviated if the interruption had not occurred. For this reason an examination of 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. If we relax this identifying assumption by estimating event study models using judges in hearing offices that will experience a retirement in the future as a control group (similar to Deshpande and Li (2019)) effects are similar.

[10]In Appendix Table A1, we show that the average effects are similar if we relax this restriction and instead include any judge who holds at least one disposition both before and after the retirement event.

We also estimate the following parameterized difference models:

$$\begin{aligned} Y_{jot} = &\beta_0 \textit{Month of Colleague Retirement}_{ot} \\ &+ \beta_1 \textit{After Colleague Retirement}_{ot} + \delta_j + \phi_o + \gamma_{yr} + \psi_{mo} + \varepsilon_{jot} \end{aligned} \tag{3}$$

This allows us to identify the average effect of peer judge retirements on remaining judges' outcome 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 the two separate time periods, for the exact month of the retirement and for 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 tells us about the average impact of a peer retirement on caseload and allowance rates over the next nine months. As with the event study, we also employ the Deshpande and Li (2019) approach to verify these effects are not driven by secular trends.

The impact of a peer-judge retirement could impact judges heterogeneously. In the appendix we report estimates that allow the effects in equation (3) to vary with the judge's characteristics, the retiree's characteristics, or other workplace and environmental conditions, but we find limited evidence that retiring judge's gender, race, historic allowance rate, or tenure impact 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

Before exploring the impacts on judge outcomes, 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

$$Number\ of\ Judges_{ot} = \sum_{\tau=-6}^{9} \beta_\tau(\tau\ months\ from\ retirement)_{ot} + \phi_o + \gamma_{yr} + \psi_{mo} + \varepsilon_{ot}. \quad (4)$$

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 (4) 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. In the first month after the retirement, the number of judges drops, significantly by 0.84 judges and remains lower, but slowly increases until reaching 0.5 judges lower nine months after the retirement.[11] 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.

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 2 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. Dispositions stay significantly higher for 6 months before dropping slightly and becoming no longer statistically significant. Decisions follow a similar pattern, suggesting the remaining judges face a higher workload, holding 5 percent more dispositions and making 5 percent more decisions. The point estimates for the number of awards increase by a slightly smaller amount.

In Panel B of Figure 2 we plot the event study coefficients on the natural log of the judges' monthly allowance rate (awards/decisions). Allowance rates hold constant prior to the retirement, and then start to drop around two months after the retirement event. These

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

estimates are less precise, and only significant at the five percent level in three of the nine months. However, if we pool observations into two month bins for power, the drop in the allowance rate is more precisely estimated (Appendix Figure A2). The impact of a judge retirement on office-level allowance rates is similar, suggesting the change in allowance rates is a treatment effect, and not simply driven by a change in the composition of cases or judge case pairings (see Appendix Figure A3).[12]

In Table 2 we report the average effect in the month of retirement and after retirement. Dispositions and decisions increase by 4.6 and 5.0 percent respectively in the month of retirement and 2.9 and 3.7 percent on average over the next nine months after the retirement. At the mean, this corresponds to about one additional disposition or case in the month. Average daily dispositions and decisions are around 1.6-1.8, suggesting a little more that half a day's additional disposition and decision work in a month. In the nine months following 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 the retirement. After a peer judge retires, the remaining judges face higher caseloads and make more decisions, but overall the cases they review are less likely to be awarded disability insurance. Since cases are quasi-randomly assigned, this drop is not due to a compositional change in the judge's caseload.

We further explore the role of caseload on allowance rates in Figure 3. For each judge in our analysis sample, we calculate the average change in log dispositions and log allowance rates from the 6 months before and 9 months after the retirement. We then plot this relationship in a bin-scatter plot with the percent change in dispositions (caseload) in 25 percent bins on the x-axis and average percent changes in allowance rates on the y-axis. Consistent with our results in Table 2, judges that experience a larger percent increase in disposition caseloads, report a larger average drop in allowance rates. This appears to be driven

[12]The estimated effect on allowance rates from the office-level equation analogous to the difference estimator in (3) is -0.0153 (standard error of 0.003) in period zero and -0.0075 (standard error of 0.003) in the post period implying a 0.75 percent drop in office-level allowance rates.

by judges that experienced a large change in caseloads. But, when we focus on moderate changes in caseloads where most judges in the sample fall, larger increases in caseloads are still associated with a reduction in allowance rates (Appendix Figure A4).

6.3 Extensions

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

Robustness Results are robust to various sample restrictions and controls (Appendix Table A1) and functional form (natural log versus inverse hyperbolic sine) (Appendix Table 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 (Appendix Tables A3 and A4). Results are also consistent if we compare judges in hearing offices that experience a retirement to judges in hearing offices that experience a retirement exactly one year later, to account for potential, underlying secular trends in caseloads and allowance rates, similar to Deshpande and Li (2019) (Appendix Table A5). The estimates are also robust to pre-trend specification tests (Appendix Figures A5 and A6) (Borusyak and Jaravel, 2020).

Heterogeneity and Peer Effects As seen in Appendix Table A6, effects do not vary significantly by judge's race. 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 experience also observe higher caseloads and lower allowance rates after the peer retirement. These patterns are consistent with an increase in caseload negatively impacting allowance rates. 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, potentially a dismissal (see Appendix Table A7).[13]

7 Implications for the Disability Program

These patterns have potential implications for the size of the SSA disability insurance program. During our analysis sample, judges make 32.4 decisions a month. A one percent lower allowance rate after the retirement would suggest that 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. At average monthly benefits of $1,233.70 for disabled workers, this would imply nearly $246 million of forgone annual benefits. Claimants denied after a judge retirement might be more negatively selected than the average awardee. However even at the median ($1,135.00) or 25th percentile benefit ($900) this would still imply $226 or $179.5 million of forgone annual benefits respectively (Social Security Administration, 2019).

Higher caseloads could also increase applicant wait times. SSA reports that annually nearly 7,800 people die waiting for an ALJ decision (Gross et al., 2018). Using monthly, hearing office-level measures of applicant wait times, we estimate the impact of a judge retirement on wait times similar to equation (4). After a judge retirement, office-level wait times trend upward to approximately 0.1 additional months, but are imprecisely estimated (Appendix Figure A7). Using the national annual count of applicants that die waiting for an ALJ disability insurance decision and national average wait times across all hearing offices (Gross et al., 2018), we estimate that between 2010 and 2017 one additional month of waiting is associated with 546 additional deaths. Based on this back of the envelope calculation, the

[13]We also explore judge/retiree gender parity, and detailed retiree race categories, but see limited differences (Appendix Tables A8 and A9). We also examine differences by office-level characteristics in Appendix Table A10.

wait time point estimates associated with judge retirements would lead to an insignificant point estimate of 55 additional deaths each year if applied across all hearing offices. The increased caseload associated with retirements does not significantly increase applicant wait times or deaths while waiting for SSDI appeals.[14]

8 Implications for Judge Fixed Effects Estimation Strategies

Heterogeneous responses across judges to a peer judge retirement has potential implications for estimation strategies employing a judge fixed effects design. The judge fixed effects estimation strategy exploits differences in judge strictness and leniency (the θ_j's from equation (1)) to isolate exogenous variation in case outcomes and instrument for a claimant, applicant, or defendant, receiving a treatment. This strategy has been used to uncover a range of economic relationships, including the effects disability insurance receipt on family take-up of welfare (Dahl et al., 2014), and disability insurance receipt on labor force participation (French and Song, 2014; Maestas et al., 2013).[15] For this instrumental variables strategy to identify a local average treatment effect (LATE), both the exclusion restriction and monotonicity must hold. Monotonicity requires that if judge A is more likely to award claimants disability than judge B, every individual awarded disability by judge B would have been awarded by judge A if they had handled the case (Frandsen et al., 2019; Kling, 2006). In other words, the leniency of judges can be ordered. If a change in the workplace environment (α_{ot}) interacts with judge characteristics ($\frac{\partial^2 F(\delta_i,\theta_j,\alpha_{ot})}{\partial\theta_j\partial\alpha_{ot}} \neq 0$ from equation (1)), leading to heterogeneous impacts on allowance rates ("judge leniency") the leniency of judges might be re-ordered temporarily. This would violate monotonicity. We find that when a peer judge retires, remaining judges' allowance rates fall, with larger impacts for judges that have less experience or that are female. These heterogeneous effects could violate monotonicity,

[14] Autor et al. (2017) find that an increase in time waiting for SSDI benefits reduces labor force participation and annual earnings. Because the increase in wait time is small, this likely would have negligible impacts on labor force participation of applicants.

[15] Frandsen et al. (2019) list 24 papers that use judge fixed effects to explore a range of topics, including incarceration, pre-trial detention, bankruptcy, foster care, patents, and disability insurance.

meaning the estimated effects from the IV do not identify a LATE. In cases like this, researchers might need to rely on weaker assumptions such as average monotonicity proposed by Frandsen et al. (2019) to interpret effects, or estimate interacted heterogeneous effect models as suggested by Mueller-Smith (2015). Previous evidence documenting that temperature, sports losses, or judge corps gender composition impacts judges' decisions (Eren and Mocan, 2018, 2020; Heyes and Saberian, 2019) could have similar implications.

9 Conclusion

We explore the impact of peer judge retirement on remaining judges' caseloads and allowance rates in the SSA disability judicial corps. After a peer retirement, only about 50 percent of the judges have been replaced by 9 months. As such, remaining judges see an increase in both disposition and decision caseloads on the order of 2.9-3.7 percent. During the 9 months after the retirement, the remaining judges make fewer positive decisions, resulting in a one percent drop in the judge's monthly allowance rate. This would suggest that some disability insurance claimants that would have been awarded disability insurance under different circumstances, were not because of the change in the work environment.

Judges who are less experienced exhibit the largest increase in caseloads and the largest reduction in allowance rates. There is also suggestive evidence that female judges experience a relatively larger increase in caseloads and reduction in allowance rates. These results largely seem 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. Importantly, these are impacts that 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.

Dahl, Gordon, Andreas Kostol, and Magne Mogstad, "Family Welfare Cultures," *Quarterly Journal of Economics*, 2014, *129* (4), 1711–1752.

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.

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.

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, "Annual Statistical Report on the Social Security Disability Insurance Program, 2018," 2019. SSA Publication No. 13-11826.

Ulmer, Jeffery T. and Brian Johnson, "Sentencing in Context: A Multilevel Analysis," *Criminology*, 2004, *42* (1), 137–178.

Walker, Doug, "Answer the Call to Public Service, Become an Administrative Law Judge," 2016. https://blog.ssa.gov/answer-the-call-to-public-service-become-an-administrative-law-judge/.

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

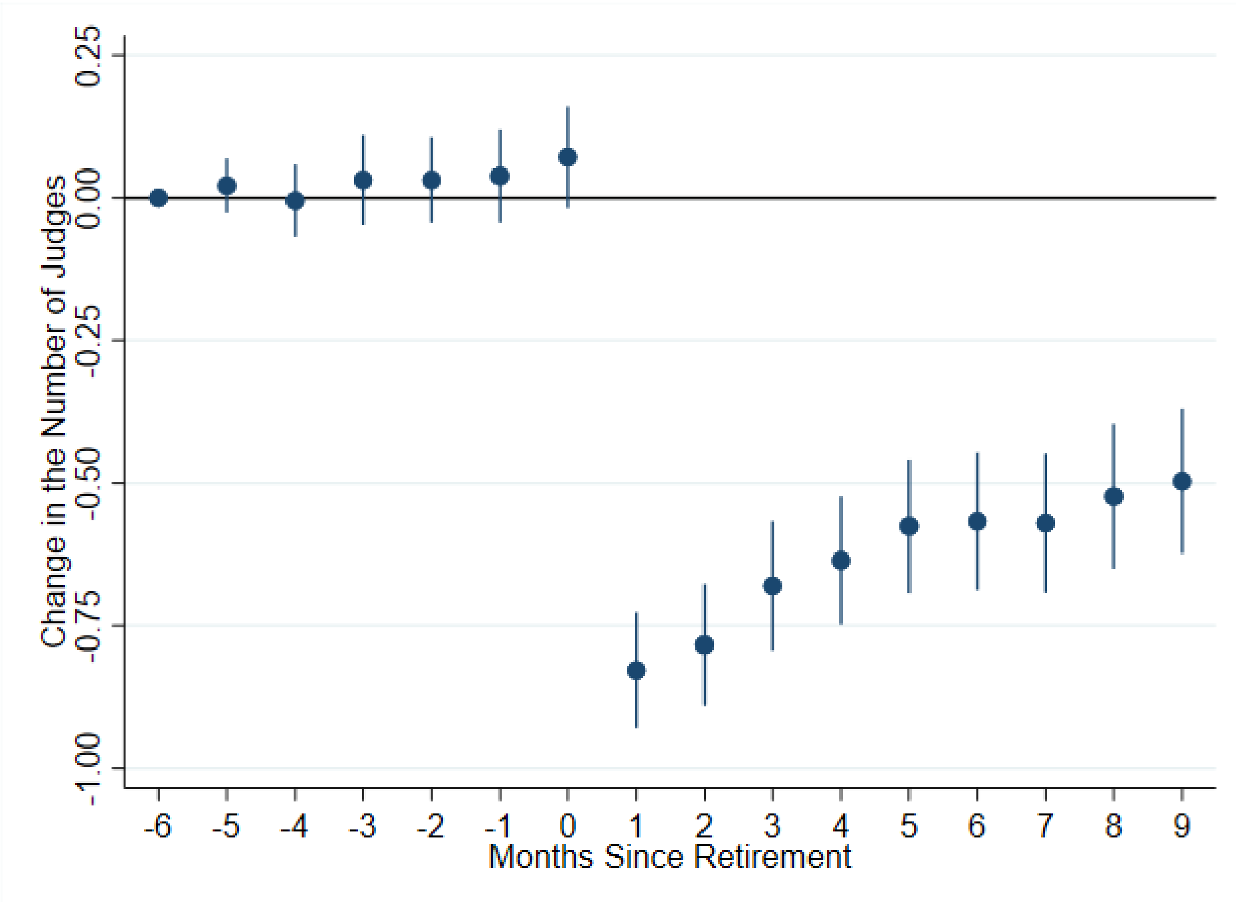

Figure 1: Impact of Judge Retirement on Office-level Number of Judges

Notes: Observation at the Hearing Office by month level. Each point is an event study coefficient obtained from estimating equation (4), 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.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

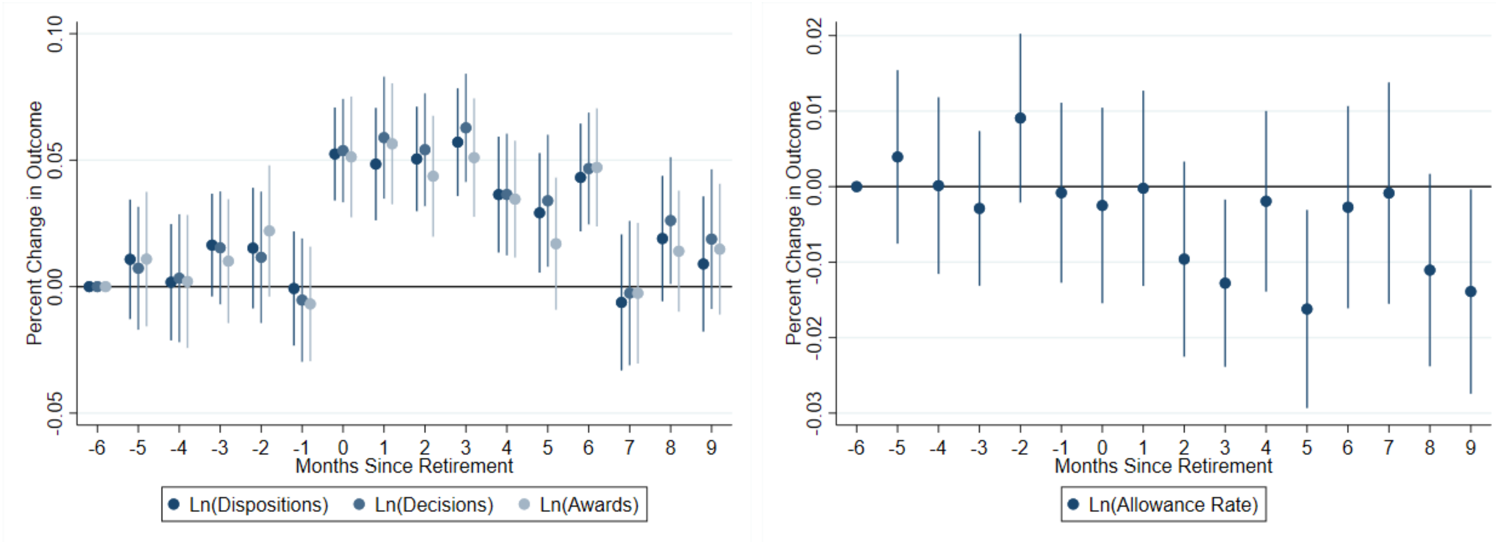

Figure 2: Impact of Judge Retirement on Judge-level Caseload and Allowance Rates

Notes: Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (2), 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.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

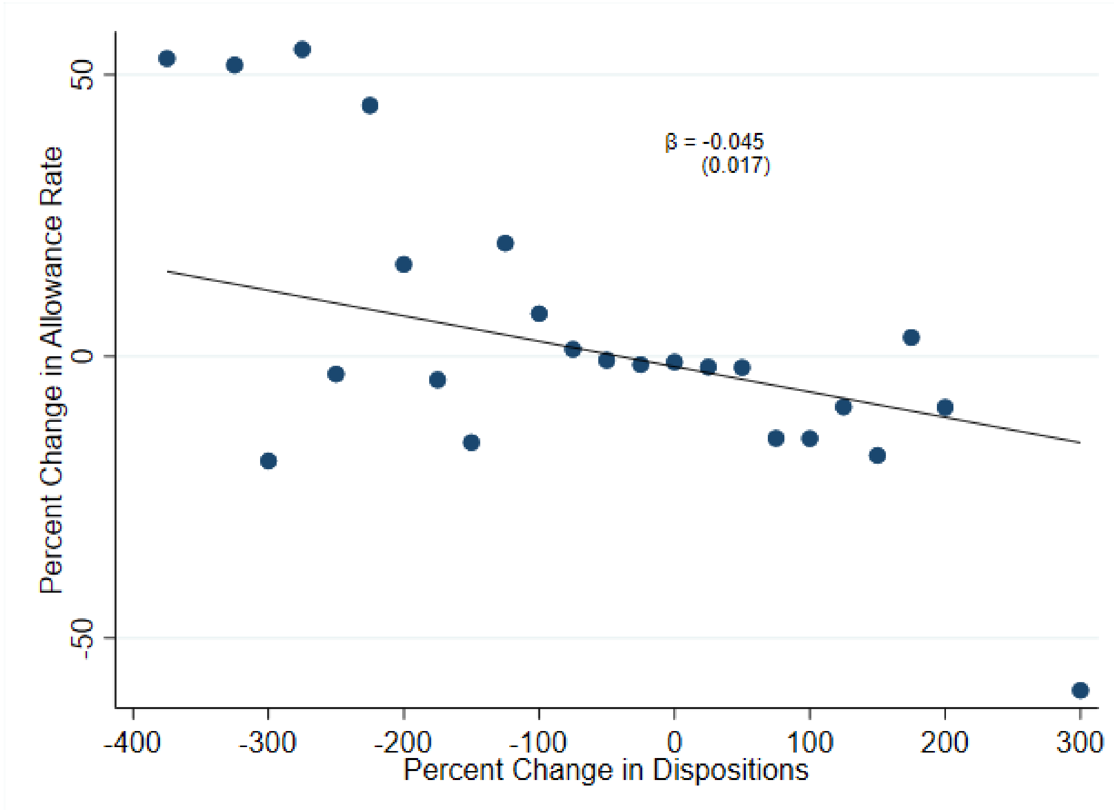

Figure 3: Potential Mechanisms: Relationship between Caseload and Allowance Rates Changes After a Peer Retirement

Notes: Each point represents the average percent change in allowance rates from the pre- to post-retirement period by the percent change in pre- to post-retirement dispositions. The average is constructed over 25 percent bins along the x-axis.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

Table 1: Judge Monthly Summary Statistics

		Event 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
Prob. White Surname	77.47	77.76	77.76	77.76	78.75
Prob. Black Surname	13.59	13.58	13.58	13.58	12.06
Prob. Hispanic Surname	6.50	6.28	6.28	6.28	6.82
Race Surname Missing	0.11	0.11	0.11	0.11	0.14
Observation Year	2014.75	2014.80	2014.39	2015.05	2014.34
Monthly Observations	158,065	90,144	33,804	56,340	4,488

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.

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

	Natural Log of							
	Dispositions	Decisions	Awards	Allowance Rate	Dispositions	Decisions	Awards	Allowance Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Month of Colleague Retirement	0.045	0.049	0.045	-0.004	1.799	1.501	0.782	-0.003
	(0.008)	(0.008)	(0.010)	(0.005)	(0.311)	(0.272)	(0.181)	(0.002)
After Colleague Retirement	0.026	0.033	0.025	-0.009	0.888	0.882	0.305	-0.005
	(0.005)	(0.005)	(0.005)	(0.003)	(0.149)	(0.133)	(0.078)	(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.

Online Appendix A: Additional Tables and Figures

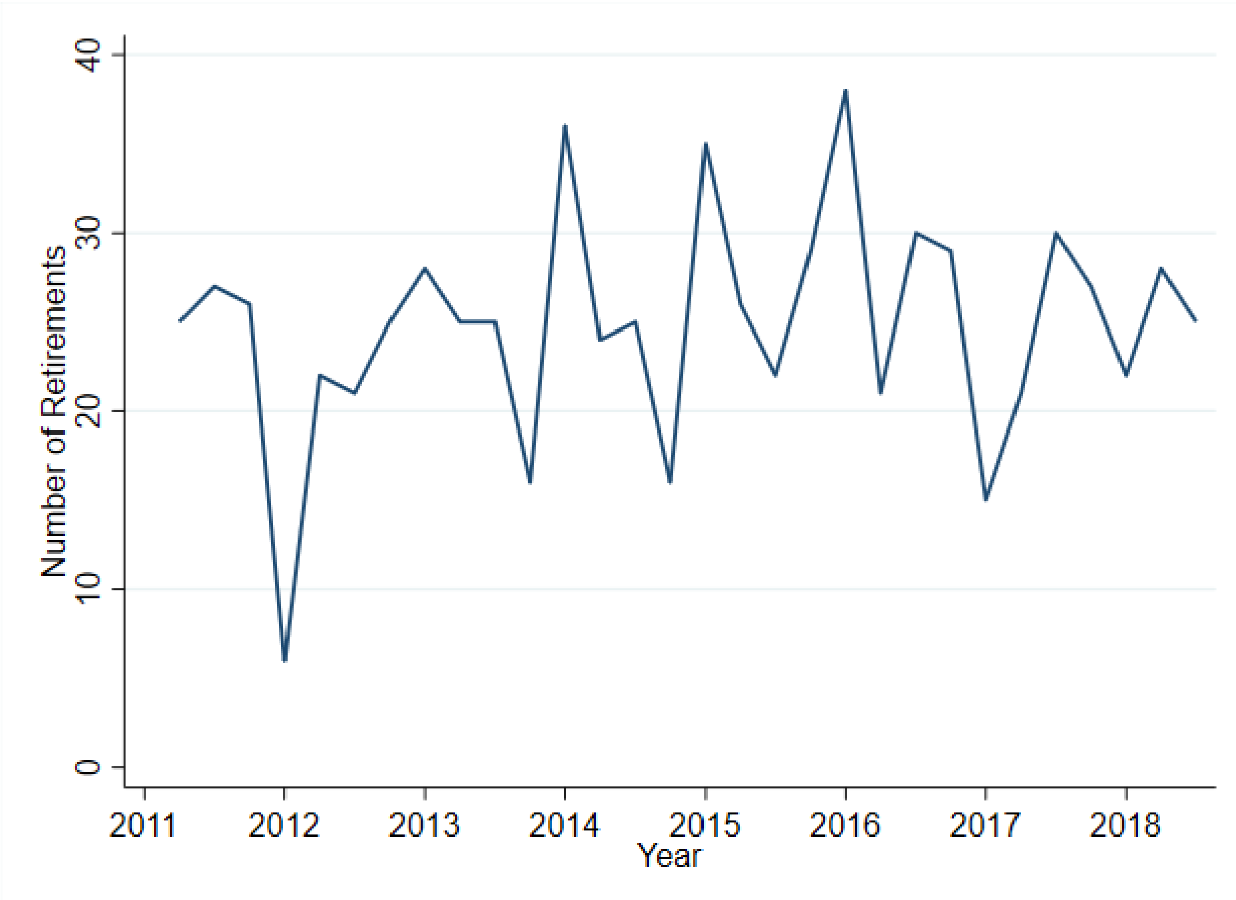

Figure A1: Number of Retirement Events

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

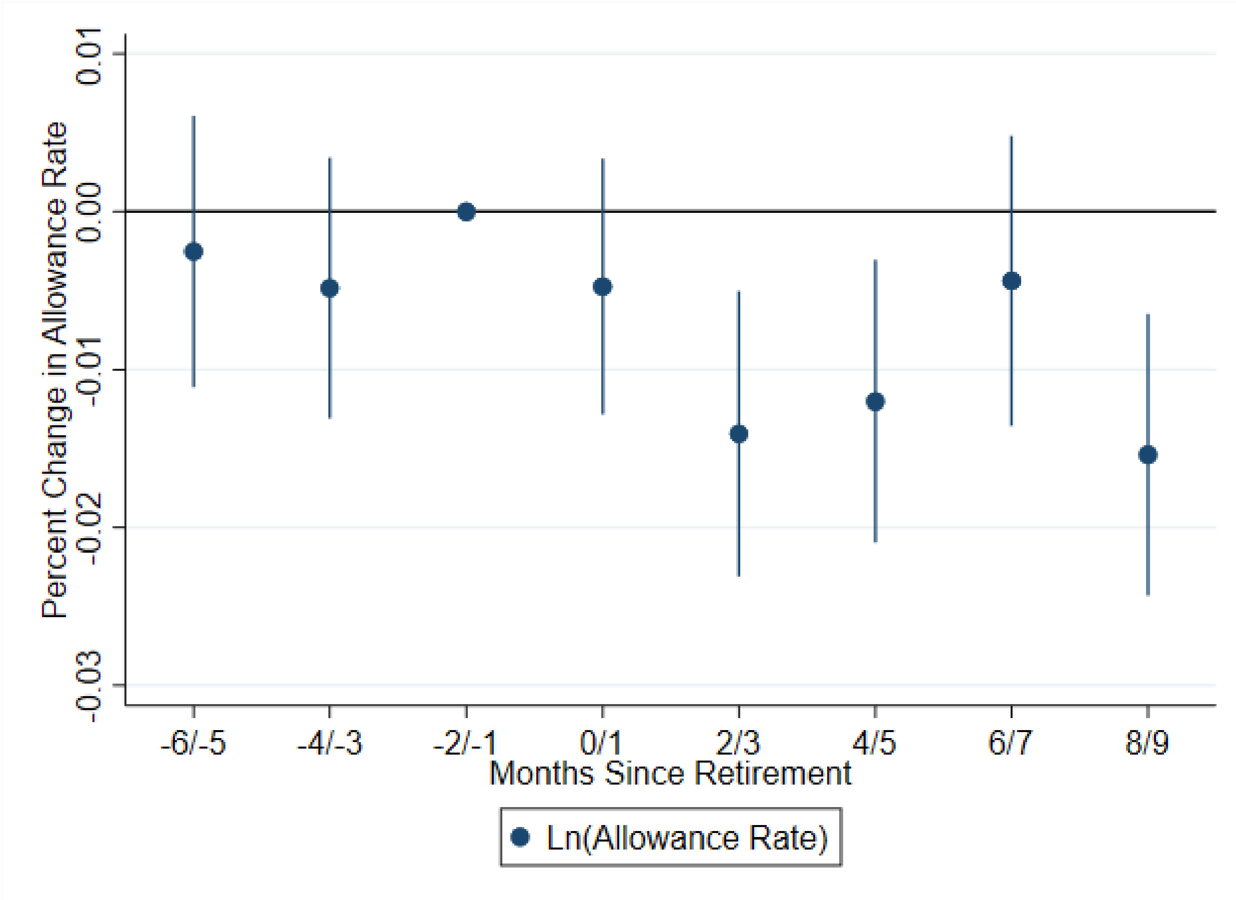

Figure A2: Impact of Judge Retirement on Judge-level Allowance Rate, 2 Week Bins

Notes: Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (2), 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.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

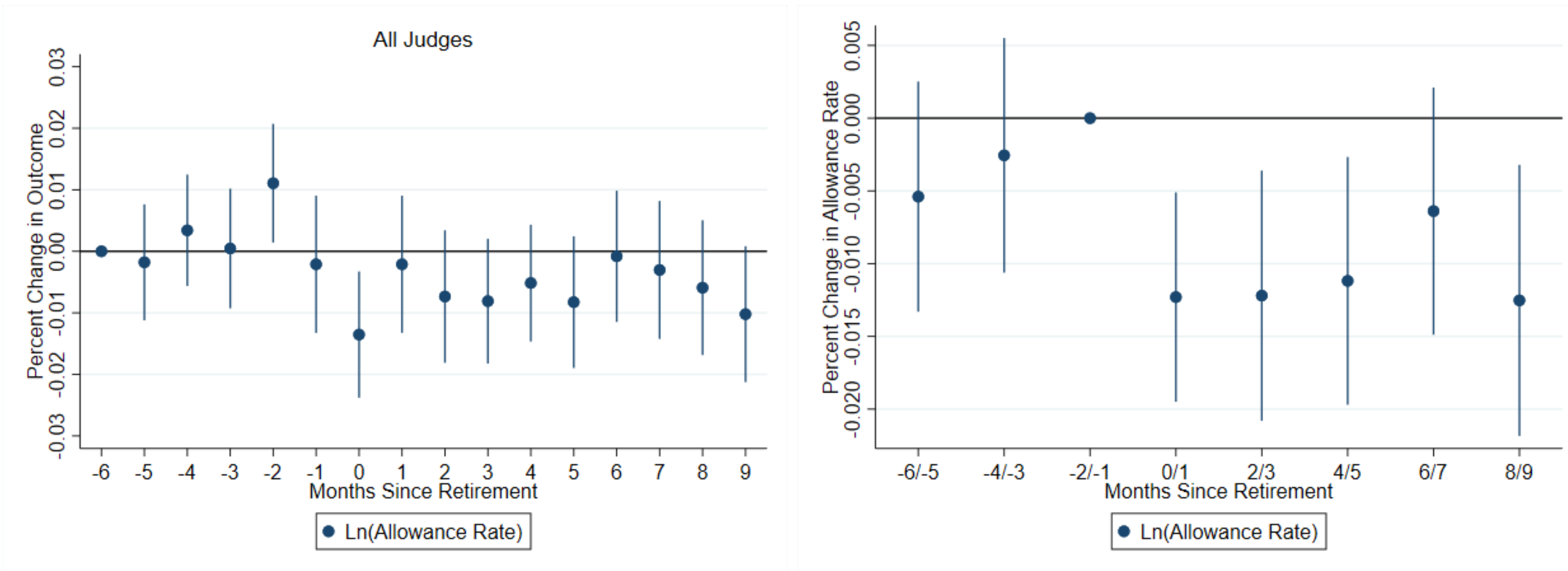

Figure A3: Impact of Judge Retirement on Office-level Allowance Rate

Notes: Observation at the Hearing Office by month level. Each point is an event study coefficient obtained from estimating equation (4), where the outcome is the natural log of the monthly office-level allowance rate. Decisions for all judges (including the retiring judge) are included in the measure. 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.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

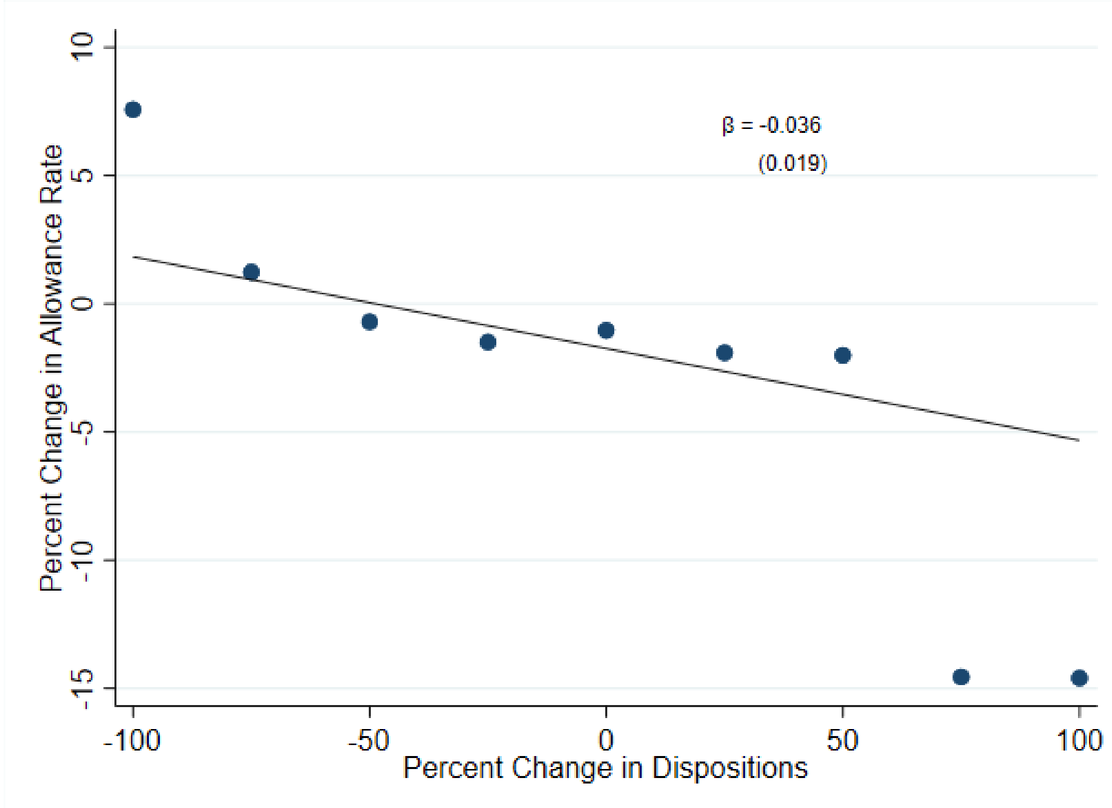

Figure A4: Potential Mechanisms: Relationship between Caseload and Allowance Rates Changes After a Peer Retirement, Removing Outliers

Notes: Each point represents the average percent change in allowance rates from the pre- to post-retirement period by the percent change in pre- to post-retirement dispositions (left) or by the retiring judge's historic allowance rate (right). The average is constructed over 25 percent bins along the x-axis.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

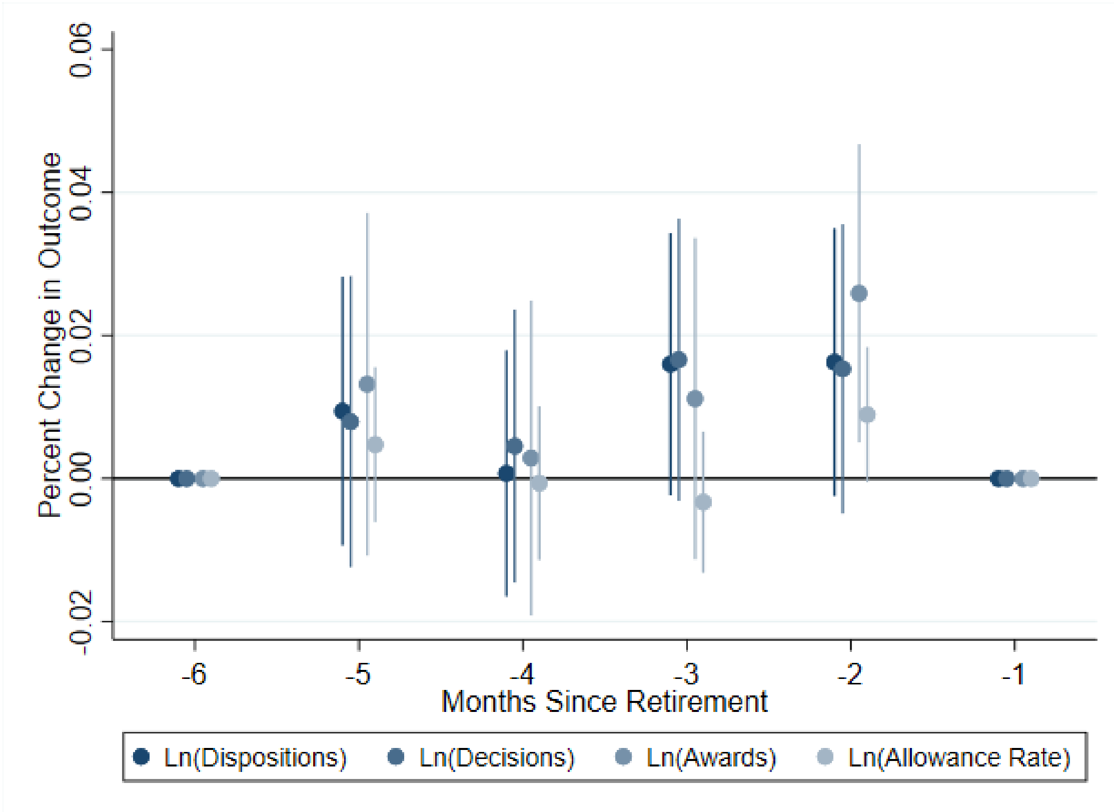

Figure A5: Testing Flatness of Pre-Trends

Notes: This test builds on (Borusyak and Jaravel, 2020) to explore the pre-trend. Observation at the judge by month level. Each point is an event study coefficient obtained from estimating an equation similar to equation (2), but only the pre-period indicators are included and the indicators for both six and one months prior are excluded. 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. 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.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

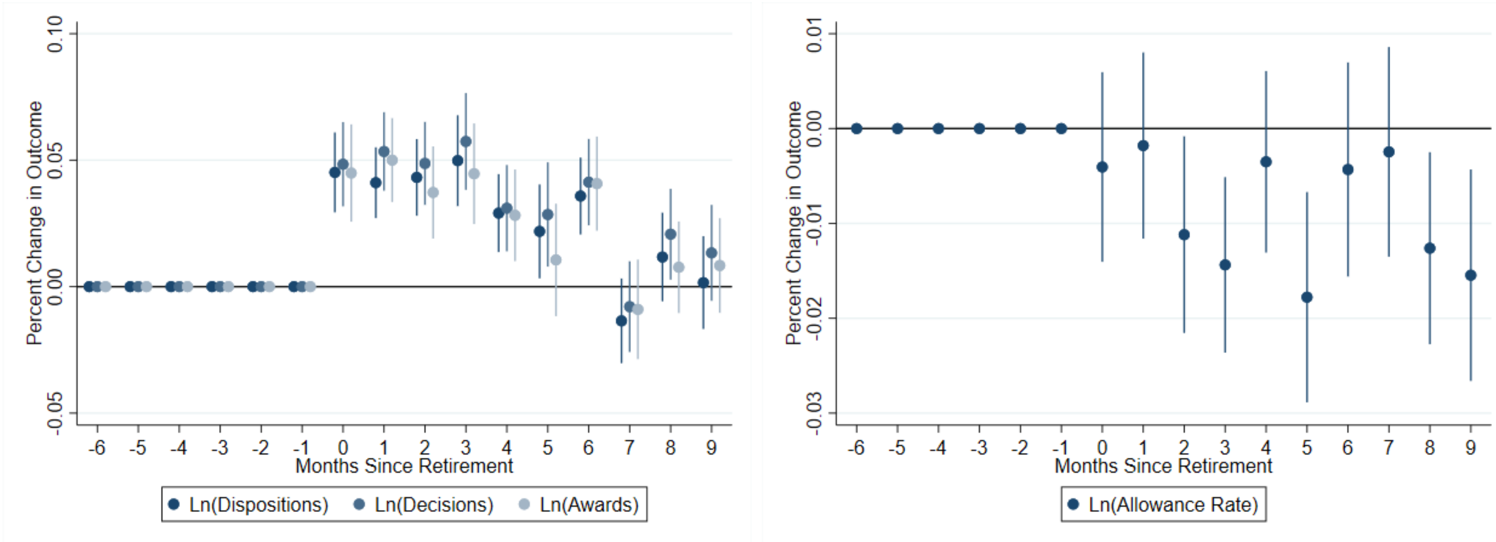

Figure A6: Impact of Judge Retirement on Judge-level Caseload and Allowance Rates with Fixed Pre-Trend

Notes: This test builds on (Borusyak and Jaravel, 2020) to explore robustness of effects to the pre-trend. Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (2), 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.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

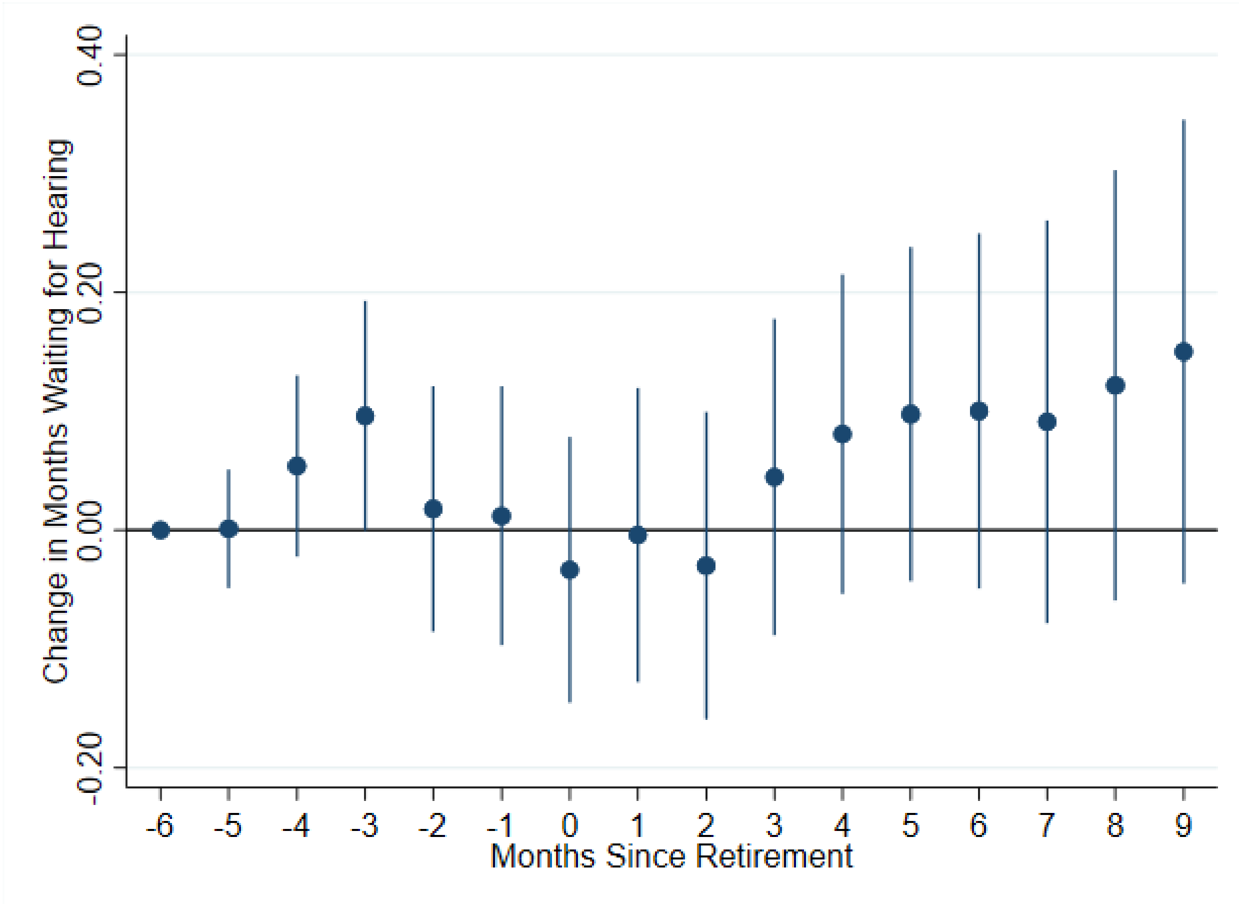

Figure A7: Impact of Judge Retirement on Office-level Wait Time Before Hearing is Held

Notes: Observation at the Hearing Office by month level. Each point is an event study coefficient obtained from estimating equation (2), where the outcome is the Hearing Office-level average number of months cases that were heard this month had to wait after appeal was filed. 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.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

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

	Natural Log of							
	Dispositions (1)	Decisions (2)	Awards (3)	Allowance Rate (4)	Dispositions (5)	Decisions (6)	Awards (7)	Allowance Rate (8)
Panel A.	Include National Hearing Centers							
Month of Colleague Retirement	0.044	0.046	0.044	-0.003	1.696	1.415	0.742	-0.003
	(0.008)	(0.008)	(0.010)	(0.005)	(0.312)	(0.275)	(0.177)	(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 B.	Include Labor Market, Industry, and Demographic Controls							
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
Panel C.	Include Everyone with Any Pre- and Post- Observations							
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 Event Windows							
Month of Colleague Retirement	0.050	0.055	0.058	0.001	2.091	1.768	1.039	-0.001
	(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.	Hearing Office and Year-by-Month Fixed 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)
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
Panel F.	Retirement Event and Month Fixed Effects							
Month of Colleague Retirement	0.035	0.035	0.032	-0.004	1.423	1.109	0.543	-0.004
	(0.008)	(0.008)	(0.009)	(0.005)	(0.306)	(0.267)	(0.175)	(0.002)
After Colleague Retirement	0.008	0.009	-0.001	-0.012	0.207	0.124	-0.206	-0.007
	(0.006)	(0.006)	(0.005)	(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 includes hearing office and month-by-year fixed effects. Panel F includes month and retirement event-level fixed effects.

Table A2: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rate, Inverse Hyperbolic Sine rather than Natural Log

	Inverse Hyperbolic Sine of							
	Dispositions	Decisions	Awards	Allowance Rate	Dispositions	Decisions	Awards	Allowance Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Month of Colleague Retirement	0.042	0.043	0.037	-0.003	1.799	1.501	0.782	-0.003
	(0.010)	(0.010)	(0.011)	(0.002)	(0.311)	(0.272)	(0.181)	(0.002)
After Colleague Retirement	0.011	0.018	0.010	-0.005	0.888	0.882	0.305	-0.005
	(0.007)	(0.007)	(0.007)	(0.001)	(0.149)	(0.133)	(0.078)	(0.002)
Dependent Mean (in Levels)	40.17	32.43	18.32	0.56	40.17	32.43	18.32	0.56
Observations	90,144	90,144	90,144	89,205	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.

Table A3: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rate, Removing MSPB Dismissals

	Natural Log of							
	Dispositions	Decisions	Awards	Allowance Rate	Dispositions	Decisions	Awards	Allowance Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Retiree Not Disciplined or Reviewed by MSPB							
Month of Colleague Retirement	0.046	0.049	0.045	-0.005	1.789	1.485	0.767	-0.004
	(0.008)	(0.008)	(0.010)	(0.005)	(0.304)	(0.265)	(0.177)	(0.002)
After Colleague Retirement	0.026	0.033	0.025	-0.009	0.879	0.873	0.303	-0.005
	(0.005)	(0.005)	(0.005)	(0.003)	(0.147)	(0.132)	(0.079)	(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
	4 Retiree Judges Disciplined or Reviewed by MSPB							
Month of Colleague Retirement	0.087	0.094	0.192	0.098	5.145	5.884	4.628	0.039
	(0.138)	(0.129)	(0.104)	(0.049)	(2.592)	(1.978)	(0.499)	(0.030)
After Colleague Retirement	0.112	0.085	0.011	-0.074	1.305	1.699	-1.259	-0.045
	(0.114)	(0.096)	(0.103)	(0.019)	(2.551)	(2.078)	(1.721)	(0.013)
Dependent Mean (in Levels)	44.83	36.88	20.50	0.54	44.76	36.76	20.43	0.54
Observations	607	606	606	606	608	608	608	606

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 top panel includes retirement events where there is no evidence the judge was disciplined or removed by the US Merit System Protection Board (MSPB). The bottom panel includes the four retirement events where there is record of a MSPB case for the retiring ALJ. There is one other judge who had a case heard through the MSPB, but his retirement event is not in our sample window. Judge, Hearing Office, year, and month fixed effects are included. Standard errors corrected for clustering at the Hearing Office level.

Table A4: Heterogeneous Impact of Colleague Retirement on Judge-level Caseload and Allowance Rate by Time Period

	Natural Log of							
	Dispositions	Decisions	Awards	Allowance Rate	Dispositions	Decisions	Awards	Allowance Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(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
	(0.007)	(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.

Table A5: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rate, Retirement-Event Stacked Panel

	Natural Log of							
	Dispositions	Decisions	Awards	Allowance Rate	Dispositions	Decisions	Awards	Allowance Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Stacked Baseline Specification							
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
	Stacked Baseline Specification with Year and Month 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

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.

Table A6: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rate by Judge Characteristics

	Natural Log of							
	Dispositions (1)	Decisions (2)	Awards (3)	Allowance Rate (4)	Dispositions (5)	Decisions (6)	Awards (7)	Allowance Rate (8)
Month of Colleague Retirement	0.038	0.042	0.045	-0.001	1.741	1.423	0.744	-0.003
	(0.009)	(0.010)	(0.010)	(0.005)	(0.363)	(0.314)	(0.209)	(0.002)
After Colleague Retirement	0.021	0.027	0.022	-0.006	0.651	0.665	0.213	-0.004
	(0.005)	(0.005)	(0.006)	(0.004)	(0.158)	(0.142)	(0.090)	(0.002)
Month of Colleague Retirement	0.021	0.019	0.001	-0.010	0.192	0.232	0.114	0.000
*Female	(0.011)	(0.012)	(0.013)	(0.009)	(0.363)	(0.316)	(0.206)	(0.004)
After Colleague Retirement	0.013	0.018	0.011	-0.009	0.712	0.651	0.277	-0.004
*Female	(0.009)	(0.009)	(0.009)	(0.005)	(0.237)	(0.199)	(0.134)	(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	0.045	0.048	0.043	-0.005	1.786	1.492	0.758	-0.004
	(0.008)	(0.009)	(0.010)	(0.005)	(0.319)	(0.279)	(0.183)	(0.002)
After Colleague Retirement	0.025	0.032	0.025	-0.009	0.860	0.858	0.299	-0.005
	(0.005)	(0.005)	(0.005)	(0.003)	(0.147)	(0.132)	(0.080)	(0.002)
Month of Colleague Retirement	0.005	0.014	0.029	0.015	0.179	0.114	0.314	0.007
*Non-White Surname	(0.020)	(0.022)	(0.026)	(0.013)	(0.633)	(0.586)	(0.382)	(0.006)
After Colleague Retirement	0.011	0.011	0.010	-0.002	0.369	0.310	0.075	-0.003
*Non-White Surname	(0.013)	(0.013)	(0.014)	(0.009)	(0.342)	(0.299)	(0.231)	(0.004)
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	0.036	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)
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	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. 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.

Table A7: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rate by Retiree's Characteristics

	Natural Log of							
	Dispositions	Decisions	Awards	Allowance Rate	Dispositions	Decisions	Awards	Allowance Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Month of Colleague Retirement	0.055	0.058	0.059	0.000	2.118	1.799	1.007	-0.002
	(0.010)	(0.011)	(0.012)	(0.006)	(0.381)	(0.335)	(0.219)	(0.003)
After Colleague Retirement	0.027	0.034	0.026	-0.009	0.970	0.949	0.355	-0.005
	(0.005)	(0.006)	(0.006)	(0.003)	(0.164)	(0.151)	(0.094)	(0.002)
Month of Colleague Retirement	-0.039	-0.037	-0.057	-0.018	-1.270	-1.189	-0.894	-0.005
*Retiree Female	(0.019)	(0.021)	(0.021)	(0.011)	(0.657)	(0.571)	(0.354)	(0.005)
After Colleague Retirement	-0.004	-0.003	-0.002	-0.001	-0.320	-0.265	-0.198	0.000
*Retiree Female	(0.010)	(0.012)	(0.013)	(0.006)	(0.321)	(0.291)	(0.186)	(0.003)
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	0.048	0.052	0.046	-0.006	1.870	1.566	0.778	-0.005
	(0.008)	(0.009)	(0.010)	(0.005)	(0.313)	(0.278)	(0.188)	(0.002)
After Colleague Retirement	0.027	0.035	0.026	-0.010	0.920	0.912	0.316	-0.006
	(0.005)	(0.005)	(0.005)	(0.003)	(0.153)	(0.140)	(0.085)	(0.001)
Month of Colleague Retirement	-0.027	-0.033	-0.008	0.025	-0.781	-0.720	0.044	0.014
*Retiree Non-White Surname	(0.028)	(0.031)	(0.032)	(0.015)	(1.129)	(1.049)	(0.690)	(0.007)
After Colleague Retirement	-0.020	-0.024	-0.012	0.007	-0.389	-0.357	-0.125	0.003
*Retiree Non-White Surname	(0.014)	(0.015)	(0.015)	(0.010)	(0.513)	(0.433)	(0.312)	(0.005)
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	0.042	0.044	0.043	-0.002	1.666	1.365	0.708	-0.003
	(0.008)	(0.009)	(0.010)	(0.005)	(0.293)	(0.260)	(0.166)	(0.002)
After Colleague Retirement	0.027	0.035	0.029	-0.008	0.988	0.979	0.378	-0.005
	(0.005)	(0.005)	(0.005)	(0.003)	(0.159)	(0.142)	(0.084)	(0.002)
Month of Colleague Retirement	0.032	0.047	0.020	-0.025	1.327	1.333	0.592	-0.011
*Retiree < 1 Year as Judge	(0.029)	(0.030)	(0.029)	(0.013)	(1.092)	(0.979)	(0.651)	(0.007)
After Colleague Retirement	-0.002	-0.013	-0.037	-0.024	-0.820	-0.853	-1.012	-0.015
*Retiree < 1 Year as Judge	(0.014)	(0.015)	(0.017)	(0.009)	(0.492)	(0.469)	(0.340)	(0.005)
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	0.046	0.049	0.045	-0.004	1.823	1.519	0.793	-0.003
	(0.008)	(0.008)	(0.010)	(0.005)	(0.310)	(0.274)	(0.182)	(0.002)
After Colleague Retirement	0.025	0.033	0.025	-0.009	0.881	0.878	0.304	-0.005
	(0.005)	(0.005)	(0.005)	(0.003)	(0.148)	(0.134)	(0.078)	(0.002)
Month of Colleague Retirement	-0.014	-0.025	0.013	0.034	-2.505	-1.866	-1.029	0.008
*Retiree Pre- Award Rate Minus Office Award Rate	(0.051)	(0.058)	(0.060)	(0.034)	(1.999)	(1.835)	(1.083)	(0.016)
After Colleague Retirement	0.052	0.052	0.042	-0.011	1.399	0.986	0.437	-0.006
*Retiree Pre- Award Rate Minus Office Award Rate	(0.034)	(0.039)	(0.038)	(0.021)	(1.153)	(1.072)	(0.592)	(0.010)
Dependent Mean (in Levels)	40.47	32.77	18.61	0.56	40.16	32.43	18.32	0.56
Observations	89,375	89,125	88,673	88,673	90,064	90,064	90,064	89,125

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.

Table A8: Impact of Colleague Retirement on Caseload and Allowance Rates by Gender Parity

	Natural 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
	(0.010)	(0.011)	(0.012)	(0.006)	(0.426)	(0.375)	(0.245)	(0.003)
After Colleague Retirement	0.027	0.034	0.028	-0.006	0.889	0.873	0.328	-0.004
	(0.006)	(0.006)	(0.007)	(0.004)	(0.181)	(0.164)	(0.106)	(0.002)
Month of Collegue Retirement	0.015	0.011	-0.008	-0.009	0.010	-0.024	-0.048	-0.001
*Retiree Male, Female Judge	(0.013)	(0.014)	(0.016)	(0.010)	(0.437)	(0.375)	(0.246)	(0.005)
After Colleague Retirement	0.012	0.015	0.006	-0.011	0.592	0.522	0.253	-0.004
*Retiree Male, Female Judge	(0.010)	(0.009)	(0.010)	(0.005)	(0.266)	(0.227)	(0.150)	(0.002)
Month of Colleague Retirement	-0.002	0.001	-0.024	-0.018	-0.585	-0.366	-0.241	0.001
*Retiree Female, Female Judge	(0.024)	(0.026)	(0.026)	(0.015)	(0.762)	(0.644)	(0.414)	(0.007)
After Colleague Retirement	-0.006	0.002	0.002	-0.002	-0.050	0.052	-0.048	-0.001
*Retiree Female, Female Judge	(0.017)	(0.019)	(0.018)	(0.008)	(0.470)	(0.439)	(0.247)	(0.004)
Month of Colleague Retirement	-0.046	-0.048	-0.069	-0.014	-1.637	-1.642	-1.115	-0.004
*Retiree Female, Male Judge	(0.019)	(0.021)	(0.023)	(0.012)	(0.703)	(0.624)	(0.403)	(0.006)
After Colleague Retirement	-0.008	-0.010	-0.013	-0.003	-0.647	-0.599	-0.318	0.000
*Retiree Female, Male Judge	(0.014)	(0.015)	(0.015)	(0.006)	(0.409)	(0.360)	(0.226)	(0.003)
Dependent Mean (in Levels)	40.5	32.8	18.5	0.6	40.2	32.4	18.2	0.6
Observations	94,905	94,628	94,141	94,141	95,656	95,656	95,656	94,628

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.

Table A9: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rate, Detailed Race Heterogeneity

	Natural 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	0.051	0.046	-0.008	1.733	1.516	0.648	-0.005
	(0.012)	(0.012)	(0.013)	(0.007)	(0.405)	(0.353)	(0.239)	(0.003)
After Colleague Retirement	0.022	0.031	0.022	-0.010	0.719	0.731	0.187	-0.005
	(0.008)	(0.008)	(0.008)	(0.004)	(0.215)	(0.191)	(0.125)	(0.002)
Month of Colleague Retirement	0.001	0.000	0.000	0.000	0.016	0.011	0.012	0.000
*Retiree Black Surname	(0.000)	(0.001)	(0.001)	(0.000)	(0.015)	(0.013)	(0.009)	(0.000)
After Colleague Retirement	0.000	0.000	0.000	0.000	0.012	0.011	0.008	-0.000
*Retiree Black Surname	(0.000)	(0.000)	(0.000)	(0.000)	(0.009)	(0.008)	(0.006)	(0.000)
Month of Colleague Retirement	-0.000	-0.000	-0.000	0.000	-0.006	-0.011	-0.002	0.000
*Retiree Hispanic Surname	(0.000)	(0.000)	(0.001)	(0.000)	(0.013)	(0.011)	(0.007)	(0.000)
After Colleague Retirement	0.000	-0.000	0.000	0.000	0.002	0.001	0.003	0.000
*Retiree Hispanic Surname	(0.000)	(0.000)	(0.000)	(0.000)	(0.007)	(0.006)	(0.004)	(0.000)
Dependent Mean (in Levels)	40.65	32.95	18.49	0.56	40.34	32.61	18.21	0.56
Observations	68,472	68,277	67,933	67,933	68,992	68,992	68,992	68,277
Month of Colleague Retirement	0.050	0.054	0.050	-0.000	1.974	1.552	0.904	-0.001
	(0.010)	(0.011)	(0.015)	(0.008)	(0.395)	(0.344)	(0.235)	(0.004)
After Colleague Retirement	0.027	0.034	0.027	-0.009	0.918	0.888	0.291	-0.005
	(0.008)	(0.008)	(0.008)	(0.005)	(0.236)	(0.202)	(0.116)	(0.002)
Month of Colleague Retirement	0.004	0.007	0.007	-0.006	0.190	0.267	0.030	-0.005
*Prob. Retiree Surname White Q2	(0.013)	(0.014)	(0.017)	(0.011)	(0.538)	(0.477)	(0.277)	(0.005)
After Colleague Retirement	-0.000	0.000	0.006	0.006	0.175	0.170	0.173	0.001
*Prob. Retiree Surname White Q2	(0.010)	(0.010)	(0.012)	(0.007)	(0.318)	(0.264)	(0.172)	(0.004)
Month of Colleague Retirement	0.002	0.004	0.009	0.001	-0.257	0.070	-0.154	-0.001
*Prob. Retiree Surname White Q3	(0.017)	(0.018)	(0.022)	(0.011)	(0.448)	(0.388)	(0.269)	(0.005)
After Colleague Retirement	0.008	0.006	0.002	-0.003	0.230	0.169	0.085	-0.001
*Prob. Retiree Surname White Q3	(0.011)	(0.010)	(0.012)	(0.006)	(0.307)	(0.254)	(0.174)	(0.003)
Month of Colleague Retirement	-0.026	-0.036	-0.036	-0.011	-0.719	-0.578	-0.421	-0.006
*Prob. Retiree Surname White Q4	(0.016)	(0.017)	(0.020)	(0.010)	(0.559)	(0.480)	(0.329)	(0.005)
After Colleague Retirement	-0.011	-0.012	-0.014	-0.005	-0.544	-0.373	-0.204	-0.001
*Prob. Retiree Surname White Q4	(0.010)	(0.010)	(0.011)	(0.006)	(0.303)	(0.265)	(0.177)	(0.003)
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. In the bottom panel, retiree judges are split into quartiles by probability that surname is associated with being white. Standard errors corrected for clustering at the Hearing Office level.

Table A10: Impact of Colleague Retirement on Judge-level Caseload and Allowance Rate by Hearing Office Characteristics

	Natural Log of							
	Dispositions (1)	Decisions (2)	Awards (3)	Allowance Rate (4)	Dispositions (5)	Decisions (6)	Awards (7)	Allowance Rate (8)
Month of Colleague Retirement	0.044	0.047	0.043	-0.004	1.765	1.465	0.749	-0.003
	(0.008)	(0.008)	(0.010)	(0.005)	(0.315)	(0.273)	(0.181)	(0.002)
After Colleague Retirement	0.021	0.027	0.018	-0.011	0.774	0.737	0.183	-0.005
	(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)
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 Colleague Retirement	0.044	0.047	0.044	-0.004	1.761	1.468	0.762	-0.003
	(0.008)	(0.008)	(0.010)	(0.005)	(0.308)	(0.271)	(0.180)	(0.002)
After Colleague Retirement	0.024	0.031	0.024	-0.009	0.833	0.834	0.276	-0.005
	(0.005)	(0.005)	(0.005)	(0.003)	(0.155)	(0.140)	(0.082)	(0.002)
Month of Colleague Retirement	-0.000	-0.000	-0.000	-0.000	-0.004	-0.003	-0.001	-0.000
*De-Meaned Average Monthly Dispositions	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)	(0.002)	(0.002)	(0.000)
After Colleague Retirement	-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)	(0.001)	(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 Retirement	0.045	0.048	0.045	-0.004	1.789	1.494	0.780	-0.003
	(0.008)	(0.008)	(0.010)	(0.005)	(0.306)	(0.270)	(0.179)	(0.002)
After Colleague Retirement	0.025	0.032	0.025	-0.009	0.865	0.867	0.300	-0.005
	(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)
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. 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.

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

In Table A1, we explore the impacts on dispositions, decisions, and allowance rates are similar if we include retirement events from the National Hearing Offices, control for local labor market conditions and demographics in the hearing office catchment area, relax our balanced panel requirement to include anyone with both a pre- and post- observation, exclude events where the event window overlaps, include year-by-month fixed effects rather than year and month fixed effects, or include retirement event and month fixed effects. The impact on allowance rates range from negative 1.3 percent to negative 0.7 percent. If we 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 Appendix Table A2).

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. (Appendix Table B1). Judge retirements do not affect the share of office level cases held by video (Appendix Figure B1).

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 similar if we exclude known dismissal or disciplined retirement events (Appendix Table A3). Among non-dismissed retirement events there is a similar increase in dispositions and decisions and a one percent drop in allowance rates. Estimates for 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. We also examine how the effects of a peer judge retirement differ over time. Beginning in 2010 and 2011 the SSA undertook various ALJ reforms aimed at standardizing outcomes across judges. 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 (Appendix Figure B2). We estimate effects separately for retirement events from 2011-2014 and 2015-2018 and 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 (Appendix Table A4).

To account for potential, underlying secular trends in caseloads and allowance rates we estimate an alternative model, similar to Deshpande and Li (2019). 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.[16] 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 judge-level 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

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

there will be a retirement exactly one year later as follows

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

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) are also included 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 (2).[17]

As seen in Appendix Figure B3 judges in the treated hearing office experience an increase in dispositions and decisions after the retirement, although this increase is only sustained for four months. There is also a drop in the allowance rate starting around week two, that is even more pronounced than in the baseline model. Results from the parameterized difference in differences model are also similar (Appendix Table A5). 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.

Our estimates are also 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 (2), but only include pre-period months and omit both the first and last month indicators, we see

[17]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.

the pre-trends are flat, with no significant trend. Only one point (log awards in month t-2) is significantly different from zero (Appendix Figure A5). 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 (Appendix Figure A6).

Heterogeneity

We also explore heterogeneity by remaining judge characteristics, retiring judge characteristics, and limited characteristics of the hearing office. To do this, we estimate a variant of equation 3, as follows

$$\begin{aligned} Y_{jot} = {} & \beta_0 \textit{Month of Colleague Retirement}_{ot} + \beta_1 \textit{After Colleague Retirement}_{ot} \\ & + \beta_3 \textit{Month of Colleague Retirement}_{ot} * Char_{jo} + \beta_4 \textit{After Colleague Retirement}_{ot} \\ & * Char_{jo} + \beta_5 Char_{jo} + \delta_j + \phi_o + \gamma_{yr} + \psi_{mo} + \varepsilon_{jot} \end{aligned} \tag{6}$$

Where $Char_{jo}$ is a judge, retiree, or office specific characteristic. The θ_0 and θ_1 coefficients will allow us to test if impacts vary by the judge's characteristics (such as gender, race, 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 judge race or gender, the direct effect of $Char_{jo}$ will be absorbed by the fixed effects.

By Judge's Characteristics. We explore how effects vary by the judge's characteristics to test if certain groups are more responsive in Table A6. 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.[18] 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.

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. We explore how effects vary by the retiring judge's characteristics in Table A7 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. 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 Appendix Tables A8 and A9. 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

[18] 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.

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 A10 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. Effects do not significantly vary by the number of judges in the hearing office during the pre-retirement period. The impact of a peer retirement does not appear to significantly vary with the local setting.

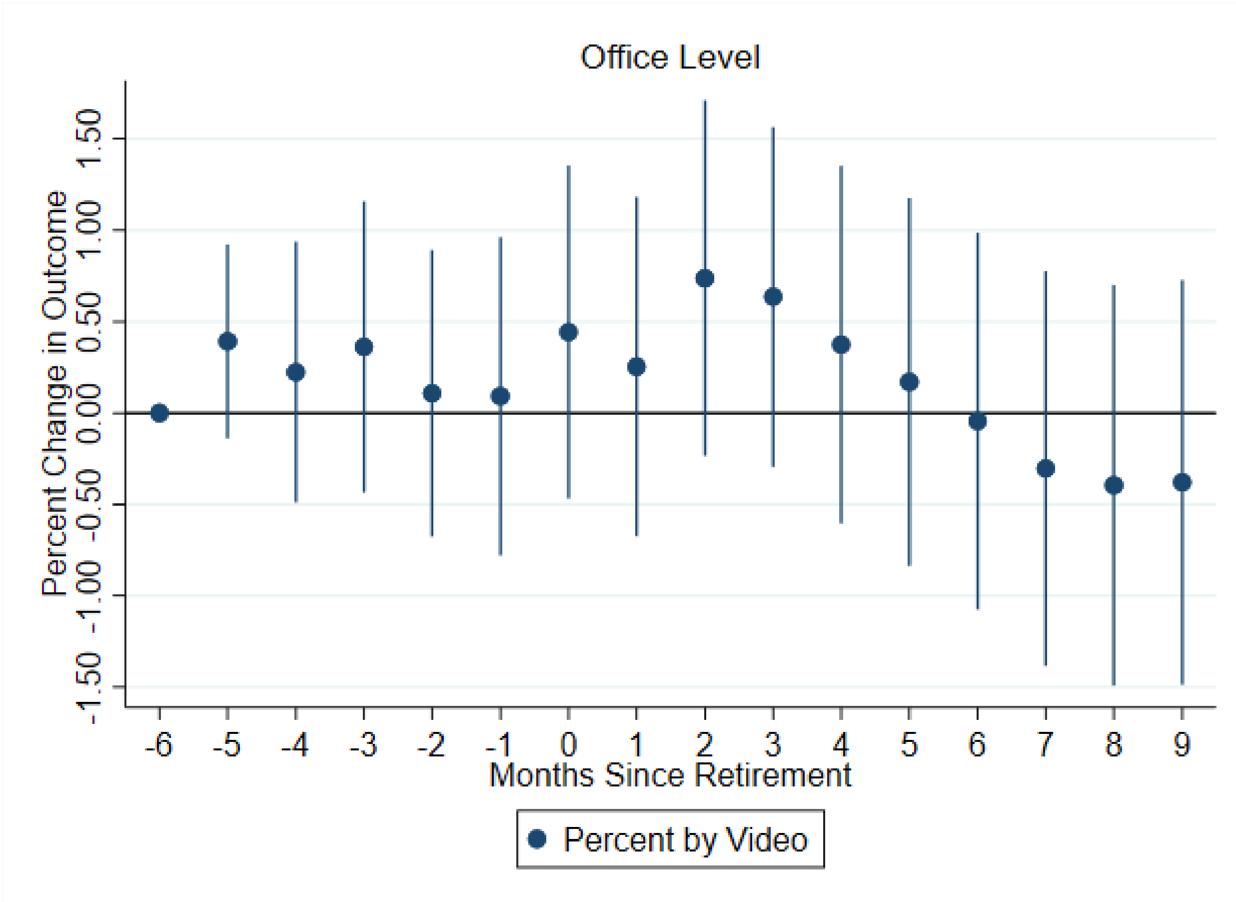

Figure B1: Impact of Judge Retirement on Office-level Hearings Held by Video

Notes: Observation at the Hearing Office by month level. Each point is an event study coefficient obtained from estimating equation (4), where the outcome is the natural log of the share of cases held over video. Decisions for all judges (including the retiring judge) are included in the measure. 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.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

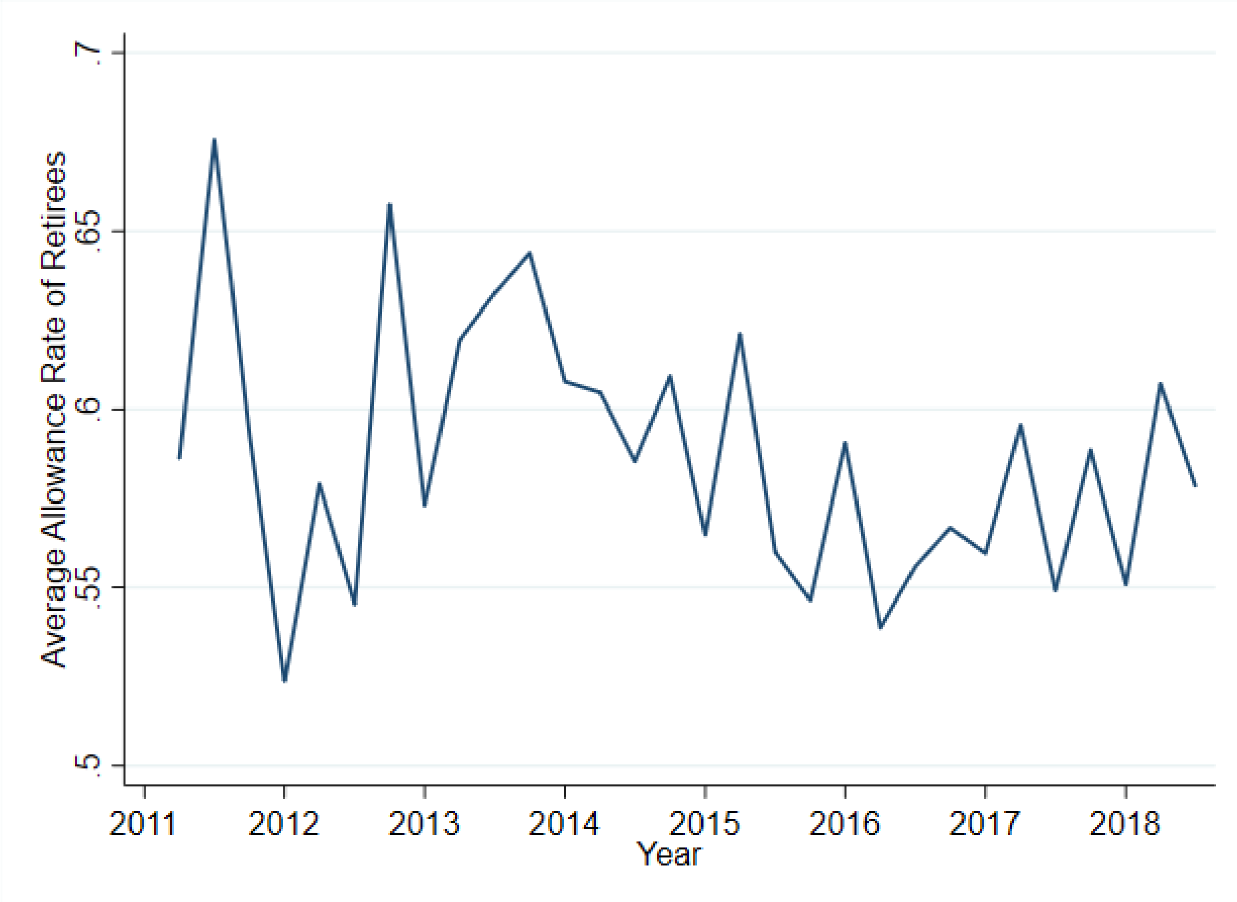

Figure B2: Historical Allowance Rate of Retiring Judges by Quarter of Retirement

Notes: Each point represents the historic allowance rate from September 2010 to retirement, averaged across all judges that retired in the given quarter.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

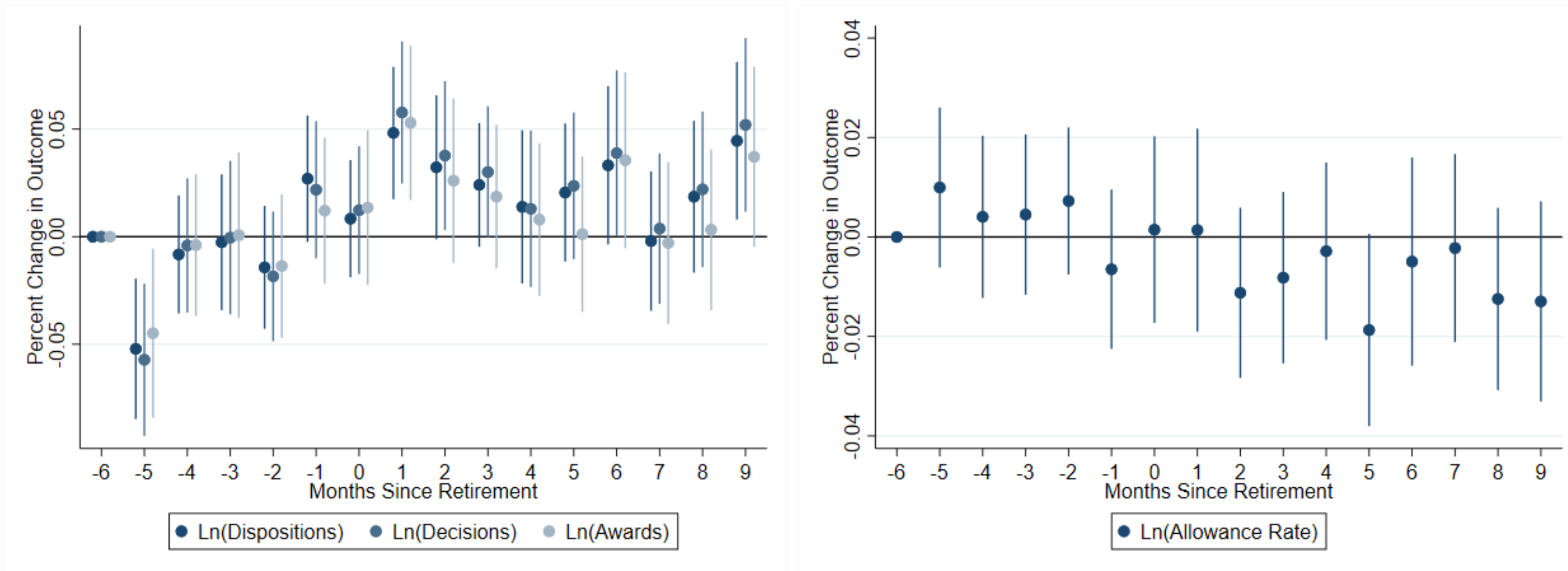

Figure B3: Impact of Judge Retirement on Judge-level Caseload, Stacked Event Panel Including Future Retirement Events as Controls

Notes: Observation at the judge by month level. Each point is an event study coefficient obtained from estimating equation (5), 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.

Source: Author's own calculations using the monthly SSA ALJ Disposition Data.

Table B1: Impact of Colleague Retirement on Other Judge-level Outcomes

	Natural Log of					
	Awards to Dispositions Rate (1)	Fully Favorable Rate (2)	Partially Favorable Rate (3)	Awards to Dispositions Rate (4)	Fully Favorable Rate (5)	Partially Favorable Rate (6)
Month of Colleague Retirement	-0.003	-0.006	-0.021	-0.002	-0.002	-0.001
	(0.006)	(0.006)	(0.009)	(0.002)	(0.002)	(0.001)
After Colleague Retirement	-0.003	-0.010	-0.014	-0.002	-0.005	-0.000
	(0.003)	(0.004)	(0.006)	(0.001)	(0.002)	(0.001)
Dependent Mean (in Levels)	0.45	0.49	0.09	0.45	0.49	0.07
Observations	88,753	88,417	69,668	89,455	89,205	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. 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.