Mandatory Retirement for Judges Improved Performance on U.S. State Supreme Courts

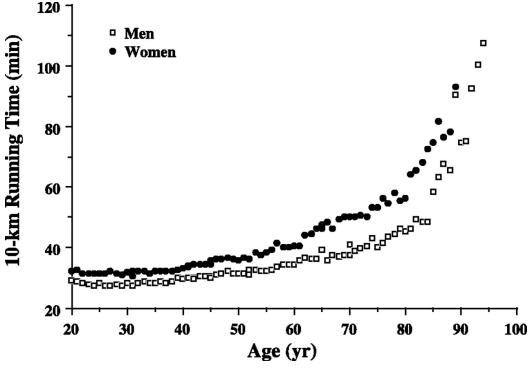
Elliott Ash, ETH Zurich
W. Bentley MacLeod, Columbia University and NBER

NBER Summer Institute, July 2020

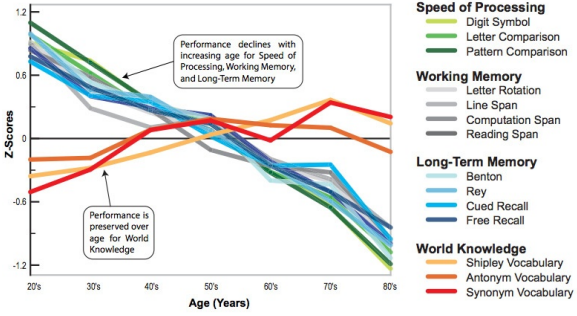
We thank Yisehak Abraham, Ankeet Ball, Josh Brown, Josh Burton, Matthew Buck, Eammonn Campbell, Zoey Chopra, Daniel Deibler, Seth Fromer, Gohar Harutyunyan, Archan Hazra, Montague Hung, Dong Hyeun, Mithun Kamath, James Kim, Michael Kurish, Jennifer Kutsunai, Steven Lau, Sharon Liao, Claudia Marangon, Sarah MacDougall, Justin McNamee, Sourabh Mishra, Brendan Moore, Arielle Napoli, Karen Orchansky, Bryn Paslawski, Olga Peshko, Matteo Pinna, Quinton Robbins, Ricardo Rogriguez, Jerry Shi, Xiaofeng Shi, Carol Shou, Alex Swift, Holly Toczko, Tom Verderame, Sam Waters, Sophie Wilkowske, John Yang, Geoffrey Zee, Fred Zhu, and Jon Zytnick for research assistance. Columbia University's Program for Economic Research, Columbia Law School, Princeton University's Center for Health and Wellbeing, and the National Science Foundation Grant SES-1260875 provided financial support for this research.

Aging and Performance Decline

A. 10-K Running Time

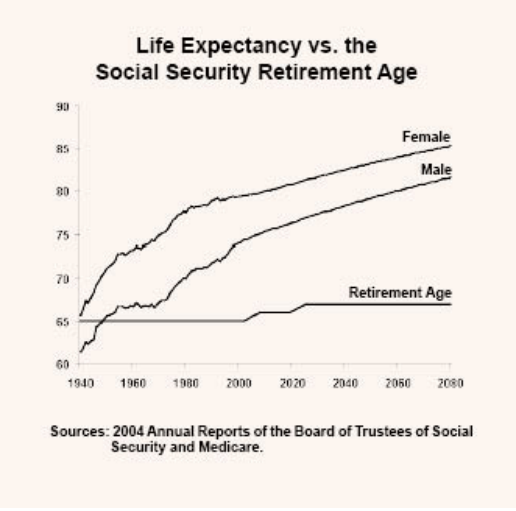


B. Cognitive Performance



- Large literature in medicine/psychology documenting changes in physical and cognitive ability due to aging (e.g. Desjardins and Warnke 2012).

Life Expectancy over Time



- Should a longer life after age 65 be allocated to leisure or to work?

U.S. Age Discrimination in Employment Act of 1967

It is therefore the purpose of this chapter to promote employment of older persons based on their ability rather than age; to prohibit arbitrary age discrimination in employment ...

The Measurement Problem

- The law requires employment to be based upon ability rather than age — the challenge is how to measure ability and performance of professionals:
  - Senior Management
  - University Faculty
  - Physicians
  - Judges (our case)
- In contrast to jobs with clear performance measures, many/most jobs entail some form of subjective evaluation, which in turn can lead to conflict regarding employment (MacLeod and Parent 1999, MacLeod 2003, Macleod, MacLeod, Valle Lara and Zehnder 2019).

Why Judges?

- We address these issues by focusing on appellate judges.
- Attractive features of this setting (e.g. Posner 2008):
  - Judges work in these positions for many years and typically retire from them.
  - The nature of tasks does not vary across the career, and does not vary over a period of decades.
  - Judges do not have much control over their workload (portfolio of cases).
  - There is no performance pay, and minimal rewards based on tenure.
  - Judge output (judicial opinions) consists of published documents, from which we can produce consistent quantitative measures of performance across the lifespan.
  - Variation across states and over time in mandatory retirement rules.

Why Mandatory Judicial Retirement?

- 1999 Report on Mandatory Retirement:
  - "In upholding mandatory retirement laws, courts routinely cite the difficulty of removing older judges with impaired mental faculties. To be sure, the embarrassing, expensive and protracted process of deciding which judges are senile and which are not is obviated by an objective age demarcation."

When Is a Judge Too Old to Judge?

Out of more than 1,200 active and senior federal judges, 16 percent will be 80 or older by the end of 2017, and 39 will be at least 90.

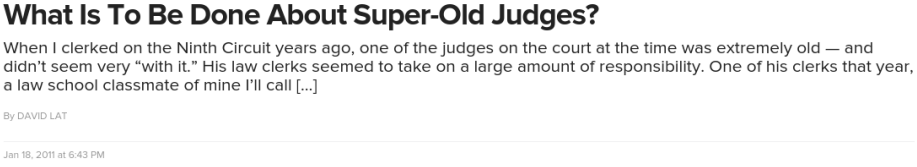
SHARE TWEET

Aaron Kase
Jul 20 2017, 6:30pm

The Oldest Bench Ever

Extreme aging in the federal judiciary—and the trouble it causes.

By Joseph Goldstein



What Is To Be Done About Super-Old Judges?

When I clerked on the Ninth Circuit years ago, one of the judges on the court at the time was extremely old — and didn't seem very "with it." His law clerks seemed to take on a large amount of responsibility. One of his clerks that year, a law school classmate of mine I'll call [...]

By DAVID LAT

Jan 18, 2011 at 6:43 PM

Judge Retirement Rules by State

A. Status Quo Rules at Period Start (1947)

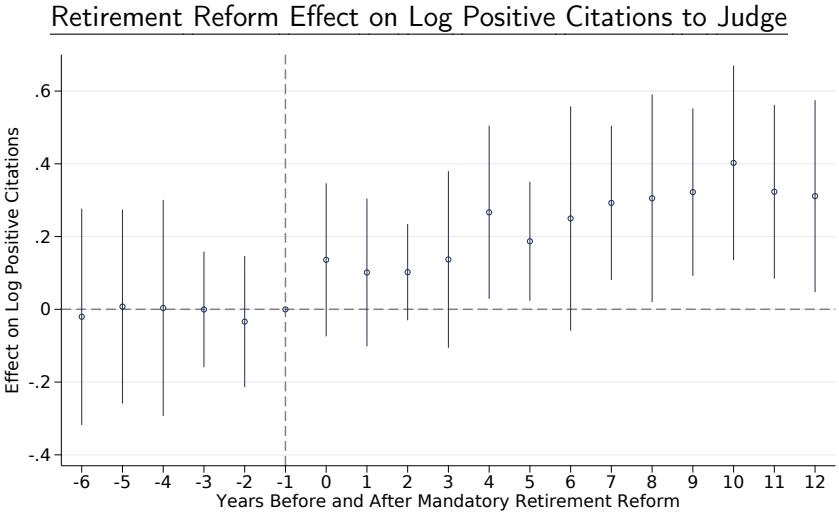
Retirement Rule	List of States
No Mandatory Retirement	AR, CA, DE, GA, ID, KY, ME, MS, MT, ND, NE, NM, NV, OK, RI, TN, WI, WV
Retirement at Age 70	AK, HI, LA, MD, MA, MI, MO, NH, NJ, NY, OH
Retirement at Age 72	NC, SC
Retirement at Age 75	IL, IN, TX, UT

B. Retirement Rule Changes, 1948-1993

Age Reform		List of States (with Year Enacted)
Before	After	
None	70	AL (1973), AZ (1992), CT (1974), FL (1972), MN (1973), PA (1968), VA (1970), WI (1955), WY (1972)
None	72	CO (1962), IA (1965), WA (1952)
None	75	KS (1993), OR (1960)

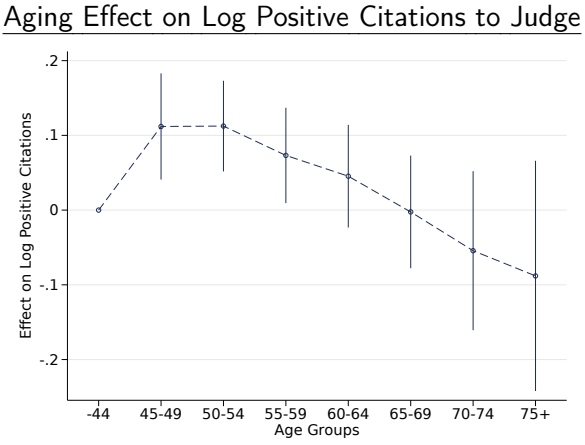
14 reforms in the time period of our data.

Result 1: Mandatory Retirement Improves Performance



Judge performance (log positive citations for a judge in a year) before/after reforms implementing retirement ages of 70, 72 or 75. Time series is a coefficient plot from the event study regression, with coefficents estimated relative to the year before the reform. Regression includes court and year fixed effects and court-specific event windows. 95% confidence intervals constructed with standard errors clustered by state.

Result 2: Aging Reduces Performance



Dynamic coefficient plots for estimates of five-year age group differences in log positive cites per judge-year, relative to the age < 45 group. Observation is a judge working in a year. All graphs contain court-year interacted fixed effects, first year baselines, and cohort fixed effects.95% confidence intervals constructed using standard errors clustered by state.

Jump to Data Description

Related Literatures

- Large literature in economics on aging and retirement (Gustman and Steinmeier 1986; Stock and Wise 1990; Lumsdaine and Mitchell 1999; Ashenfelter and Card 2002; Gustman and Steinmeier, 2005; Diermeier et al. 2005; Keane and Merlo 2010; Frederiksen and Flaherty Manchester, 2019).
- Massive interdisciplinary literature on aging and productivity (Medoff and Abraham, 1980; Abraham and Farber, 1987; Levin and Stephan, 1991; Oster and Hamermesh, 1998; Tanaka and Higuchi 1998; Blundell and Macurdy, 1999; Choudhry et al., 2005; Ballesteros et al 2009; Desjardins and Warnke, 2012; Ramscar et al. 2014; Borsch-Supan and Weiss 2016).
- Mixed evidence on aging and productivity for judges (Posner 1995; Smyth & Bhattacharya 2003; Teitelbaum 2006; Dimitrova-Grajzl et al. 2012).
  - Posner (1995) notes that judging has a "late-peaked, sustained" performance profile.
- Small literature on judge retirement choice (Nixon and Haskin 2000; Choi, Gulati, and Posner 2012).

Related Literatures

- Large literature in economics on aging and retirement (Gustman and Steinmeier 1986; Stock and Wise 1990; Lumsdaine and Mitchell 1999; Ashenfelter and Card 2002; Gustman and Steinmeier, 2005; Diermeier et al. 2005; Keane and Merlo 2010; Frederiksen and Flaherty Manchester, 2019).
- Massive interdisciplinary literature on aging and productivity (Medoff and Abraham, 1980; Abraham and Farber, 1987; Levin and Stephan, 1991; Oster and Hamermesh, 1998; Tanaka and Higuchi 1998; Blundell and Macurdy, 1999; Choudhry et al., 2005; Ballesteros et al 2009; Desjardins and Warnke, 2012; Ramscar et al. 2014; Borsch-Supan and Weiss 2016).
- Mixed evidence on aging and productivity for judges (Posner 1995; Smyth & Bhattacharya 2003; Teitelbaum 2006; Dimitrova-Grajzl et al. 2012).
  - Posner (1995) notes that judging has a "late-peaked, sustained" performance profile.
- Small literature on judge retirement choice (Nixon and Haskin 2000; Choi, Gulati, and Posner 2012).

State Courts of Last Resort

- "State supreme courts" operate similarly to U.S. Supreme Court but at state level:
  - After decision in trial court, losing party can appeal, and eventually his/her appeal may be accepted for review by the state supreme court.
  - State supreme court judges rehear the case and review submitted briefs. Judges vote whether to affirm or reverse the lower decision.
  - One of the majority judges writes an opinion explaining the decision.
  - Those decisions are binding on the state, can be cited by all courts.

State Courts of Last Resort

- "State supreme courts" operate similarly to U.S. Supreme Court but at state level:
  - After decision in trial court, losing party can appeal, and eventually his/her appeal may be accepted for review by the state supreme court.
  - State supreme court judges rehear the case and review submitted briefs. Judges vote whether to affirm or reverse the lower decision.
  - One of the majority judges writes an opinion explaining the decision.
  - Those decisions are binding on the state, can be cited by all courts.

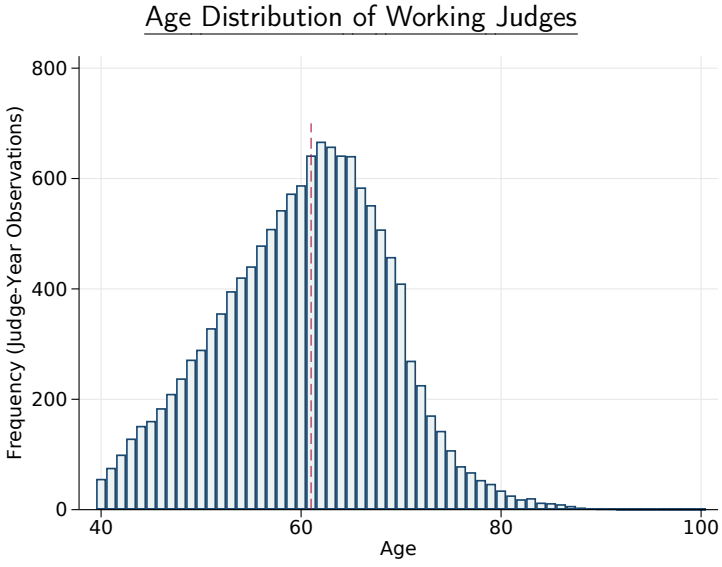
Data Overview

- We analyze a unique data set on state appellate courts.
- Previous data sets:
  - State Court Data Project: 520 judges, four years (1995-1998)
  - Choi-Gulati-Posner Group: 408 judges, three years (1998-2000)
- Our data set:
  - 1553 judges
  - 48 years (1947-1994)
- 50 states, 52 courts (Oklahoma and Texas each have two high courts each)
- 1,025,461 cases
- 1,126,560 opinions (including discretionary opinions)
- 15,486 judge-years

Judge Biographies

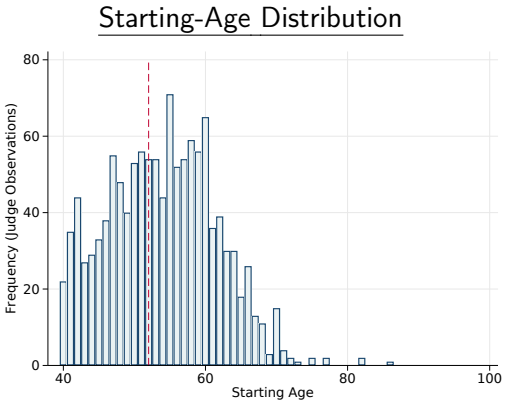
- Judge biographies:
  - Comprehensive data on judge birthdates and deathdates, how judgeships ended, and judge retirement policies.
  - Manually collected by RA's from court web sites, obits, Marquis Who's Who, etc.
- Average career length is 12 years.
- Less than 3% of judges are "promoted," where promotion is defined as becoming governor or joining a federal court.

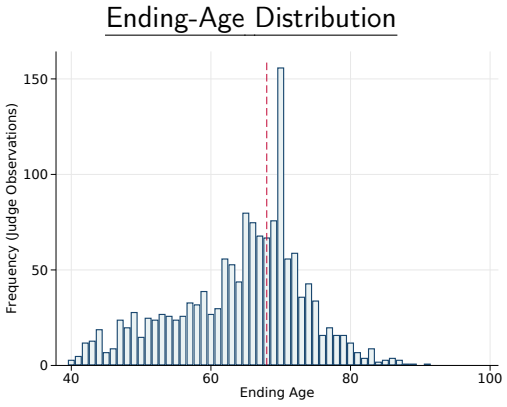
Age Distribution of Working State Supreme Court Judges



Mean: 60.4, Median: 61 (indicated by vertical dashed line), S.D.: 9.05.

Starting and Ending Age Distributions





Median indicated by vertical dashed line.

- Jump to Effect of Retirement Rule

Judge Retirement Rules by State

A. Status Quo Rules at Period Start (1947)

Retirement Rule	List of States
No Mandatory Retirement	AR, CA, DE, GA, ID, KY, ME, MS, MT, ND, NE, NM, NV, OK, RI, TN, WI, WV
Retirement at Age 70	AK, HI, LA, MD, MA, MI, MO, NH, NJ, NY, OH
Retirement at Age 72	NC, SC
Retirement at Age 75	IL, IN, TX, UT

B. Retirement Rule Changes, 1948-1993

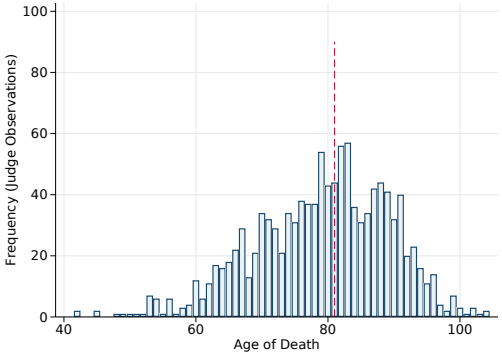
Age Reform		List of States (with Year Enacted)
Before	After	
None	70	AL (1973), AZ (1992), CT (1974), FL (1972), MN (1973), PA (1968), VA (1970), WI (1955), WY (1972)
None	72	CO (1962), IA (1965), WA (1952)
None	75	KS (1993), OR (1960)

14 reforms in the time period of our data.

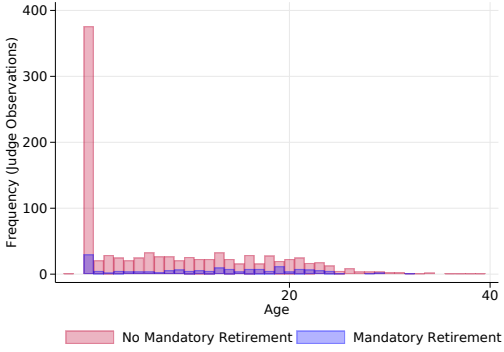
Retirement Rates by Age, by Mandatory Retirement Age

End of the judge life cycle

Judge Age-at-Death Distribution



Distribution of Years Between Termination and Death, With/Without Mandatory Retirement



Performance Data

- Judicial Opinion Data:
  - All authored opinions between 1947 and 1994, collected from Bloomberg Law.
  - Exclude per curiam / memorandum opinions that do not have a named author
- Measuring Judge Performance:
  - log positive citations by later judges in a year
    - citations of cases occur over long time period (median ten years delay), so judges don't have much interpersonal influence on citations.
    - also restrict to citations from judges in other states.
- Quantity vs. Quality:
  - quantity: number of opinions
  - quality: citations per opinion written

Performance Data

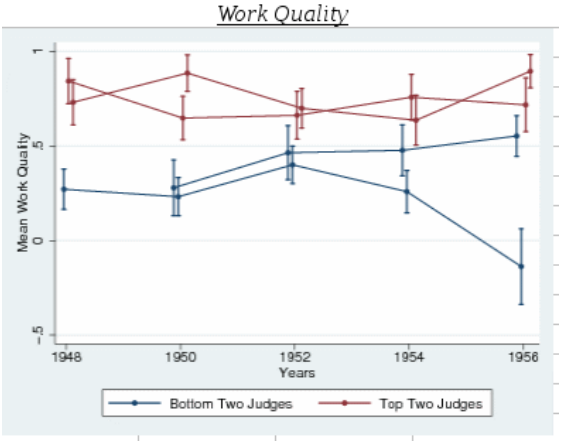
- Judicial Opinion Data:
  - All authored opinions between 1947 and 1994, collected from Bloomberg Law.
  - Exclude per curiam / memorandum opinions that do not have a named author
- Measuring Judge Performance:
  - log positive citations by later judges in a year
    - citations of cases occur over long time period (median ten years delay), so judges don't have much interpersonal influence on citations.
    - also restrict to citations from judges in other states.
- Quantity vs. Quality:
  - quantity: number of opinions
  - quality: citations per opinion written

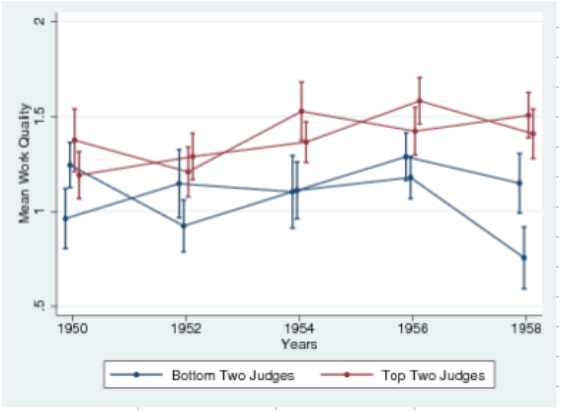
Previous analysis of these outcomes

- Ash and MacLeod (2015):
  - Judges respond to relaxation of time pressure with higher work quality.
  - Consistent with "intrinsic motivation" or "professionalism"
- Ash and MacLeod (2020):
  - Nonpartisan elections and merit systems select better judges than partisan elections.

Is there a Judge Fixed Effect?

Ash and MacLeod (2020) show that judge fixed effects explain a large amount of variation in quality, and that judges have persistence over time:





- Massachusetts, 1947-1956, and California, 1949-1958. Normalized within judge.

Case Quality Correlated with Bar Association Evaluations

	Logit Estimate for Effect on "Good Judge" Designation		
	(1)	(2)	(3)
Output	0.154		-0.0771
	(1.046)		(1.100)
Quality		1.059**	1.076**
		(0.363)	(0.112)
State Fixed Effects	X	X	X
Year Fixed Effects	X	X	X

N= 51 judge-bienniums for set of judges in Pennsylvania, Texas, and Washington for the years 1987 through 1994. Outcome is an indicator for being a "good" judge has defined in Lim and Snyder (2015). Standard errors clustered by state in parentheses. + p < .1, * p < .05, ** p < .01.

From Ash and MacLeod (2020).

Differences-in-Differences Approach

Judge i, state s, year t:

$$y_{ist} = \alpha_s + \alpha_t + \rho M_{st} + \alpha_s \cdot t + X_s^0 \cdot \alpha_t + X_{ist}'\beta + \varepsilon_{ist} \quad (1)$$

- y_{ist} = judge age or annual performance metric (e.g. log citations).
- α_s = court fixed effects, α_t = time fixed effects, $\alpha_s \cdot t$ = state trends
- M_{st} = treatment indicators
  - use "treatment windows" of (e.g.) ten years before/after reform
- $X_s^0 \cdot \alpha_t$, initial-period characteristics (retention rules, case types, age distribution, interacted with year fixed effects.
- X_{ist}, other time-varying controls.
- Cluster standard errors by state.

Differences-in-Differences Approach

Judge i, state s, year t:

$$y_{ist} = \alpha_s + \alpha_t + \rho M_{st} + \alpha_s \cdot t + X_s^0 \cdot \alpha_t + X_{ist}'\beta + \varepsilon_{ist} \quad (1)$$

- y_{ist} = judge age or annual performance metric (e.g. log citations).
- α_s = court fixed effects, α_t = time fixed effects, $\alpha_s \cdot t$ = state trends
- M_{st} = treatment indicators
  - use "treatment windows" of (e.g.) ten years before/after reform
- $X_s^0 \cdot \alpha_t$, initial-period characteristics (retention rules, case types, age distribution, interacted with year fixed effects.
- X_{ist}, other time-varying controls.
- Cluster standard errors by state.

Effect of Retirement Reform on Court Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Age (Mean)		Age (Min)	Age (Q25)	Age (Median)	Age (Q75)	Age (Max)
Ret. Reform	-2.065*	-3.924**	-3.444**	-3.804**	-4.386**	-3.956**	-3.381*
	(0.886)	(0.796)	(1.239)	(1.287)	(0.833)	(1.098)	(1.312)
Year FE, Court FE	X	X	X	X	X	X	X
Windows/Trends		X	X	X	X	X	X
N	14775	14775	14775	14775	14775	14775	14775
R-sq	0.494	0.648	0.598	0.569	0.614	0.616	0.638

DD effect of mandatory retirement reform on judge age statistics in ten years after reform, relative to ten years before reform. Observation is a judge working in a year. "Ret. Reform" is a treatment indicator for the ten years after the introduction of mandatory retirement. Dependent variables are computed at the court-year level. In particular, "Age (Mean)" is the average age of judges in each court and year, "Age (Min)" the minimum age, "Age (Q25)" the age at the 25th percentile, "Age (Median)" the median age, "Age (Q75)" the age at the 75th percentile and "Age (Max)" the maximum age. Court Treat Windows means court-specific treatment windows (ten years before and after reform). Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

Jump to Effect of Age on Performance

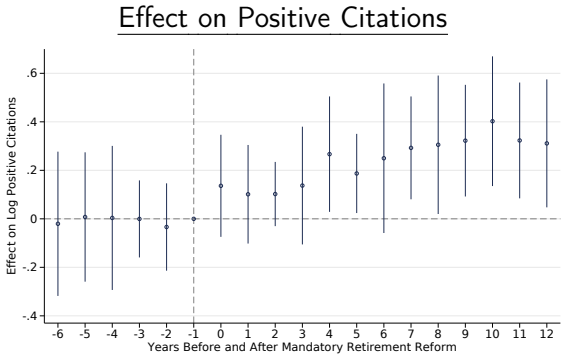
Event-Study Approach

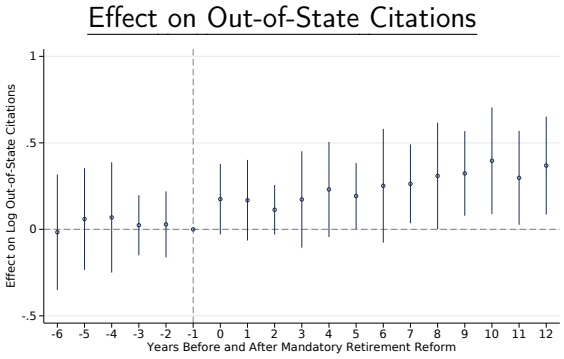
To formally test for parallel trends, we estimate

$$y_{ist} = \alpha_s + \alpha_t + \sum_{k=-6, k\neq -1}^{12} \rho_k M_{st}^k + \alpha_s \cdot t + X_s^0 \cdot \alpha_t + X_{ist}' \beta + \varepsilon_{ist} \quad (2)$$

- M_{st}^k = event-study year indicators, equalling one for years relative to reform (indexed by k)
  - Parallel trends are consistent with $\hat{\rho}_k = 0$ for $k < -1$.
  - For $k >= 0$, $\hat{\rho}_k$ will elucidate the dynamics of the difference-in-differences effect.

Effect of Retirement Reform on Performance, Event Study





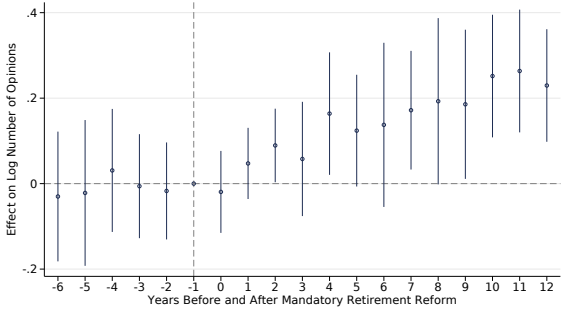
Effect of Mandatory Retirement Reform on Log Citations

	(1)	(2)	(3)	(4)	(5)
	Effect on Log Positive Cites per Judge				
Retirement Reform	0.228**	0.253**	0.237**	0.322**	0.328**
	(0.0756)	(0.0836)	(0.0818)	(0.0899)	(0.0880)
Court FE, Year FE	X	X	X	X	X
Court Trends/Windows		X	X	X	X
Init Court Rules × Year FE			X	X	X
Init Case Types × Year FE				X	X
Init Age × Year FE					X
N	15010	15010	15010	15010	15010
R-sq	0.460	0.526	0.538	0.555	0.565

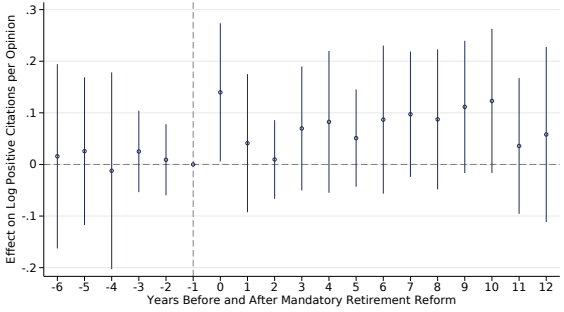
Notes. DD effect of mandatory retirement reform on log positive citations to a judge's opinions in ten years after reform, relative to ten years before reform. Observation is a judge working in a year. "Ret. Reform" is a treatment indicator for the ten years after the introduction of mandatory retirement. Court Treat Windows means court-specific treatment windows (ten years before and after reform). "Init X" × year FE means initial values are interacted with year. "Init Court Rules" includes a state's 1947 rules for judge selection/retention system, admin office, intermediate appellate court, number of judges, and term length. "Init Case Types" includes a court's 1947 average values for case characteristics (legal area and related industries). "Init Age" includes the initial mean and standard deviation for judge age on the court. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

Is it quantity or quality? Both.

Effect on Quantity: Log Number of Opinions



Effect on Quality: Positive Cites per Opinion



Judge performance effects before and after reforms implementing retirement ages of 70, 72 or 75. Top panel outcome is log number of opinions by a judge in a year; bottom panel is average log positive citations per opinon. Time series is a coefficient plot from the event study regression, with coefficents estimated relative to the year before the reform. Regression includes court and year fixed effects and court-specific event windows. 95% confidence intervals constructed with standard errors clustered by state.

Jump to Age Effect

Regression Table

Robustness for Main Retirement Reform Results

1. Result holds (about half the effect size) with judge fixed effects Regression Table
2. Effects not driven by types of cases ruled on Regression Table
3. Additional measures of performance Regression Table
4. Only treated states Figure
5. Only the event study window Figure
6. More pre-periods Figure
7. Dropping each state individually Figure
8. Separating the age reforms Regression Table
9. Alternative weighting/clustering Weighting Clustering
10. Time-varying controls Regression Table
11. Add lagged dependent variable Table

Quadratic Age Model

Judge i, state s, year t:

$$y_{ist} = \alpha_{st} + \alpha_i^0 + \gamma_1 A_{ist} + \gamma_2 A_{ist}^2 + X_{ist}'\beta + \varepsilon_{ist} \tag{3}$$

- y_{ist} = annual performance metric (e.g. log citations).
  - for this analysis, also use within-court-year rank percentile in citations.
- α_{st} = court-year interacted fixed effects
- α_i^0 = baseline outcome value for juge's first year on court
- A_{ist} = judge age (in years)
- X_{ist}, other time-varying controls.
- Cluster standard errors by state.

Quadratic Age Model

Judge i, state s, year t:

$$y_{ist} = \alpha_{st} + \alpha_i^0 + \gamma_1 A_{ist} + \gamma_2 A_{ist}^2 + X_{ist}'\beta + \varepsilon_{ist} \tag{3}$$

- y_{ist} = annual performance metric (e.g. log citations).
  - for this analysis, also use within-court-year rank percentile in citations.
- α_{st} = court-year interacted fixed effects
- α_i^0 = baseline outcome value for juge's first year on court
- A_{ist} = judge age (in years)
- X_{ist}, other time-varying controls.
- Cluster standard errors by state.

Performance Falls with Age

	(1)	(2)	(3)	(4)	(5)	(6)
		Log Positive Cites			Rank Percentile Cites	
Judge Age (Years)	-0.00797**	-0.00790**	-0.00702**	0.0351*	-0.00428**	0.0185+
	(0.00140)	(0.00114)	(0.00127)	(0.0133)	(0.000833)	(0.00939)
Age Squared				-0.000356**		-0.000192*
				(0.000118)		(0.0000824)
Court-Year FE	X	X	X	X	X	X
First-Year Baseline		X	X	X	X	X
Cohort FE / Trends			X	X	X	X
N	13655	13655	13655	13655	13646	13646
R-sq	0.674	0.694	0.701	0.702	0.112	0.115

Observation is a judge working in a year. "Log Positive Cites" is log of positive cites to a judge in a year. "Rank Percentile Cites" means judges are uniformly distributed between zero and one based on number of positive citations within court-year (0 is lowest, 1 is highest). Court-Year FE is interacted court-year fixed effects. First-Year Baseline means a judge's value for the outcome in their first year on the court is included as a control. Cohort FE means fixed effect for decade that the judge started on the court. Cohort Trends means judge starting-year interacted with court fixed effect. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

Jump to Conclusions

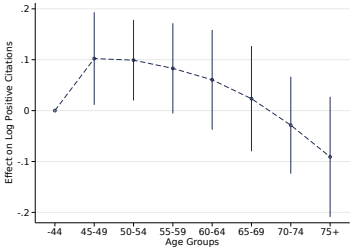
Dynamic Age Groups Model

We plot out the dynamic changes in performance over the lifespan by estimating

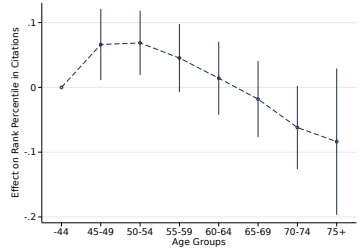
$$y_{ist} = \alpha_{st} + \alpha_i^0 + \sum_{g \in G} \gamma_g A_{ist}^g + X_{ist}'\beta + \varepsilon_{ist} \tag{4}$$

- A_{ist}^g, $g \in G$, equaling one when judge i is in age group g.
  - Age groups are 0-44, 45-49, 50-54, 55-59, 60-64, 65-60, 61-64, 65-69, 70-74, and 75+.
  - Distribution over Age Groups

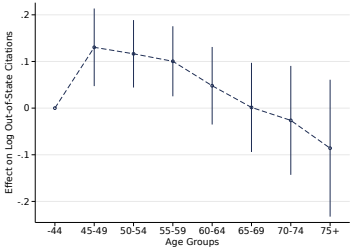
(A) Log Positive Cites



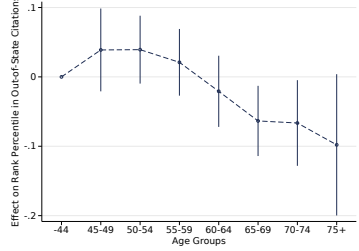
(B) Rank Percentile Positive Cites



(C) Log Out-of-State Cites

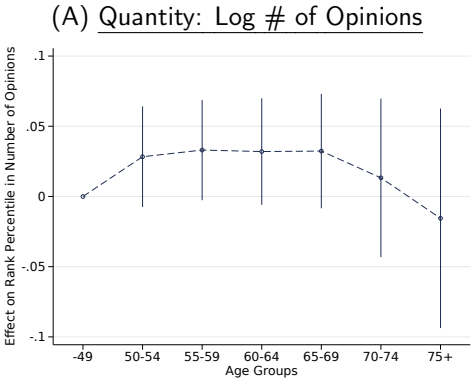


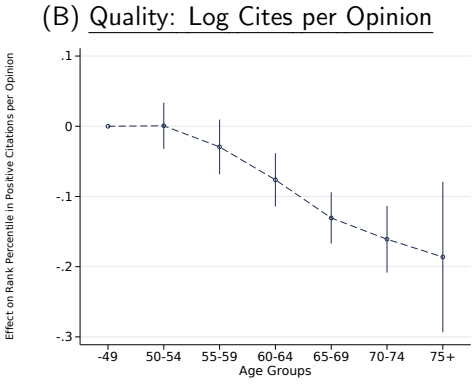
(D) Rank Percentile Out-of-State Cites



Dynamic coefficient plots for estimates of five-year age group differences, relative to the age < 45 group. Observation is a judge working in a year. All graphs contain court-year interacted fixed effects, first year baselines, and cohort fixed effects. Outcomes are in logs or rank percentiles, as indicated. 95% confidence intervals constructed using standard errors clustered by state.

Aging affects quality, not quantity





Dynamic coefficient plots for estimates of five-year age group differences, relative to the age < 45 group. Observation is a judge working in a year. All graphs contain court-year interacted fixed effects, first year baselines, and cohort fixed effects. 95% confidence intervals constructed using standard errors clustered by state.

Robustness for Aging and Performance Results

1. Additional measures of performance Regression Table
2. No effect on types of cases assigned Regression Table
3. Years-on-court fixed effects Figure
4. Separately with/without mandatory retirement rule Regression Table
5. Before/after 1970 Button
6. Alternative weighting/clustering Button
7. 2SLS specification Regression Table

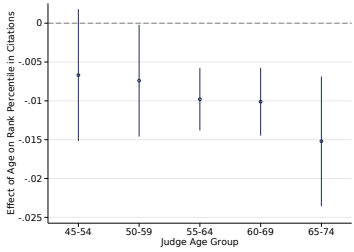
Why not judge fixed effects?

- Age (in years) is perfectly linear within judge.
  - Therefore, γ is not identified when including both judge fixed effects and year (or court-year) fixed effects.
- There is no straight-forward estimation approach that would allow judge fixed effects and also account for the large global variation in citations over time across all courts.
  - Further, with judge fixed effects the age and experience effect cannot be distinguished.

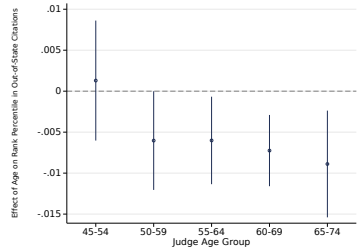
Analysis in Age-Balanced Samples

- A potential issue with the previous specifications is that they are applied to an unbalanced sample of judges.
  - Judges start and end at different ages, so the estimated effects could be driven by selection of different types of judges into different starting and ending ages.
- Next we produce regressions using balanced samples of judges:
  - take overlapping ten-year age groups across the lifespan
  - (45-54, 50-59, 55-64, 60-69, 65-74)
  - limit to judges that worked continuously in that age group.
  - Distribution
- Estimate age-performance trend restricting to the balanced samples.
  - preferred outcome is rank percentile in citations, which is comparable across courts and over time.

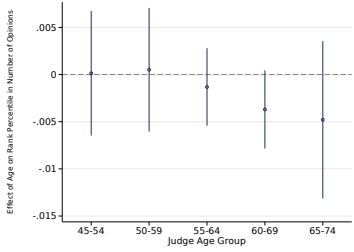
(A) Rank Percentile Positive Cites



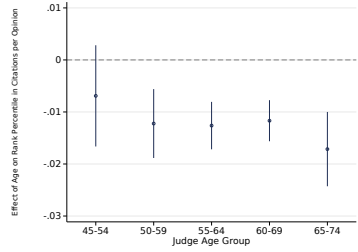
(B) Rank Percentile Out-of-State Cites



(C) Rank Percentile # of Opinions

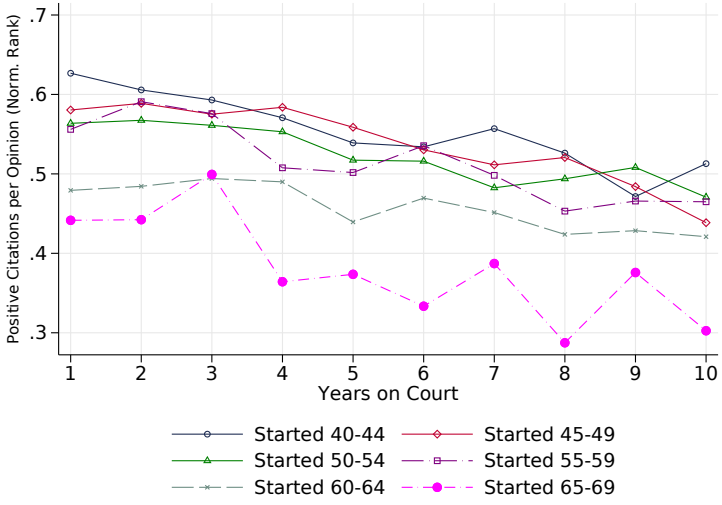


(D) Rank Percentile Cites per Opinion



Performance-Age estimates for separate balanced samples of judges based on age group. Observation is a judge working in a year. Outcomes are in rank percentiles, regressed on age (in years) for the specified group. 95% confidence intervals constructed using standard errors clustered by state.

Effect of Starting Age in First Years on Court



Notes. Time series for average rank percentile (within court year) in positive citations for the first ten years of a judge career, separately by starting age (indicated in legend).

Better Judges at a Given Age will Stay on Court Longer

	(1)	(2)	(3)	(4)	(5)
	Effect on Log Positive Cites				
Judge Start Age	-0.00700**	-0.000885			-0.00147
	(0.00155)	(0.00232)			(0.00222)
Judge End Age			-0.000612	**0.0151****	**0.0152****
			(0.00132)	**(0.00177)**	**(0.00181)**
Court-Year FE	X	X	X	X	X
Age FE		X		X	X
N	13655	13643	14618	13643	13643
adj. R-sq	0.672	0.678	0.668	0.682	0.682

Effect of judge start age and end age on judge performance. Observation is a judge working in a year. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

Summary

- Appellate courts provide an attractive setting for empirical work on the aging and productivity upon professionals.
- The introduction of a mandatory retirement age increases performance, both in terms of quantity and quality.
- Physical aging is associated with a reduction in quality over the lifespan, particularly the last few years.

Policy Implications

- Results support efforts to introduce mandatory retirement at the federal level.
- Could use these results, and a few assumptions, to compute an optimal retirement age.
- Future work can further unpack quality by looking at language used in opinions.

Thanks!

Elliott Ash | ashe@ethz.ch |elliottash.com

W. Bentley MacLeod | bentley.macleod@columbia.edu | wbmacleod.net

Policy Implications

- Results support efforts to introduce mandatory retirement at the federal level.
- Could use these results, and a few assumptions, to compute an optimal retirement age.
- Future work can further unpack quality by looking at language used in opinions.

Thanks!

Elliott Ash | ashe@ethz.ch |elliottash.com

W. Bentley MacLeod | bentley.macleod@columbia.edu | wbmacleod.net

6 Appendix Slides

Effect of Reform on Quantity/Quality

	(1)	(2)	(3)	(4)	(5)	(6)
	Quantity (Log Number of Opinions)			Quality (Log Positive Cites per Case)		
Judge Age	-0.000991	-0.00310**	0.0296**	-0.00601**	-0.00328**	0.00319
	(0.000768)	(0.000798)	(0.00581)	(0.000781)	(0.000711)	(0.00707)
Age Squared			-0.000276**			-0.0000547
			(0.0000492)			(0.0000627)
Court-Year FE	X	X	X	X	X	X
1st-Year Base		X	X		X	X
Cohort FE/Trend		X	X		X	X
N	13655	13655	13655	13655	13655	13655
R-sq	0.684	0.711	0.713	0.804	0.827	0.827

Observation is a judge working in a year. Outcomes are in logs: number of opinions, or citations per opinion. Court-Year FE is interacted court-year fixed effects. 1st-Year Base means a judge's value for the outcome in their first year on the court is included as a control. Cohort FE means fixed effect for decade that the judge started on the court. Cohort Trends means judge starting-year interacted with court fixed effect. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

Back

Effect of Mandatory Retirement Reform, Judge Fixed Effects

	(1)	(2)	(3)	(4)	(5)
	Effect on Log Cites per Judge				
Retirement Reform	0.175*	0.173*	0.170*	0.221*	0.210*
	(0.0768)	(0.0726)	(0.0728)	(0.0872)	(0.0801)
Court FE, Year FE	X	X	X	X	X
Judge FE	X	X	X	X	X
Court Trends/Windows		X	X	X	X
Init Court Rules × Year FE			X	X	X
Init Case Types × Year FE				X	X
Init Age × Year FE					X
N	14905	14905	14905	14905	14905
R-sq	0.675	0.683	0.691	0.700	0.710

Observation is a judge working in a year. "Retirement Reform" is an indicator for the ten years after the introduction of mandatory retirement. Court Treat Windows means court-specific treatment windows (ten years before and after reform). "Init X" × year FE means initial values are interacted with year. "Init Court Rules" includes a state's 1947 rules for judge selection/retention system, admin office, intermediate appellate court, number of judges, and term length. "Init Case Types" includes a court's 1947 average values for case characteristics (legal area and related industries). "Init Age" includes the initial mean and standard deviation for judge age on the court. Standard errors clustered by state in parentheses.

+ p<.0.1, * p<0.05, ** p<0.01. Back

Effect of Mandatory Retirement Reform, Case Characteristics

	(1)	(2)	(3)	(4)
	Share Criminal Cases		Case Importance	
Retirement Reform	0.0400*	0.0301*	0.0570	0.0350
	(0.0157)	(0.0133)	(0.0376)	(0.0300)
Year FE, Court FE	X	X	X	X
Court Treat Windows		X		X
N	15010	15010	15010	15010
R-sq	0.596	0.649	0.394	0.472

Observation is a judge working in a year. "Retirement Reform" is an indicator for the ten years after the introduction of mandatory retirement. Court Treat Windows means court-specific treatment windows (ten years before and after reform). "Share Criinal Cases" is the share on criminal law. "Case Importance" is the predicted citations to a case based on case characteristics (legal area and related industries). Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01. Back

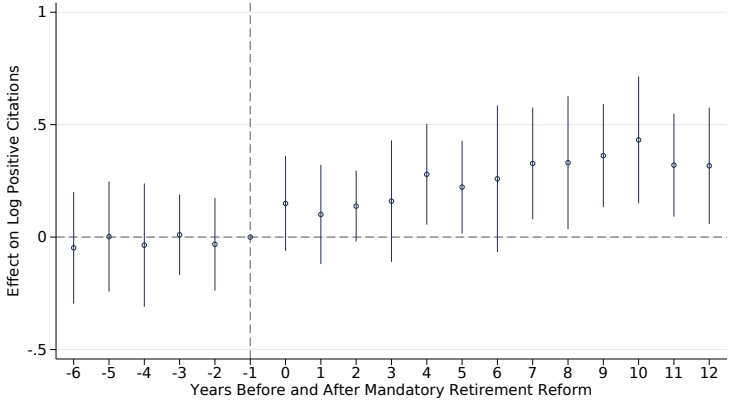
Effect of Mandatory Retirement Reform, Other Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	# of Opinions		Work Output		Work Quality		Out-of-State Cites	
70/72 Retire Reform	0.136*	0.112+	0.0751	0.0775	0.0855+	0.0926*	0.173	0.191+
	(0.0538)	(0.0574)	(0.0695)	(0.0656)	(0.0484)	(0.0439)	(0.117)	(0.0978)
Year FE	X	X	X	X	X	X	X	X
Court FE	X	X	X	X	X	X	X	X
Court Treat Windows		X		X		X		X
Court Trends		X		X		X		X
Rule Controls		X		X		X		X
N	15010	13863	15010	13863	15010	13863	15010	13863
R-sq	0.325	0.512	0.266	0.386	0.649	0.718	0.471	0.521

Observation is a judge working in a year. "70/72 Retirement Reform" is an indicator for the ten yeras after the introduction of mandatory retirment at ages 70 or 72. "# of Opinions" is the number of majority opinions written by a judge in a year. "Work Output" is log number of words writen in a year. "Work Quality" is number of citations per published opinion. "Total Out-of-State Cites" is Court Treat Windows means court-specific treatment windows (ten years before and after reform). Rule controls include dummies for changes to the electoral system, number of judges, and expenditures on judicial system. Time-served controls include a quadratic in years on court. Case controls means the first five principal components of the matrix of controls for legal topic and related industries. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

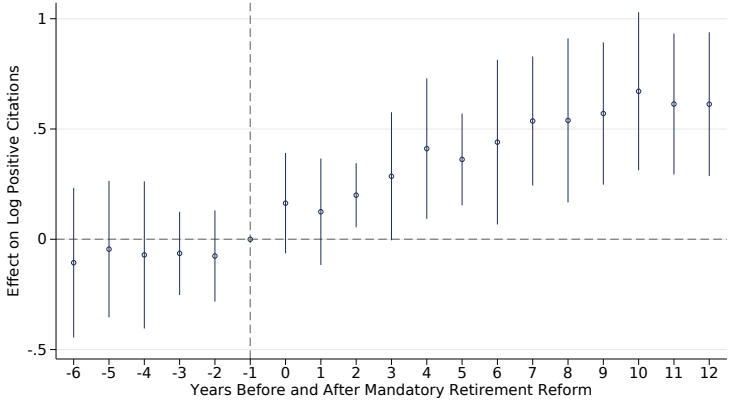
Back

Effect of Mandatory Retirement Reform, Treated States



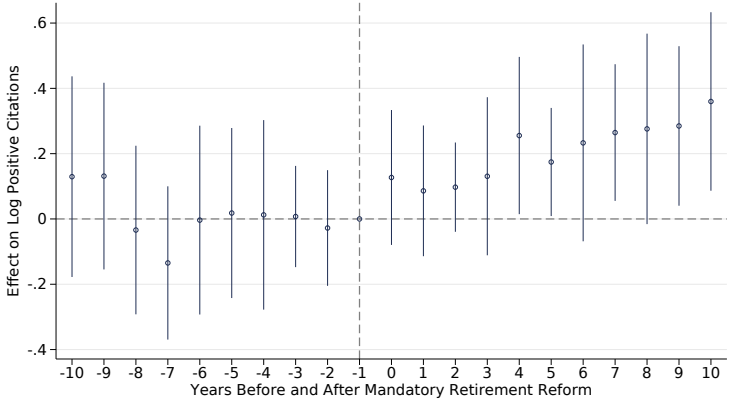
Back

Effect of Mandatory Retirement Reform, Event Study Window



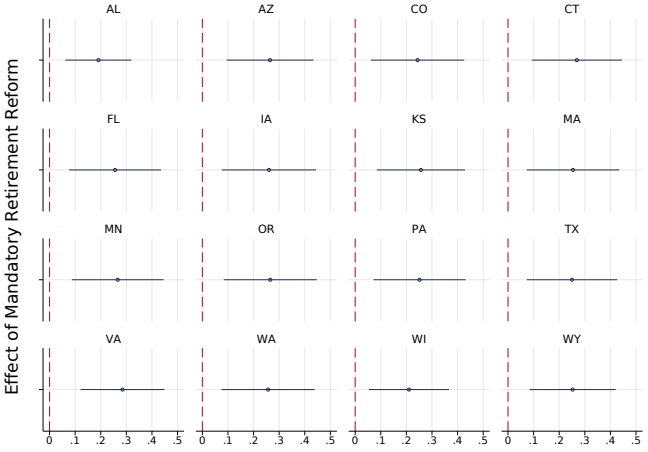
Back

Effect of Mandatory Retirement Reform, More Pre-Periods



Back

Effect of Mandatory Retirement Reform, Dropping each State Individually



Back

Effect of Mandatory Retirement Reform, Separating Age reforms

	(1)	(2)	(3)	(4)	(5)	(6)
Maximum Age	70		72		75	
Retirement Reform	0.212+	0.291*	0.396**	0.310**	0.155+	0.165**
	(0.117)	(0.133)	(0.113)	(0.0773)	(0.0899)	(0.0384)
Court FE, Year FE	X	X	X	X	X	X
Court Trends/Windows		X		X		X
N	15010	15010	15010	15010	15010	15010
R-sq	0.459	0.524	0.459	0.518	0.458	0.517

Notes. Observation is a judge working in a year. "Retirement Reform" is an indicator for the ten years after the introduction of mandatory retirement. Court Treat Windows means court-specific treatment windows (ten years before and after reform). Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

Back

Effect of Mandatory Retirement Reform, Alternative Weighting

	(1)	(2)	(3)	(4)
	Effect on Log Cites per Judge			
Retirement Reform	0.211**	0.253**	0.133+	0.181*
	(0.0750)	(0.0858)	(0.0736)	(0.0802)
Weighting	# of Opinions		Judges Equal	
Court FE, Year FE	X	X	X	X
Court Trends/Windows		X		X
N	15010	15010	14997	14997
R-sq	0.496	0.569	0.421	0.499

Observation is a judge working in a year. "Retirement Reform" is an indicator for the ten years after the introduction of mandatory retirement. Court Treat Windows means court-specific treatment windows (ten years before and after reform). Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

Back

Effect of Mandatory Retirement Reform, Alternative Clustering

	(1)	(2)	(3)	(4)	(5)	(6)
Clustering Group	State and Year		Judge		None (Robust)	
Retirement Reform	0.228**	0.253**	0.228**	0.253**	0.228**	0.253**
	(0.0458)	(0.0436)	(0.0438)	(0.0412)	(0.0291)	(0.0300)
Court FE, Year FE	X	X	X	X	X	X
Court Trends/Windows		X		X		X
N	15010	15010	15010	15010	15010	15010
R-sq	0.460	0.526	0.460	0.526	0.460	0.526

Notes. Observation is a judge working in a year. "Retirement Reform" is an indicator for the ten years after the introduction of mandatory retirement. Court Treat Windows means court-specific treatment windows (ten years before and after reform). Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

Back

Effect of Mandatory Retirement Reform, Time Varying Controls

	(1)	(2)	(3)	(4)	(5)	(6)
	Effect on Log Positive Cites per Judge					
Retirement Reform	0.106+	0.148*	0.251**	0.271**	0.202**	0.225**
	(0.0572)	(0.0703)	(0.0886)	(0.0804)	(0.0725)	(0.0804)
Court FE, Year FE	X	X	X	X	X	X
Court Trends/Windows		X		X		X
Case Controls	X	X				
Rule Controls			X	X		
Judge Experience FE					X	X
N	13304	13304	13304		13304	13304
R-sq	0.585	0.609	0.618		0.630	0.638

Observation is a judge working in a year. "Retirement Reform" is an indicator for the ten years after the introduction of mandatory retirement. Court Treat Windows means court-specific treatment windows (ten years before and after reform).Case controls means the first five principal components of the matrix of controls for legal topic and related industries. Rule controls means rules for selection and retention of juges and other institutional items. Judge Experience FE means fixed effects for yeras on the court. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01. Back

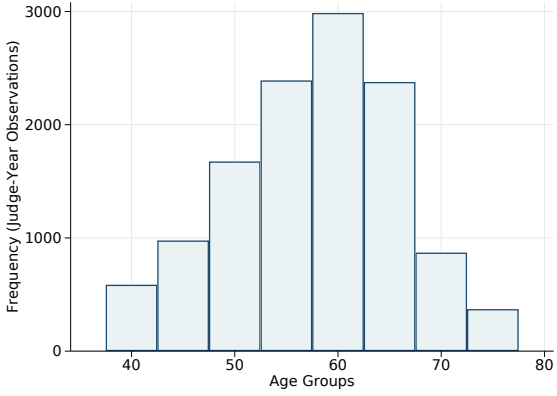
Effect of Mandatory Retirement Reform, Lagged Outcome

	(1)	(2)	(3)	(4)	(5)
	Effect on Log Positive Cites per Judge				
Retirement Reform	0.143**	0.172*	0.165*	0.209**	0.206**
	(0.0504)	(0.0644)	(0.0661)	(0.0711)	(0.0635)
Court FE, Year FE	X	X	X	X	X
Lagged y_{ist}	X	X	X	X	X
Court Trends/Windows		X	X	X	X
Init Court Rules × Year FE			X	X	X
Init Case Types × Year FE				X	X
Init Age × Year FE					X
N	13304	13304	13304	13304	13304
R-sq	0.585	0.609	0.618	0.630	0.638

Observation is a judge working in a year. "Retirement Reform" is an indicator for the ten years after the introduction of mandatory retirement. Includes lagged outcome variable by judge in the regression. Court Treat Windows means court-specific treatment windows (ten years before and after reform). "Init X" × year FE means initial values are interacted with year. "Init Court Rules" includes a state's 1947 rules for judge selection/retention system, admin office, intermediate appellate court, number of judges, and term length. "Init Case Types" includes a court's 1947 average values for case characteristics (legal area and related industries). "Init Age" includes the initial mean and standard deviation for judge age on the court. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01.

Back

Distribution of Age Groups



Number of judge-year observations in each five-year age group for the life cycle coefficient plots. Back

Effect of Aging, Additional Measures of Performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Cites in 10 Years		All Cites		Discuss Cites		Out-of-State Cites	
Judge Age	-0.00693**	0.0375**	-0.00739**	0.0348*	-0.00836**	0.0373**	-0.00885**	0.0323*
	(0.00135)	(0.0127)	(0.00129)	(0.0138)	(0.00116)	(0.0119)	(0.00143)	(0.0139)
Age Squared		-0.000375**		-0.000356**		-0.000386**		-0.000348**
		(0.000112)		(0.000122)		(0.000105)		(0.000123)
Court-Year FE	X	X	X	X	X	X	X	X
1st-Year Base	X	X	X	X	X	X	X	X
Cohort FE/Trend	X	X	X	X	X	X	X	X
N	13655	13655	13655	13655	13655	13655	13655	13655
R-sq	0.768	0.769	0.697	0.698	0.698	0.690	0.691	0.674

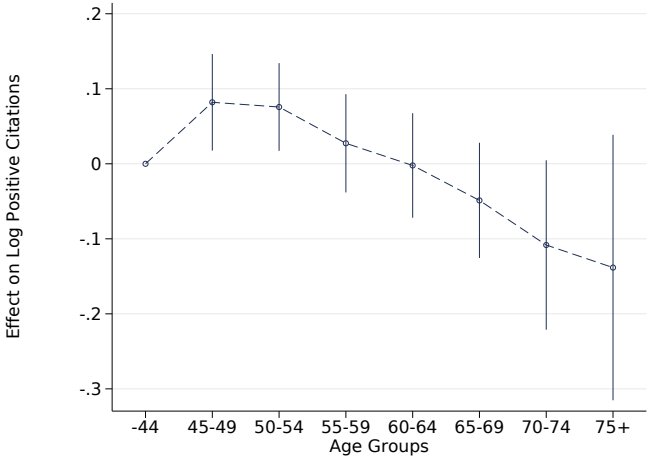
Observation is a judge working in a year. Outcomes are in logs. "Cites in 10 years" is log of positive cites to a judge in a year, within ten years of a case. "All Cites" includes negative and distinguishing (not just positive) cites. "Discuss cites" means the case was positively discussed and applied. "Out-of-state cites" means citations from courts in other states. Court-Year FE is interacted court-year fixed effects. 1st-Year Base means a judge's value for the outcome in their first year on the court is included as a control. Cohort FE means fixed effect for decade that the judge started on the court. Cohort Trends means judge starting-year interacted with court fixed effect. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01. Back

No selection into case types by age

	(1)	(2)	(3)	(4)	(5)
			Case Type		
	Crim	Civil	Admin	Con Law	Pred. Cites
Age × Random	0.00427	-0.00435	-0.0164	-0.0196	-0.00127
	(0.00845)	(0.00700)	(0.0115)	(0.0129)	(0.00188)
Age × Not Rand	0.0265	-0.0198	-0.00161	-0.0131	-0.0000133
	(0.0209)	(0.0230)	(0.0176)	(0.0194)	(0.00229)
Court-Year FE	X	X	X	X	X
N	13643	13643	13607	13632	13599
adj. R-sq	0.140	0.209	-0.062	-0.042	0.397

"Random" means random-assignment states, "Not Rand" means discretionary assignment. Age is standardized within court-year. "Crim" means proportion of cases on criminal law in a year (respectively for civil, administrative, and constitutional law). "Pred. Cites" means predicted case quality from OLS regression with case characteristics (legal area and related industries). Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01. Back

Effect of Aging, Years-on-Court Fixed Effects



Dynamic coefficient plots for estimates of five-year age group differences, relative to the age < 45 group. Outcom is log positive cites. Observation is a judge working in a year. Panel A has no fixed effects; Panel B contains court-year interacted fixed effects, first year baselines, cohort fixed effects, and fixed effects for number of years on court. 95% confidence intervals constructed using standard errors clustered by state.

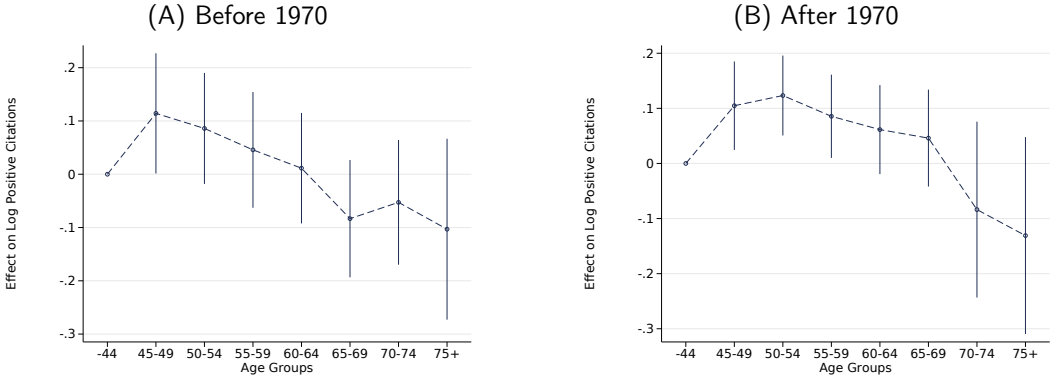
Back

Effect of Aging, with/without Mandatory Retirement

	(1)	(2)	(3)	(4)	(5)	(6)
	Voluntary Retirement			Mandatory Retirement		
	Log Cites		Rank	Log Cites		Rank
Judge Age (Years)	-0.00962**	0.0181	-0.00576**	-0.00615**	0.0519**	-0.00390**
	(0.00320)	(0.0249)	(0.00147)	(0.00160)	(0.0137)	(0.00122)
Age Squared		-0.000231			-0.000494**	
		(0.000227)			(0.000119)	
Court-Year FE	X	X	X	X	X	X
First-Year Baseline	X	X	X	X	X	X
Cohort FE / Trends	X	X	X	X	X	X
N	4688	4688	4688	8967	8967	8967
R-sq	0.692	0.692	0.059	0.613	0.616	0.043

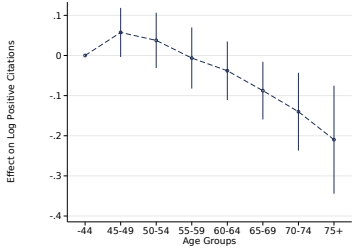
Observation is a judge working in a year. "Log Positive Cites" is log of positive cites to a judge in a year. "Rank Percentile Cites" means judges are uniformly distributed between zero and one based on number of positive citations within court-year (0 is lowest, 1 is highest). Court-Year FE is interacted court-year fixed effects. First-Year Baseline means a judge's value for the outcome in their first year on the court is included as a control. Cohort FE means fixed effect for decade that the judge started on the court. Cohort Trends means judge starting-year interacted with court fixed effect. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01. Back

Effect of Aging, Before/After 1970

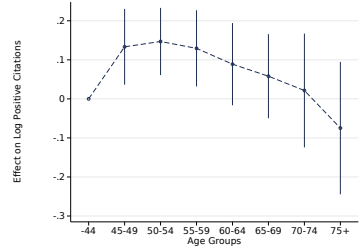


Dynamic coefficient plots for estimates of five-year age group differences, relative to the age < 45 group. Observation is a judge working in a year. All graphs contain court-year interacted fixed effects, first year baselines, and cohort fixed effects. Outcomes are log positive cites to a judge in a year. 95% confidence intervals constructed using standard errors clustered by state. Back

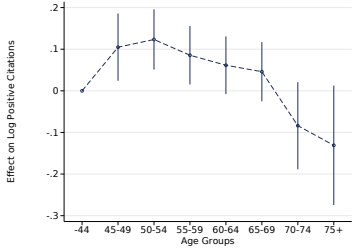
(A) Weighting by Number of Opinions



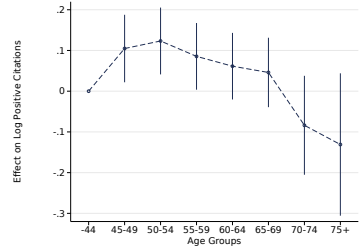
(B) Weighting by Inverse Career Length



(C) Clustering at State-Year Level



(D) Clustering at Judge Level

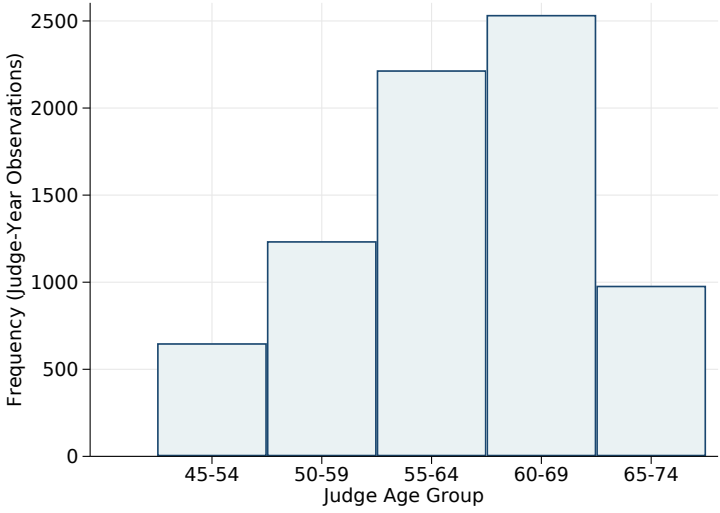


Effect of Aging, 2SLS Specification

	(1)	(2)	(3)	(4)	(5)
	Effect on Log Positive Cites per Judge				
Judge Age	-0.0916**	-0.139**	-0.0766**	-0.0893**	-0.0888**
	(0.0278)	(0.0395)	(0.0232)	(0.0249)	(0.0244)
Court FE, Year FE	X	X	X	X	X
Court Trends/Windows		X	X	X	X
Init Court Rules × Year FE			X	X	X
Init Case Types × Year FE				X	X
Init Age × Year FE					X
Cragg-Donald F-stat	44.526	39.694	43.803	46.821	45.943
Kleibergen-Paap F-stat	10.372	7.416	17.520	16.991	21.625
N	15010	15010	15010	15010	15010
R-sq	0.460	0.526	0.538	0.555	0.565

2SLS estimates for effect of age on performance, instrumenting with the retirement reforms. Observation is a judge working in a year. "Retirement Reform" is an indicator for the ten years after the introduction of mandatory retirement. Court Treat Windows means court-specific treatment windows (ten years before and after reform). "Init X" × year FE means initial values are interacted with year. "Init Court Rules" includes a state's 1947 rules for judge selection/retention system, admin office, intermediate appellate court, number of judges, and term length. "Init Case Types" includes a court's 1947 average values for case characteristics (legal area and related industries). "Init Age" includes the initial mean and standard deviation for judge age on the court. Standard errors clustered by state in parentheses. + p<.0.1, * p<0.05, ** p<0.01. Back

Age-Balanced Sample Distribution



Histogram of age groups for balanced sample analysis. Number of judge-years in each overlapping ten-year balanced sample. Back