

# THE SIX-MONTH LIST AND THE UNINTENDED CONSEQUENCES OF JUDICIAL ACCOUNTABILITY

*Miguel F. P. de Figueiredo, Alexandra D. Lahav  
& Peter Siegelman*<sup>†</sup>

*A little-known mechanism instituted to improve judicial accountability and speed up the work of the federal judiciary has led to unintended consequences, many of them unfortunate. Federal district court judges are subject to a soft deadline known as the Six-Month List (the List). By law, every judge's backlog (cases older than three years and motions pending more than six months) is made public twice a year. Because judges have life tenure and fixed salaries, a mere reporting requirement should not influence their behavior. But it does. Using the complete record of all federal civil cases between 1980 and 2017 and a hand-coded sample of sum-*

---

<sup>†</sup> The authors are affiliated with the University of Connecticut Law School, where de Figueiredo is Associate Professor, Lahav is Ellen Ash Peters Professor, and Siegelman is Phillip I. Blumberg Professor. Siegelman is the corresponding author: peter.siegelman@law.uconn.edu. We thank Hon. Nancy Gertner for the suggestion that motivated this Article. We are extraordinarily grateful to Prof. William Hubbard of the University of Chicago Law School for sharing with us his cleaned and carefully documented civil litigation data, as well as for his insightful suggestions. We are also very grateful to the many anonymous informants who helped us understand the mechanics and consequences of the Six-Month List. Anne Rajotte at the UConn Law School library made important contributions to this work. We received helpful comments from Oren Bar-Gill, Eric Beal, Deborah Beim, John de Figueiredo, Tim Fisher, Brian Galle, Sean Griffith, Michael Higgins, William H.J. Hubbard, Benjamin Johnson, Tom Miceli, Sachin Pandya, Jonathan Petkun, Hon. Loretta A. Preska, Connor Raso, Hon. Lee Rosenthal, Kathy Seger-son, David Super, Josh Teitelbaum, Steve Thel, David Vladeck, Jonathan Vogel, Tobias Wolff, Megan Wright, Hon. William Young, and Adam Zimmerman. Not all of them agree with all our conclusions, for which we alone are responsible; none of those listed were interviewed for this study. Seminar and conference participants (including students) at UVA, Fordham, UConn Law School, UConn Economics, Georgetown, Penn State, Stanford, the Political Economy and Public Law (PEPL) Conference, the Conference on Empirical Legal Studies (CELS), and the American Law & Economics Association (ALEA) Annual Meeting all made helpful suggestions. Finally, thanks for outstanding research assistance to a group of motions coders: Katherine Blouin, Kimberly Bosse, Tarek Chatila, Ramy Esmail, Maegan Faitsch, Aniella Fignon, Adam Fountaine, Zachary Gain, Katherine Graichen, Patrick Greenhalgh, Jonathan Hague, Michael Karpman, Christopher Kelley, Jesse King, Michael Kochol, Luke Martin, Ashley McWilliams, Tiffany Montouth, Bradley Morley, Stephani Roberts, Dylan Shaw, Patricia Shields, Stacy Skankey, and Julia Steere. Special thanks to super-coder extraordinaire, Amanda Carpenter. Matthew Hall provided terrific programming help.

mary judgment motions, we demonstrate that the List leads judges to close substantially more cases and decide more motions in the week immediately before it is compiled. While average motion processing time is shortened by ten to thirty days, duration is actually lengthened for some motions (those for which the deadline is least pressing). Moreover, we find suggestive evidence that the List has substantive consequences: in an effort to comply with the List, judges may be making more errors. Theory suggests that giving judges an incentive for faster case processing is probably a mistake. But because this incentive is Congressionally-mandated, it cannot be eliminated by the judiciary unilaterally. We offer an alternative mechanism that will limit distortions until Congress acts to relieve the federal courts of this unnecessary burden.

INTRODUCTION .....	365
I. BACKGROUND .....	372
A. The List's History and Motivation .....	372
B. How the List Works .....	375
C. Evaluations of the CJRA and the List .....	377
D. Literature Review .....	379
II. DATA AND METHODS .....	383
A. Data Sources: Case and Motion Level .....	383
1. <i>Case-Level Data</i> .....	383
2. <i>Motions Data</i> .....	384
B. Methods .....	385
1. <i>Difference-in-Differences</i> .....	385
2. <i>Quasi-Random Assignment</i> .....	386
a. <i>Assessing Randomness in Filing Dates</i> .	388
b. <i>Further Measurement Issues</i> .....	390
III. THE EFFECTS OF THE LIST .....	390
A. Bunching .....	392
1. <i>Case-Level Bunching</i> .....	392
a. <i>Spikes in Volume of Case Closures</i> .....	392
b. <i>No Bunching of Case</i> .....	397
2. <i>Motion-Level Bunching</i> .....	397
a. <i>Bunching in Motion Dispositions</i> .....	398
b. <i>Controlling for Age?</i> .....	401
B. Duration: Does the List Speed-Up Case and Motion Processing? .....	402
1. <i>Case-level Effects</i> .....	403
a. <i>"Durational Bunching"</i> .....	403
b. <i>Using Filing Dates to Test Duration Reduction</i> .....	405
2. <i>Motions</i> .....	409
a. <i>Duration-Reducing Effects of the List</i> ...	409

- b. *Which Motions are Processed More Rapidly?* ..... 411
- c. *Does Reducing Motion Processing Time Speed up the Case Overall?* ..... 414
- d. *Generalizing the Results* ..... 415
- C. *Outcomes: Does the List Change Decisions?* .. 417
  - 1. *Case-Level Evidence: The Plaintiff Win Rate* ..... 418
  - 2. *Case-Level Evidence: The Remand Rate* .... 421
  - 3. *Motion Evidence: Grant/Denial Rates and Other Dispositions*..... 425
    - a. *Overall Results* ..... 426
    - b. *Interaction of Treatment and Party Filing Motion* ..... 431
  - 4. *Discussion* ..... 432
    - a. *List Week Drops in Win Rates at the Case Level* ..... 433
    - b. *List Week Spikes in Motion Denials* ..... 434
    - c. *Treatment Effects on Motion Grants*..... 435
- IV. THE SIX-MONTH LIST AS PUBLIC POLICY: DESIGNING GOOD INCENTIVES ..... 437
  - A. *Cost Benefit Analysis* ..... 438
    - 1. *Costs* ..... 439
    - 2. *Benefits* ..... 441
  - B. *Complying with the CJRA* ..... 442
    - 1. *Problems with the Current Incentives*..... 443
      - a. *Incentive Mismatch & Excess Heterogeneity* ..... 444
      - b. *Excessive Duration Risk Borne by Judges* ..... 445
    - 2. *Better Incentives (if any)*..... 446
- CONCLUSION ..... 448

INTRODUCTION

Who can be against accountability?

We are, at least in the case of judicial incentives.<sup>1</sup> This position may seem shocking, but we are in good company, including Nobel prize winning economist Bengt Holmstrom.

---

<sup>1</sup> We use “incentive” in the tradition of the personnel economics literature to mean an explicit mechanism that is designed by a principal to punish or reward an agent for certain conduct. Judges may face other *implicit* incentives (e.g., to bolster their reputation so as to increase their odds of being selected for a Circuit Court judgeship), but the Six-Month List is the only express incentive for federal judges, as far as we know.

His theoretical work demonstrates that in settings where there are multiple objectives—some of which are unmeasurable—optimal policy may require no incentive at all, even if there are some job components where measurement and incentives are possible.<sup>2</sup> When only some aspects of performance can be measured and rewarded, the best choice may be “to pay a fixed wage independent of measured performance.”<sup>3</sup> In such settings, performance-based pay may cause the worker to devote too much effort to the incentivized task at the expense of what cannot be rewarded. In this Article, we show how a simple shaming mechanism instituted to improve judicial accountability and thereby speed up the work of the federal judiciary has had unintended, and mostly bad, consequences, just as Holmstrom’s work predicts. For judges, no incentive at all is likely to be the optimal approach because the system cannot measure the quality of judicial decisions.

In 1990, Congress turned its eye on the federal courts because of an alleged “crisis”: judges were taking too long to handle routine matters that could and should be resolved more rapidly.<sup>4</sup> The complaints seemed pretty bad. One story involved a fifty-six-year old man who suffered a personal injury, but whose trial was delayed for four years and rescheduled three times, causing his litigation costs to skyrocket.<sup>5</sup> But these anecdotes were just that: anecdotes.<sup>6</sup> As the Chief Judge

---

<sup>2</sup> See generally Bengt Holmstrom & Paul Milgrom, *Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design*, 7 J. L. ECON. & ORG. 24 (1991) (providing model of optimal incentives when some conduct is unobservable and cannot be rewarded or punished); Bengt Holmstrom, *Pay for Performance and Beyond*, 107 AMER. ECON. REV. 1753, 1768-69 (2017) (Nobel lecture) (concluding that in settings where quality is not measurable and hence not rewardable, optimal policy may require no incentive at all). Holmstrom and Milgrom’s basic insight has been further developed by an extensive literature in personnel economics on the design of optimal incentives. See sources *infra* note 63.

<sup>3</sup> Holmstrom & Milgrom, *supra* note 2, at 26.

<sup>4</sup> See Stephen Labaton, *Business and the Law; Biden’s Challenge to Federal Courts*, N.Y. TIMES (Apr. 16, 1990), <https://www.nytimes.com/1990/04/16/business/business-and-the-law-biden-s-challenge-to-federal-courts.html> [<https://perma.cc/8THH-S6ML>].

<sup>5</sup> Joseph R. Biden, Jr., *Equal, Accessible, Affordable Justice Under Law: The Civil Justice Reform Act of 1990*, 1 CORNELL J. L. & PUB. POL. 1, 2 (1992). Biden was the CJRA’s sponsor.

<sup>6</sup> In fact, according to our calculations based on AO dataset described *infra*, among cases filed in 1989, the average case terminated in 370 days, two-thirds of all cases closed in less than one year, and only 15% of cases took more than two years to close. To be sure, there were—and still are—some judges with substantial backlogs of unresolved cases and motions. But such judges are very few in number and their backlogs cannot possibly be a significant source of average (system-wide) delays.

of the Second Circuit, James L. Oakes, told a journalist at the time: “They are trying to take examples of judges who have been bogged down and extrapolate it to apply to the whole judiciary . . . . To the extent that it does lay blame on the judges, it’s a bad rap.”<sup>7</sup>

By the standards of modern social science, none of these stories offered anything resembling compelling evidence that slow judges were an important cause of delays,<sup>8</sup> or that imposing an incentive on all judges to move civil cases more quickly would be sound policy.<sup>9</sup> Still, Congress concluded the delays were a problem and that a significant cause was the federal judiciary’s lack of accountability.<sup>10</sup> The result was the Civil Justice Reform Act (CJRA),<sup>11</sup> a law that tried to force dilatory judges to get in line by publicizing judicial delays. That is, the law mandated that each judge have his or her backlog of pending cases and motions publicly revealed twice a year; the publication is therefore universally known as the “Six-Month List.”

This aspect of the CJRA is a perfect example of anecdotal policymaking—legislators relied on virtually no quantitative or statistical evidence in assessing the size or the causes of the problem they set out to address. Neither did they consider what other consequences might result from the statute. But just as Holmstrom’s theory would predict, the law has led judges to distort their behavior, and to focus inordinately on the semi-annual deadlines Congress created under the CJRA.<sup>12</sup> As Holmstrom explained, one risk of instituting incen-

---

<sup>7</sup> Labaton, *supra* note 4.

<sup>8</sup> We do not have data on the distribution of judicial backlogs in 1990; but after consulting the September 30, 2016 CJRA report, we calculated that 57% of the 1,085 active federal district court judges had *no* motions pending more than six months, and 77% had fewer than five such motions pending. See *CJRA Table 8—Reports of Motions Pending Over Six Months For Period Ending September 30, 2016*, U.S. COURTS, [https://www.uscourts.gov/sites/default/files/data\\_tables/cjra\\_8\\_0930.2016.pdf](https://www.uscourts.gov/sites/default/files/data_tables/cjra_8_0930.2016.pdf) [<https://perma.cc/XJ6D-LMGH>]. We note that these are statistics from the post-CJRA era, so the distribution might have been different before the statute was passed.

<sup>9</sup> Rather, faster processing of motions or cases inherently involves trade-offs: other matters receive less attention, or work gets done less carefully on a rushed schedule, or resources are strained.

<sup>10</sup> See Labaton, *supra* note 4 (“The real problem here is that Federal judges have lifetime tenure . . . . That would make it difficult to make judges accountable and force them to follow the Biden Act.”).

<sup>11</sup> The CJRA had many components; our focus is only on its creation of the Six-Month List. 70 Pub. L. No. 101–650, §§ 101–106, 104 Stat. 5089 (1990) (codified as amended at 28 U.S.C. §§ 471–482). The Six-Month List provision is codified at 28 U.S.C.A. § 476.

<sup>12</sup> A list of motions pending more than six months before “judicial officers,” and cases pending more than three years before the same, is compiled twice a

tives such as these is that by rewarding compliance with the deadline, they tend to direct judges' attention away from other important outputs that the system does not reward, such as consistency across cases, fairness among litigants, and perhaps even decision quality.<sup>13</sup> This insight won Holmstrom the 2016 Nobel prize, and it turns out to be just as applicable to judges as to other "workers."<sup>14</sup> In fact, we show that although they are insulated from review and have lifetime tenure, judges—just like students, military recruiters, and many other ordinary "employees"<sup>15</sup>—work to deadline and sometimes even delay projects until they are due. And as they speed up their work to meet deadline pressures, the quality of that work often suffers.

As far as we know, the Six-Month List is the only current judicial incentive; but as the discussion below shows, it plays a significant role in the operation of the federal courts. Federal judges are unique, in that the Constitution protects them from the common incentives that we associate with ordinary employment, such as firing or salary reduction.<sup>16</sup> They are not eligible for merit raises, and are not subjected to merit review; judicial pay is not even dependent on a judges' length of service. Rather, all district court judges earn the same salary,

---

year, on March 31 and September 30. Cases and motions are actually listed separately, as are bench trials pending more than six months, bankruptcy appeals pending more than six months, and Social Security appeal cases pending more than six months. The List is available online, at *Civil Justice Reform Act Report*, U.S. COURTS, <https://www.uscourts.gov/statistics-reports/analysis-reports/civil-justice-reform-act-report> [<https://perma.cc/C5CK-E5Q4>] (last visited Nov. 11, 2016). Judges are allowed to append one of several "explanation codes" (or "Status Codes"), such as "Recently Received from Another Judge," or "Decision/Opinion in Draft." See *Civil Justice Reform Act Status Codes*, FED. JUD. CTR., <https://www.fjc.gov/sites/default/files/2017/CJRA-I-2%20CJRA%20Status%20Codes.pdf> [<https://perma.cc/WY7K-JBZT>]; U.S. COURTS, CIVIL JUSTICE REFORM ACT OF 1990, at 3–4, [https://www.uscourts.gov/sites/default/files/cjra\\_na\\_0930.2018\\_1.pdf](https://www.uscourts.gov/sites/default/files/cjra_na_0930.2018_1.pdf) [<https://perma.cc/TDE2-FS7X>]. Appendix Figure A provides a sample page from a recent Six-Month List.

<sup>13</sup> See Holmstrom & Milgrom, *supra* note 2, at 26 ("[A]n increase in an agent's compensation in any one task will cause some reallocation of attention away from other tasks.").

<sup>14</sup> *Id.* (e.g., teachers). Ironically, Holmstrom published his article explaining this insight in 1991, just as the CJRA was taking effect.

<sup>15</sup> See generally Dan Ariely & Klaus Wertenbroch, *Procrastination, Deadlines, and Performance: Self-Control by Precommitment*, 13 PSYCHOL. SCI. 219 (2002) (providing theoretical model of how imperfectly rational actors can use precommitment devices to solve procrastination problems).

<sup>16</sup> Judges "hold their Offices during good Behaviour, and . . . receive for their Services a Compensation, which shall not be diminished during their Continuance in Office." U.S. CONST. art. III, § 1.

regardless of tenure.<sup>17</sup> Judges who fail to live up to their professional obligations are not punished except in extreme circumstances; likewise, those who work diligently are not materially rewarded.<sup>18</sup> Instead of incentives, we expect judges to be motivated by a commitment to justice, expressed in their professional values and dedication to duty and craft.<sup>19</sup>

Our empirical study confirms, however, that the List has significant effects on judicial behavior. In the words of one informant: “[w]e live and die by the List around here.”<sup>20</sup> For example, Figure 1 reveals the dramatic spikes in case closures during List weeks, the thirteenth and thirty-ninth weeks of the year, immediately preceding the compilation of the March 31 and September 30 Lists.<sup>21</sup>

---

<sup>17</sup> See *Judicial Compensation*, U.S. COURTS, <http://www.uscourts.gov/judges-judgeships/judicial-compensation> [<https://perma.cc/VP29-7ZWZ>] (last visited Oct. 31, 2016) (disclosing “how much federal judges are paid currently since 1968”).

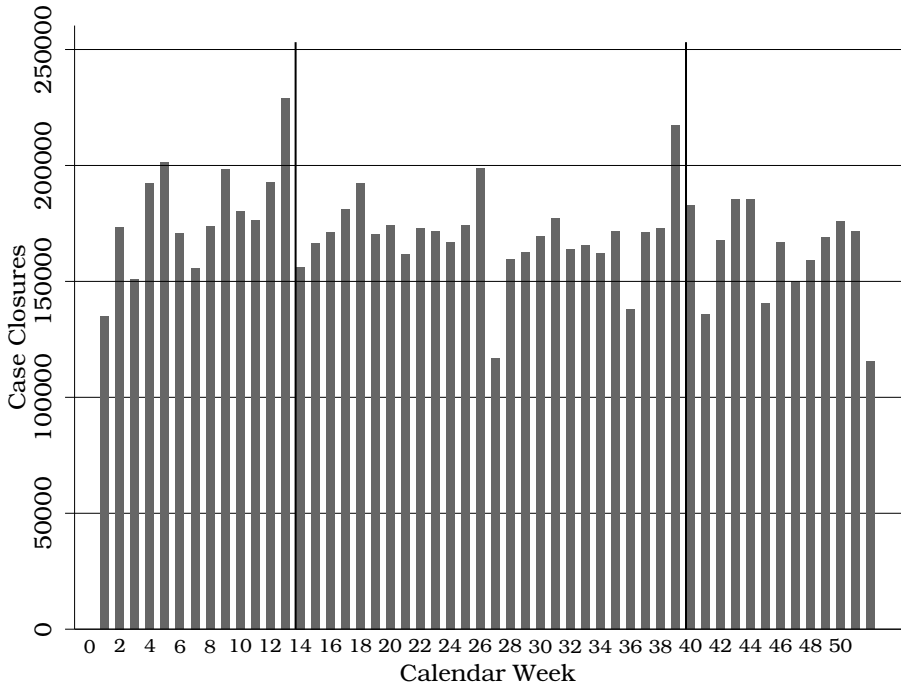
<sup>18</sup> As former Seventh Circuit Judge Richard A. Posner put it, “[A]lmost the whole thrust of the rules governing compensation and other terms and conditions of judicial employment is to *divorce* judicial action from incentives—to take away the carrots and sticks, the different benefits and costs associated with different behaviors, that determine human action . . . .” Richard A. Posner, *What do Judges and Justices Maximize? (The Same Thing Everyone Else Does)*, 3 SUP. CT. ECON. REV. 1, 2 (1993) (emphasis added). Posner’s focus is on appellate judges, but his point seems equally applicable to district court judges. *Id.* at 4–5.

<sup>19</sup> See *id.* at 40.

<sup>20</sup> Interview with operations manager at a U.S. District Court, June 2016 (on file with author).

<sup>21</sup> There is also an uptick in decisions on motions during List weeks.

FIGURE 1: VOLUME OF CASE CLOSURES, BY CALENDAR WEEK:  
1980–2017



These findings might be expected by the student of human behavior who knows that many people delay their work until a deadline looms. But it strikes us as somewhat surprising that Article III judges, with remarkably little at stake, should be so susceptible to peer pressure.

Using an innovative study design, we do find some evidence that the List has reduced motion processing time, albeit in uneven and unanticipated ways. “Eligible” motions<sup>22</sup> filed just before the cutoff for the next List are subject to a six-month deadline, while those filed just after the cutoff have more than a year before they are eligible to appear on the List.<sup>23</sup> As a result, motions in the first group are, on average, processed ten to thirty days faster than those in the second.<sup>24</sup>

<sup>22</sup> By “eligible,” we mean motions pending before a U.S. district or magistrate judge that will appear on the List if the motion remains pending for more than six months.

<sup>23</sup> For instance, if a motion is filed in a United States District Court on September 1, 2017, the motion would not yet be pending for a full seven months by March 31, 2018 (when the next Six-Month List would be calculated). Thus, it would not be eligible to appear on that List; instead, the judge would have an extra six months (until September 30, 2018) to resolve the motion before it could appear on the List.

<sup>24</sup> See Figures 4A, 4B, 4C, *infra* section III.A.2.



Importantly, however, the reduction in processing time is only experienced by motions that would have taken roughly six months to resolve in the absence of a deadline. Motions that have very short or very long durations are unaffected by the List. We also find that judges apparently “mothball” some motions they were not able to resolve by the March deadline, pushing off a decision until the September deadline nears. In sum, even though it may reduce the average motion processing time, the six-month deadline does so in a way that introduces considerable heterogeneity in the administration of justice.

Even more surprising is the effects of the List on *outcomes*, which more directly implicate justice concerns. We find suggestive evidence that the List negatively affects accuracy. Some of our evidence is causal. For example, exposure to the List results in an increase in defendant wins (a combination of grants of defendant-filed summary judgment motions and of motions to dismiss) and an increase in settlements in cases where plaintiffs filed summary judgment motions.<sup>25</sup> Other evidence is more correlational. For example, the plaintiff win rate at the case level declines substantially in List weeks, and cases decided in List weeks have a higher appellate remand rate than those decided at other times.<sup>26</sup> The risk that the List may affect outcomes by increasing error arises from a possible tradeoff between speed and accuracy, a problem that was never addressed in the initial decision to implement this incentive.

In Part I, we describe the legislative history and theoretical background. In Part II, we describe our study design. Part III explains our findings and their significance. Part IV argues that it may be better to eliminate incentives altogether in situations such as this one, where the CJRA’s objective (speed) is easy to measure, but other valuable objectives (accuracy, impartiality) are unmeasurable. Recognizing that the Six-Month List is legally required, we suggest a stopgap measure that the federal courts can adopt until such time as it can be reconsidered. The main lessons of our analysis are that legislators ought to be wary of anecdotal policymaking and that in the

---

<sup>25</sup> See Figure 14, *infra* section III.A.2 (relying on AO/FJC data regarding, among other things, civil cases filed and terminated from 1970 to 1987, as well cases filed and terminated from 1988 to present). See *Integrated Database (IDB)*, FED. JUD. CTR., <https://www.fjc.gov/research/idb> [<https://perma.cc/6XSJ-XRCS>].

<sup>26</sup> See Table 6, *infra* subsection III.C.3(b) (relying on AO / FJC data regarding, among other things, civil cases filed and terminated from 1970 to 1987, as well cases filed and terminated from 1988 to present). See *Integrated Database (IDB)*, *supra* note 25).

case of judges in particular, no incentive may well be better than an incomplete incentive. As Judge Posner has explained, there are good reasons “to divorce judicial action from incentives.”<sup>27</sup>

## I BACKGROUND

In this part, we describe the origins of the Six-Month List and how it works. We also review the extant literature evaluating the List, incentives in the workplace, and judicial behavior. Readers familiar with this history or who are interested only in the empirical findings may wish to skip to Parts II and III.

### A. The List’s History and Motivation

The Civil Justice Reform Act of 1990 (CJRA) made the publication of the Six-Month List mandatory, but the idea of publicizing judicial backlogs was not wholly new at the time. Prior to 1990, the Judicial Conference of the United States required internal reporting on a quarterly basis of cases older than three years and matters under advisement for more than sixty days.<sup>28</sup> There was also a patchwork of formal and informal mechanisms for coping with delays before the advent of the CJRA.<sup>29</sup> As we will see when we examine the data, there were quarterly spikes in case dispositions prior to 1991, which we attribute to these quasi-formal practices.<sup>30</sup> What the CJRA did was to bring more publicity, uniformity of reporting requirements across circuits, and a biannual rather than quarterly reporting scheme.<sup>31</sup>

The stated purpose of the CJRA was to speed-up case disposition time in the federal courts and to reduce delays,

<sup>27</sup> Posner, *supra* note 18.

<sup>28</sup> See R. Lawrence Dessem, *Judicial Reporting Under the Civil Justice Reform Act: Look, Mom, No Cases!*, 54 U. PITT. L. REV. 687, 701 n.81 (1993).

<sup>29</sup> Charles Geyh noted the formal and informal methods for dealing with “indefensible” delays before the CJRA, largely dismissing the former as rarely used and essentially ineffective. Charles Gardner Geyh, *Adverse Publicity as a Means of Reducing Judicial Decision-Making Delay: Periodic Disclosure of Pending Motions, Bench Trials and Cases Under the Civil Justice Reform Act*, 41 CLEV. ST. L. REV. 511, 513–20 (1993). Among the latter were peer pressure by a judge’s colleagues and (private) communication from the Chief Circuit judge or the Chief District judge, as well as publicity and public accountability in extreme cases. *Id.* at 524–28. Geyh suggested that the Six-Month List provision of the CJRA worked to complement these informal mechanisms. *Id.* at 536.

<sup>30</sup> See Figure 3, *infra* subsection III.A.1(a).

<sup>31</sup> See Dessem, *supra* note 28, at 695 (providing helpful discussions of CJRA implementation and early public reactions).

abuses, and discovery costs.<sup>32</sup> In particular, the Six-Month List was explicitly designed to create a mild shaming sanction for dilatory judicial behavior.<sup>33</sup> As the Brookings Institution Task Force Report, on which the legislation was based, explained: “We believe that substantially expanding the availability of public information about caseloads by judge will encourage judges with significant backlogs in undecided motions and cases to resolve those matters and to move their cases along more quickly.”<sup>34</sup>

The idea emerged from a judicial survey which found that most respondents believed that it would assist judicial accountability if information regarding judicial caseloads, including delays, were published.<sup>35</sup> Of the 147 federal judges responding to that survey, 61% were reported to favor the idea,<sup>36</sup> even though only 4% of respondents believed that judicial delay was one of the most serious criticisms of the federal courts.<sup>37</sup>

Hearings on the bill confirmed that the Six-Month provision was designed to harness peer pressure. For example, Patrick Head, speaking for the business community, explained:

This bill does not have any so-called ‘teeth.’ It relies very heavily on peer group pressure and on the responsiveness of highly skilled professionals appointed for life. If a judge misses deadlines, or even consistently misses them, there is

---

<sup>32</sup> See Arthur R. Miller, *The Pretrial Rush to Judgment: Are the “Litigation Explosion,” “Liability Crisis,” and Efficiency Clichés Eroding Our Day in Court and Jury Trial Commitments?*, 78 N.Y.U. L. REV. 982, 998 (2003); Linda S. Mullenix, *Civil Justice Reform Comes to the Southern District of Texas: Creating and Implementing A Cost and Delay Reduction Plan Under the Civil Justice Reform Act of 1990*, 11 REV. LITIG. 165, 172–73 (1992).

<sup>33</sup> Biden, *supra* note 5, at 8–9.

<sup>34</sup> THE BROOKINGS INST., JUSTICE FOR ALL: REDUCING COSTS AND DELAY IN CIVIL LITIGATION 27 (1989); see also Dessem, *supra* note 28, at 690 (quoting the same).

<sup>35</sup> Dessem, *supra* note 28, at 690 (“This reporting requirement, and the entire Task Force report, was based in large part upon a survey conducted by Louis Harris and Associates, Inc. Those surveyed were asked whether, in an effort to increase judicial accountability, courts should make ‘publicly available each year the average length of cases, weighted by type of case, under each Federal judge.” (footnote omitted)).

<sup>36</sup> *Id.*

<sup>37</sup> *Id.*; see also BROOKINGS, *supra* note 34, at 25 (“The task force believes that its recommendations for increased judicial case management articulate an approach to the twin problems of cost and delay that maintains the essential requirements of due process. It is also noteworthy that the substantial majority of those who participate in the civil justice system, evidenced by the responses to the Harris survey . . . overwhelmingly support active judicial management.”).

no retribution spelled out in the legislation, nor should there be.<sup>38</sup>

Similarly, the response of Representative John Bryant to testimony from Public Citizen Litigation Group about one especially egregious delay gives a flavor of the intent behind the List: "I think the fellow ought to be sanctioned, the name of this man or woman ought to be put on the billboard outside the building here stating that they are not working like the rest of us."<sup>39</sup> There were very few dissenting voices to the final version of the bill.<sup>40</sup> Only the Seventh Circuit Bar Association objected that the lists "would serve primarily to focus judicial attention unduly upon the two statistical deadlines which would be reflected in the reports."<sup>41</sup> As it turns out, the Seventh Circuit Bar Association was right, but nobody who testified foresaw the deeper and more disturbing effects of the List.

In sum, the List was from its inception widely understood as a modest shaming mechanism meant to put pressure on judges to manage their dockets more efficiently by identifying judicial laggards.<sup>42</sup>

---

<sup>38</sup> *The Civil Justice Reform Act of 1990 and the Judicial Improvements Act of 1990, Hearings on S. 2027 and S. 2648 Before the Comm. on the Judiciary*, 101st Cong., 2d Sess., 46 (1990) [hereinafter *Senate Hearings*] (answers of Patrick Head to questions from Senator Hatch).

<sup>39</sup> Dessem, *supra* note 28, at 692–93 (quoting *Federal Courts Study Comm. Implementation Act and Civil Justice Reform Act: Hearing on H.R. 5381 and H.R. 3898 Before the Subcomm. on Courts, Intellectual Property, and the Administration of Justice of the House Comm. on the Judiciary*, 101st Cong., 2d Sess. 215, 240 (1990) (statement of Rep. Bryant)).

<sup>40</sup> *Id.* at 693 (noting that the Seventh Circuit Bar Association provided the "only outright opposition to the proposed reporting provision"). The bill originally proposed a much more onerous reporting requirement with a great deal more data. This was opposed by the Judicial Conference. In his Senate testimony, Judge Robert Peckham expressed concerns that the reporting requirement would measure the wrong things or would be misunderstood by the public. *Id.* at 691 (quoting *Senate Hearings*, *supra* note 38, at 476).

<sup>41</sup> *Id.* at 694 (quoting *Senate Hearings*, *supra* note 38, at 512, 515 (letter from Harvey M. Silets, President of the Seventh Circuit Bar Association, to Senator Joseph Biden (June 25, 1990))).

<sup>42</sup> See Biden, *supra* note 5, at 8–9. Writing shortly after the CJRA was passed, its primary author, then-Senator Joseph Biden cited with approval a letter he had recently received from a federal judge who pointed out that in his district, the List

has caused each individual judge to sharpen his focus on case management and on the timeliness of his decision making. The case termination statistic in our district has shown a substantial increase. We all recognize that peer pressure plays an important role in our everyday lives, and it likewise is important in the judicial setting.

*Id.* (quoting letter from Chief Judge Justin L. Quackenbush (E.D. Wash.) to Senator Biden, dated April 23, 1991).

## B. How the List Works

On March 31 and September 30 of every year, in every United States District Court, case management software run by the clerk of the court calculates how many open motions are listed on each judge's docket that have been pending for more than "six months."<sup>43</sup> In practice, that six months is actually 214 days, because each motion is treated as having a grace period of thirty days from filing.<sup>44</sup> Thus, a motion filed on March 1 is treated as though it was filed on March 31 for the purpose of compiling the List.<sup>45</sup> Cases pending for more than three years as of the List-compilation date are also tabulated, but without any grace period.<sup>46</sup>

Judges are acutely aware of the List and its deadlines. As one judge put it, "[n]o judge likes being on this 'report of shame.'"<sup>47</sup> Court personnel also spend time on administration related to the List.<sup>48</sup> Clerks we have spoken to report that prior

<sup>43</sup> *Civil Justice Reform Act Report*, *supra* note 12.

<sup>44</sup> See Dessem, *supra* note 28, at 695 n.48.

<sup>45</sup> The justification for this thirty-day grace period is that a judge cannot begin to adjudicate a motion until the judge has heard from the nonmoving party. The judge's clock is thus only deemed to start running when he or she is in a position to act on the motion.

<sup>46</sup> Interview with anonymous U.S. District Court operations staff (on file with author). See also III ADMIN. OFFICE OF THE U.S. COURTS, GUIDE TO JUDICIARY POLICIES AND PROCEDURES, ch. XXII, Pt. C, at 21 (1991) (Appendix B).

<sup>47</sup> Then-Senior Judge Richard Kopf (D. Neb.) pointed out on his short-lived blog that

there are computer systems in place to keep federal judges advised of how long their motions have been under advisement. The CM/ECF system . . . has been adopted by all federal courts. That system is able to produce computer runs of "motion lists" upon demand showing precisely what motions are pending and for how long. In my chambers, that list is run weekly and distributed to each of my law clerks.

Richard G. Kopf, *What to do When Your Summary Judgment Motion Goes Missing in Federal Court*, HERCULES & THE UMPIRE (Sept. 13, 2013), <https://herculesandtheumpire.com/2013/09/13/what-to-do-when-your-summary-judgment-motion-goes-missing-in-federal-court/> [https://perma.cc/YJ36-U6XS]. He added that "many courts have adopted internal Guidelines for resolving motions on a timely basis." *Id.* And, referring to the Six-Month List, he suggested that "[n]o judge likes being on this 'report of shame' and you can bet internally the judge and his or her staff are trying to resolve . . . [summary judgment motions] before being required to report [them] . . ." *Id.*

<sup>48</sup> We informally and anonymously interviewed two court clerks, four district court judges, and five lawyers in both defense and plaintiff side practice as part of our research (interviews on file with author). Importantly, however, judges do not appear to be aware of the technical aspects of how the List is calculated. According to our sources, most judges do not know about the 214-day rule, and instead think that List eligibility is based on the statutory definition under which motions are eligible if filed more than six months before the List compilation date (March 31 or September 30). In other words, judges do not seem to appreciate that they

to the deadline, they will send reminders to judges including lists of motions that, if not decided, will appear on the next List. They also report that district court judges will have their chambers produce similar reports or keep tabs on motions that may end up being on the next List. After the official report is assembled by the chief clerk of the individual District Court, each judge gets the compilation of her pending cases/motions so that she can add a “reasons code” (or “Status Code”) to the List, giving her a chance to offer an explanation for the delay.<sup>49</sup> Then the List is reproduced with the reasons code and sent to the Administrative Office of the Courts (AO) for distribution. The List is placed online in PDF form and is accessible to the general public.<sup>50</sup>

Although judges by all reports are highly attuned to it, the List imposes a very low risk of public shaming. Judicial backlogs or the List are rarely mentioned in the media; we found only sixty-five stories between 1990 and 2017.<sup>51</sup> Conservatively, if we multiply this number by a factor of ten to account for stories we might have missed, given a population of federal district court judges of roughly 700 and a twenty-seven-year span of time, that works out to just under one story per thirty judge-years. This strongly suggests that the List operates via peer pressure rather than through fear of bad publicity.

---

actually have seven months, rather than six, to close a motion. Moreover, while most of the lawyers we spoke to had heard of the Six-Month List, many believed that it worked on a rotating basis (every motion had to be resolved within six months of its filing), rather than being compiled semiannually. This strongly suggests that the List is not on anyone’s mind *at the moment a case or motion is filed*, but becomes relevant as the deadline approaches. At this point, the filing date is no longer of any real interest.

<sup>49</sup> See *Civil Justice Reform Act Status Codes*, *supra* note 12.

<sup>50</sup> *Id.* Some clerks have also reported to us that the clerk circulates the List to all the judges in that court, but we do not know whether this is the practice in every courthouse. It may depend on the discretion of Chief Judge of the particular district. Clerks and judges also report that judges in some districts are aware of which of their colleagues have a backlog of cases and may step in and help out their fellow judges to avoid appearing on the List.

<sup>51</sup> We searched Lexis, Westlaw, Proquest, and the web for mentions of Federal judges’ delay or backlog in any print media. We note that some local newspapers are not covered by these databases and disappear from the web after a short period; that could lead to an underestimate of coverage. On the other hand, our sources included the vast majority of legal newspapers, bar journals, and legal magazines that would be most likely to cover judicial backlogs. And most local papers probably would not have the staff or resources to carry out investigative reporting on their own—they would be more likely to reprint stories from bigger outlets (which *are* covered in our databases).

The AO is sensitive to how the List is calculated and, at least with respect to one set of cases, has altered the calculation method to increase fairness to judges. Specifically, our sources report that at least since 2013, in the Districts of Connecticut and New Jersey, *qui tam* cases were treated differently than other cases for purposes of the List.<sup>52</sup> The reason for this was that a *qui tam* case cannot be resolved until the government has decided whether to intervene.<sup>53</sup> Because this can take time, a judge may find the case lasting longer than the three-year deadline for inclusion on the List. The two districts initially solved this problem by administratively closing *qui tam* cases and reopening them once the government had determined whether to intervene. Subsequently, as of September 2016, the AO likewise instructed all districts to begin running the clock on List eligibility only after a *qui tam* case has been unsealed, which happens once the government decides whether to intervene.<sup>54</sup> This change in calculation demonstrates the importance of the List to all those involved in judicial administration, but it does not affect our analysis because *qui tam* cases are so few.

### C. Evaluations of the CJRA and the List

Implementation of the Six-Month List was only a small part of the changes wrought by the CJRA, which led to many innovations in case handling procedures, most notably the introduction of various alternative dispute resolution mechanisms in federal courts.<sup>55</sup> After its passage, the Act itself was the

---

<sup>52</sup> Interview with anonymous source (July 13, 2017) (on file with author).

<sup>53</sup> 31 U.S.C. §§ 3730(b)(1)–(b)(4) (2018); *see also United States v. Health Possibilities, P.S.C.*, 207 F.3d 335, 338–40 (6th Cir. 2000) (“[T]he government retains significant authority to influence the outcome of *qui tam* actions—even when it decides not to intervene” under the FCA.)

<sup>54</sup> Interview with anonymous court operations manager (Feb. 23, 2016) (on file with author).

<sup>55</sup> *See, e.g.,* Lauren Robel, *Fractured Procedure: The Civil Justice Reform Act of 1990*, 46 STAN. L. REV. 1447, 1448 (1994) (“Acting under color of the CJRA, districts have . . . required participation in alternative dispute resolution as a prerequisite to trial . . . .”); Linda S. Mullenix, *Unconstitutional Rulemaking: The Civil Justice Reform Act and Separation of Powers*, 77 MINN. L. REV. 1283, 1296–97 (1993) (“The Civil Justice Reform Act basically requires all ninety-four federal district courts to draft civil justice expense-and-delay reduction plans . . . [such as] mandatory alternative dispute resolution referrals . . . .”); Carl Tobias, *Recalibrating the Civil Justice Reform Act*, 30 HARV. J. ON LEGIS. 115, 116–17 (1993) (“The [CJRA] requires that all ninety-four federal district courts develop a civil justice expense and delay reduction plan . . . . [They] may adopt the eleven principles, guidelines, and techniques (primarily governing case management, discovery, and alternative dispute resolution).”); Carl Tobias, *Civil Justice Reform and the Balkanization of Federal Civil Procedure*, 24 ARIZ. ST. L.J. 1393, 1418–19

subject of a substantial volume of scholarship, virtually all of which discussed these other aspects of civil justice reform. Researchers at the RAND Corporation also carried out extensive and careful experimental investigations of the effects of new procedures implemented by the CJRA, but apparently did not examine the effects of the List.<sup>56</sup>

Two qualitative discussions of the effects of the List were published in the first years after it was adopted,<sup>57</sup> and we are aware of two quantitative assessments since then. A group at the University of Denver produced an extensive report on civil procedure in federal courts, based on motions data collected in six judicial districts.<sup>58</sup> Although they devoted only a few pages to the List, some of their conclusions strongly track our own, especially about case-bunching near List dates.<sup>59</sup> Our work represents a substantial advance on the Denver study in terms of methodological sophistication and the quantity and quality of evidence assembled, in addition to findings regarding the effect of the List on accuracy, which that study did not address.

Jonathan Petkun has recently written an as yet unpublished but very useful quantitative study of the List.<sup>60</sup> Using a dataset with a much larger volume of motions, Petkun was able to deploy different econometric techniques than we did, ad-

---

(1992) (“Indeed, [pursuant to CJRA,] a number of districts have required or suggested that magistrate judges or alternative dispute resolution processes be employed when resolving less complicated litigation.”).

<sup>56</sup> See generally JAMES S. KAKALIK ET AL., RAND INSTITUTE FOR CIVIL JUSTICE, IMPLEMENTATION OF THE CIVIL JUSTICE REFORM ACT IN PILOT AND COMPARISON DISTRICTS (1996) (describing what various districts did to implement the CJRA); JAMES S. KAKALIK ET AL., RAND INSTITUTE FOR CIVIL JUSTICE, AN EVALUATION OF JUDICIAL CASE MANAGEMENT UNDER THE CIVIL JUSTICE REFORM ACT (1996) (analyzing how the case management practices adopted in various districts affected cost, case disposition time, and participant satisfaction, but without mentioning the Six-Month List).

<sup>57</sup> See generally Geyh, *supra* note 29 (explaining how judicial backlogs were handled in the pre-CJRA era); Dessem, *supra* note 28 (analyzing reporting requirements under the CJRA).

<sup>58</sup> See generally INSTITUTE FOR THE ADVANCEMENT OF THE AMERICAN LEGAL SYSTEM, CIVIL CASE PROCESSING IN THE FEDERAL DISTRICT COURTS: A 21ST CENTURY ANALYSIS (2009), [https://www.uscourts.gov/sites/default/files/iaals\\_civil\\_case\\_processing\\_in\\_the\\_federal\\_district\\_courts\\_0.pdf](https://www.uscourts.gov/sites/default/files/iaals_civil_case_processing_in_the_federal_district_courts_0.pdf) [<https://perma.cc/3JYC-HZKX>]; *id.* at 19 (noting that since the RAND Institute’s study in 1996, “there have been no further studies concerning the entirety of case processing in the federal courts”).

<sup>59</sup> See *id.* at 78. This study found bunching of decisions at List dates. For example, it noted that “nearly 35% of summary judgment motions ruled on during the last two weeks of March or September had been pending for six months or more, meaning that they would have been listed on [an] individual judge’s CJRA report if not resolved before the month-end deadline.” *Id.* at 78.

<sup>60</sup> Jonathan Petkun, Can (and Should) Judges Be Shamed? Evidence from the ‘Six-Month List’ (Oct. 22, 2018) (unpublished manuscript) (on file with author), [https://jbpetskun.github.io/pages/working\\_papers/CELS\\_draft\\_20181022.pdf](https://jbpetskun.github.io/pages/working_papers/CELS_draft_20181022.pdf) [<https://perma.cc/97AB-H55M>].



dressings questions we were unable to examine. Our conclusions are largely parallel. Petkun, too, finds that the List lowers motion processing time by about thirty days, although he also finds some effect on *case* durations, which we do not. He also uncovers considerable heterogeneity in judicial responses to the List (an issue we could not examine), concluding that younger and minority judges are more responsive to the List than older white males. Finally, he concludes that the evidence for the kinds of “quality” effects we find is mixed.

#### D. Literature Review

With the exception of the two sources discussed above, no prior scholarship of which we are aware directly addresses questions of how federal judges respond to incentives in the sense we are using that term.<sup>61</sup> That is hardly surprising, given that federal judges face few—if any—incentives to which they can respond.<sup>62</sup>

A large body of literature in personnel economics focuses generally on monetary incentives linking pay and performance,<sup>63</sup> but this body of scholarship has little direct bearing on

---

<sup>61</sup> Most studies of judicial behavior concern the effect of political ideology on judging. See, e.g., LEE EPSTEIN, WILLIAM M. LANDES & RICHARD A. POSNER, *THE BEHAVIOR OF FEDERAL JUDGES* 44 (2013) (finding that a judge’s prior background predicts his or her decisions (“appointment bias”)); C.K. Rowland, Tina Traficanti & Eric Vernon, *Every Jury Trial is a Bench Trial: Judicial Engineering of Jury Disputes*, in *THE PSYCHOLOGY OF JUDICIAL DECISION MAKING* 197 (David E. Klein & Gregory Mitchell eds., 2010) (same); CASS R. SUNSTEIN, DAVID SCHKADE, LISA M. ELLMAN & ANDRES SAWICKI, *ARE JUDGES POLITICAL?: AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY* 4–16 (2006) (same); Pat K. Chew & Robert E. Kelley, *Myth of the Color-Blind Judge: An Empirical Analysis of Racial Harassment Cases*, 86 WASH. U. L. REV. 1117, 1131–35 (2009) (same); Alma Cohen, Alon Klement & Zvika Neeman, *Judicial Decision Making: A Dynamic Reputation Approach*, 44 J. LEGAL STUD. S133, S156 (2015) (same); Benoit Pierre Freyens & Xiaodong Gong, *Judicial Decision Making Under Changing Legal Standards: The Case of Dismissal Arbitration*, 133 J. ECON. BEHAV. & ORG. 108, 113 (2017) (same); Claire S.H. Lim, *Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges*, 103 AMER. ECON. REV. 1360, 1364 (2013) (same); Claire S.H. Lim, Bernardo S. Silveira & James M. Snyder, Jr., *Do Judges’ Characteristics Matter? Ethnicity, Gender, and Partisanship in Texas State Trial Courts*, 18 AMER. L. & ECON. REV. 302, 304–05, 321–40 (2016) (same).

<sup>62</sup> Apart, perhaps, from bribes, which constitute illegitimate incentives that are not relevant for our analysis. But see Ian Ayres, *The Twin Faces of Judicial Corruption: Extortion and Bribery*, 74 DENV. U. L. REV. 1231, 1242–47 (1997) (documenting extensive bribery among Chicago judges, including payments to fix murder cases). Judge Posner’s article does address judicial “incentives” at the appellate level from a theoretical point of view. Posner, *supra* note 18, at 2–5. He also notes that state judges subject to periodic elections face different incentives. *Id.* at 23–30.

<sup>63</sup> See, e.g., EDWARD P. LAZEAR & MICHAEL GIBBS, *PERSONNEL ECONOMICS IN PRACTICE* (2014) (textbook, explaining basic concepts of personnel economics); Carolyn

our work because none of the standard incentives used in other areas of employment apply to federal judges. Nevertheless, some of what we uncover does have parallels in other literatures. For example, our evidence is suggestive of peer effects in that judges seem to care a great deal about their productivity vis-à-vis that of their colleagues.<sup>64</sup> Such peer influence has been detected in many other settings—using both observational and experimental methods<sup>65</sup>—including Amazon’s Mechanical Turk (an online labor market<sup>66</sup>), nineteenth-cen-

---

J. Heinrich & Gerald Marschke, *Incentives and Their Dynamics in Public Sector Performance Management Systems*, 29 J. POLY ANALYSIS & MGMT. 183, 192 (2010) (arguing that while “[i]ncentive power is typically defined in monetary terms as the ratio of performance-contingent pay to fixed pay, with higher ratios of performance-contingent, monetary pay viewed as ‘higher-powered,’” monetary incentives are less powerful in the public sector, and workers in the public sector tend to be motivated by more intrinsic values, including ethics); Canice Prendergast, *The Provision of Incentives in Firms*, 37 J. ECON. LITERATURE 7, 55 (1999) (survey, noting the “strong effects of pay-for-performance on output, admittedly in settings where measures of overall performance were available”).

<sup>64</sup> For instance, the distinguished judge, Richard A. Posner, has written that “[t]o regard oneself and be regarded by others, especially one’s peers, as a good judge requires conformity to the accepted norms of judging.” RICHARD A. POSNER HOW JUDGES THINK 61 (2008).

<sup>65</sup> See, e.g., Oriana Bandiera, Iwan Barankay & Imran Rasul, *Social Incentives in the Workplace*, 77 REV. ECON. STUD. 417, 418, 442–43 (2010) (finding that in a piece-rate setting, working in physical proximity to an individual’s friends who are more able than the individual raises that individual’s productivity; working with friends who are less able lowers it); Armin Falk & Andrea Ichino, *Clean Evidence on Peer Effects*, 24 J. LAB. ECON. 39, 49–54 (2006) (finding that average output is higher when experimental subjects are randomly assigned to work, in parallel, in a single room than when they work in separate rooms); Emir Kamenica, *Behavioral Economics and the Psychology of Incentives*, 4 ANN. REV. ECON. 427, 443–44 (2012) (observing what peers do can also be informative about “what [one] should do or how hard [one] should work,” and can influence one’s own behavior via a desire to conform); Supreet Kaur, Michael Kremer & Sendhil Mullainathan, *Self-Control and the Development of Work Arrangements*, 100 AM. ECON. REV.: PAPERS & PROC. 624, 626 (2010) (observing that among Indian IT workers, being randomly assigned a high productivity neighbor “increases own productivity by 5 percent”); Alexandre Mas & Enrico Moretti, *Peers at Work*, 99 AM. ECON. REV. 112, 134–35, 143 (2009) (concluding that workers are more productive when they work near—and know they are observed by—others who are more productive than they are).

<sup>66</sup> John J. Horton & Richard J. Zeckhauser, *The Causes of Peer Effects in Production: Evidence from a Series of Field Experiments* 5 (Nat’l Bureau of Econ. Research, Working Paper No. 22386, 2016) (concluding that “[w]hat seems to matter for peer effects in productivity is the observability of output rather than the social interaction”).

tury factory workers,<sup>67</sup> academic journal referees,<sup>68</sup> and governmental food inspectors.<sup>69</sup>

Turning to the influence of deadlines, our finding that case and motion resolutions “bunch” just before List deadlines, turns out to be common in contexts where stronger incentives operate. Instead of spreading out their work over a long period of time, people responding to time-based incentives tend to increase productivity immediately before a deadline. This has been observed in numerous contexts, including road construction projects,<sup>70</sup> military recruitment,<sup>71</sup> commission-based sales,<sup>72</sup> journal refereeing,<sup>73</sup> patent examinations,<sup>74</sup> security background checks,<sup>75</sup> and year-end spending by U.S. Govern-

---

<sup>67</sup> The nineteenth-century utopian theorist and industrialist Robert Owen relied on pure peer effects to motivate workers in his factories, as described by Martin Bloom in *Editorial—Primary Prevention and Education: An Historical Note on Robert Owen*, 23 J. PRIMARY PREVENTION 275, 277–78 (2003) (quoting 1 ROBERT OWEN, THE LIFE OF ROBERT OWEN, WRITTEN BY HIMSELF, WITH SELECTIONS FROM HIS WRITINGS AND CORRESPONDENCE 80–81, 136 (1857)). The reference is due to Steven Tadelis.

<sup>68</sup> Raj Chetty, Emmanuel Saez & László Sándor, Abstract, *What Policies Increase Prosocial Behavior? An Experiment with Referees at the Journal of Public Economics* (Nat'l Bureau of Econ. Research, Working Paper No. w20290, 2014), <http://www.nber.org/papers/w20290> [<https://perma.cc/ER47-35R4>] (finding that that “social incentives” (adverse publicity) did have “small[ ] but significant effects on review times and [were] especially effective among tenured professors, who [were] less sensitive to deadlines and cash incentives”).

<sup>69</sup> Daniel E. Ho, *Does Peer Review Work? An Experiment of Experimentalism*, 69 STAN. L. REV. 1, 61–67 (2017) (finding that random assignment of inspectors to work in teams (as opposed to individually) increased accuracy and reduced under-detection of food safety violations).

<sup>70</sup> See Gregory Lewis & Patrick Bajari, *Moral Hazard, Incentive Contracts, and Risk: Evidence from Procurement*, 81 REV. ECON. STUD. 1201, 1202 (2014) (finding that “[w]hile the distribution of [delay] shocks is continuous, the distribution of outcomes exhibits ‘bunching’ at the project deadline, with many projects being completed exactly on time”).

<sup>71</sup> See Beth J. Asch, *Do Incentives Matter? The Case of Navy Recruiters*, 43 INDUS. & LAB. REL. REV. 89, 104–05 (1990) (showing the numbers of new recruits signed-up by each recruiter tends to rise until his or her assessment point or quota, drops off afterwards, and then begins to rise again until the next assessment).

<sup>72</sup> See Paul Oyer, *Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality*, 113 Q.J. ECON. 149, 173–81 (1998) (finding individual’s sales volumes are higher at the end of fiscal years because an additional dollar of revenue will usually be worth more to the salesperson at the end of one fiscal year than at the start of the next).

<sup>73</sup> See Chetty, *supra* note 68, at 8 (finding that journal referees tend to submit their work just on time).

<sup>74</sup> See Michael D. Frakes & Melissa F. Wasserman, *Procrastination in the Workplace: Evidence from the U.S. Patent Office* 34–37 (Nat'l Bureau of Econ. Research, Working Paper 22987, 2016), <http://www.nber.org/papers/w22987> [<https://perma.cc/B98Y-XSSK>].

<sup>75</sup> The *New York Times* reports that the private contractor responsible for conducting background security checks on applicants for governmental employ-

ment agencies.<sup>76</sup> Time-management experts refer to the tendency to delay completing any task until its deadline approaches as “Student Syndrome” or “Parkinson’s Law,” and although conclusive evidence seems to be lacking, it is believed to be a widespread phenomenon.<sup>77</sup> On the other hand, psychological research also suggests that externally-imposed deadlines can sometimes be useful in overcoming procrastination.<sup>78</sup>

Finally, tradeoffs between speed and accuracy have been identified in other professional contexts such as among doctors<sup>79</sup> and patent examiners.<sup>80</sup> We have elsewhere developed a model of responses to periodic shaming incentives that are based on a judge’s backlog. The model suggests both that the volume of decisions should rise as the backlog publication deadline nears and that decisional quality should fall at such times.<sup>81</sup>

In sum, many studies in other fields demonstrate that incentives frequently have unintended or “distortionary” conse-

ment would routinely “flush” (quickly approve, without quality checks) any background investigations pending at the end of a month in order to meet monthly quotas. Trip Gabriel, *Shortcuts Seen by Firm Doing Security Checks*, N.Y. TIMES (Sept. 27, 2013), <https://www.nytimes.com/2013/09/28/us/pressure-reported-in-rush-to-meet-security-clearances-including-edward-snowden-and-aaron-alexis.html> [<https://perma.cc/JRP6-ZCJ9>].

<sup>76</sup> See Jeffrey B. Liebman & Neale Mahoney, *Do Expiring Budgets Lead to Wasteful Year-End Spending? Evidence from Federal Procurement*, 107 AM. ECON. REV. 3510 (2017) (finding that spending in the last week of the budget year is 4.9 times higher than during other weeks, as agencies rush to spend allocated funds they would otherwise lose).

<sup>77</sup> See C. Northcote Parkinson, *Parkinson’s Law*, ECONOMIST (Nov. 19, 1955), <http://www.economist.com/node/14116121> [<https://perma.cc/S533-P8P8>] (last visited Dec. 3, 2017). Succinctly put, Parkinson’s Law is that “work expands to fill the time available for its completion.” See also Ruti Gafni & Nitza Geri, *Time Management: Procrastination Tendency in Individual and Collaborative Tasks*, 5 INTERDISC. J. INFO., KNOWLEDGE & MGT. 115, 115 (2010) (finding evidence that MBA students usually procrastinate and submit individual assignments shortly before their deadlines).

<sup>78</sup> See Ariely & Wertenbroch, *supra* note 15, at 219.

<sup>79</sup> “[Doctors] . . . may rush to complete their work, spending less time than socially optimal on tasks they . . . accept [as their shifts end]. Since workers usually have much more discretion [about how hard they work than about how many hours they work],” these qualitative distortions could be costly. David C. Chan, *The Efficiency of Slacking Off: Evidence from the Emergency Department* (Nat’l Bureau of Econ. Research, Working Paper 21002, 2015), <http://www.nber.org/papers/w21002> [<https://perma.cc/V7DJ-QCJX>].

<sup>80</sup> “[T]he rate of allowance [application approval on review] . . . is drastically lower for . . . applications reviewed [just before an examiner’s deadline] . . . relative to [applications reviewed earlier in the cycle].” Frakes & Wasserman, *supra* note 74, at 5.

<sup>81</sup> See Thomas Miceli, Kathleen Segerson & Peter Siegelman, *A Model of Judicial Effort Allocation over Time* (Oct. 22, 2017) (unpublished manuscript) (on file with author).

quences.<sup>82</sup> Our work suggests that judges are no different. This is important because the CJRA's authors naively assumed that judges would respond to an incentive for faster case- and motion-processing *only* by speeding up their decisions, leaving every other aspect of their work unaffected. As we show, that has not been true.

## II

### DATA AND METHODS

In this part we discuss the data and methods we used to assess whether judges respond to incentives and to investigate the effect of their response on litigants.

#### A. Data Sources: Case and Motion Level

##### 1. Case-Level Data

The case-level analysis that follows relies on the Administrative Office of the U.S. Courts (AO) civil terminations data, now known as the Federal Judicial Center Integrated Database. From 1980 to 1987, we use data assembled (and generously provided to us) by Professor William Hubbard.<sup>83</sup> For the period of 1988 to 2017, we use the AO/FJC data.<sup>84</sup> We supplement these data with interviews with lawyers, judges, and court administrators to whom we have promised anonymity.

The AO dataset has been used extensively to study various aspects of federal district court outcomes.<sup>85</sup> Unfortunately, it contains only minimal information about each case. In partic-

---

<sup>82</sup> For an excellent overview, see Richard Rothstein, *The Influence of Scholarship and Experience in Other Fields on Teacher Compensation Reform*, in *PERFORMANCE INCENTIVES* 87–105 (Matthew G. Springer ed., 2009).

<sup>83</sup> Professor William Hubbard of the University of Chicago Law School generously allowed us to use his carefully cleaned version of this data. For a fuller description of this dataset, see William H.J. Hubbard, *Testing for Change in Procedural Standards, with Application to Bell Atlantic v. Twombly*, 42 *J. LEGAL STUD.* 35, 49–53, 60 (2013).

<sup>84</sup> Published by the Federal Judicial Center. See *Integrated Database (IDB)*, *supra* note 25.

<sup>85</sup> Among the many articles that make use of the AO database are John J. Donohue III & Peter Siegelman, *The Changing Nature of Employment Discrimination Litigation*, 43 *STAN. L. REV.* 983, 985 n.3 (1991); Jonah B. Gelbach, Note, *Locking the Doors to Discovery? Assessing the Effects of Twombly and Iqbal on Access to Discovery*, 121 *YALE L. J.* 2270, 2324 (2012); Gillian K. Hadfield, *Where Have All the Trials Gone? Settlements, Nontrial Adjudications, and Statistical Artifacts in the Changing Disposition of Federal Civil Cases*, 1 *J. EMPIRICAL LEGAL STUD.* 705, 705 (2004); James A. Henderson, Jr. & Theodore Eisenberg, *The Quiet Revolution in Products Liability: An Empirical Study of Legal Change*, 37 *UCLA L. REV.* 479, 483 (1990); Hubbard, *supra* note 83.

ular, there are no judge identifiers. For every case, the data provide the district and office in which the case was filed, a “nature of suit” code,<sup>86</sup> the dates on which the case was filed and was terminated, the case’s procedural progress at termination, and the winning party, if there is one. The data track every civil case that terminated in any U.S. district court between 1980 and June 30, 2017, regardless of filing date. Appendix Table A provides some summary statistics for this data.

## 2. Motions Data

We also make use of a hand-coded random sample of 758 summary judgment motions filed between August 1 and September 30, 2011.

Since the AO data do not track motions, we could not sample them directly from the dataset. To assemble the motions data, we began by picking a 178-day interval of time between November 4, 2010, and May 1, 2011. We then randomly sampled days (excluding weekends) from this interval, and examined every case filed on those days to determine if it contained a summary judgment motion that was filed between August 1, 2011, and September 30, 2011, making it eligible for inclusion in our sample. We compiled a variety of information about these motions, including nature of suit code, judge name, case- and motion-disposition dates, and supplementary data about the motion disposition (such as the length of the opinion and the presence of a magistrate).<sup>87</sup> Figure 2 summarizes the sampling scheme.<sup>88</sup>

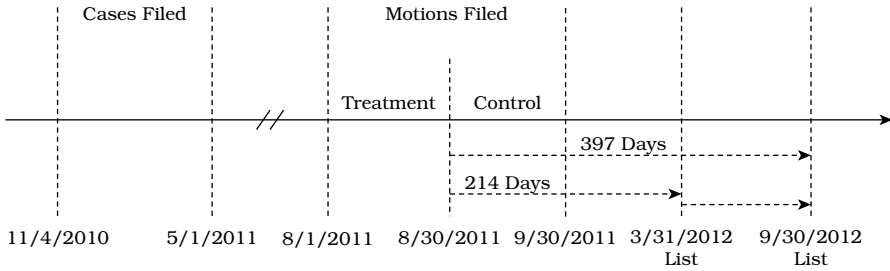
---

<sup>86</sup> The AO data includes roughly 340,000 cases that were pending as of June 30, 2017, when the data were compiled. We drop these cases from our analysis. The case-level data on which we rely are compiled from official court records, with the exception of the Nature of Suit (case type) data, which relies on information provided by plaintiffs’ attorneys on the “Cover Sheet” that accompanies each complaint. See Christina L. Boyd & David A. Hoffman, *The Use and Reliability of Federal Nature of Suit Codes*, 2017 MICH. ST. L. REV. 997, 998–99.

<sup>87</sup> All motions were double-coded by different research assistants, and anomalies were resolved by one of the authors.

<sup>88</sup> After our analysis was completed, Professor William Hubbard pointed out that our sampling scheme introduced an inappropriate difference between our control and treatment motions. To see why, we define the SJ Filing Lag as the length of time between the date a case is filed and the date on which the summary judgment motion is filed. (For example, if a case is filed on Friday, January 7, 2012, and the defendant files a motion for summary judgment on Tuesday, April 3, 2012, the SJ Filing Lag is eighty-eight days.) Because the control group motions were all filed one calendar month later than the treatment group motions (September vs. August), yet both sets of motions came from the same sample of cases, the treatment group motions will tend to have a larger SJ Filing lag purely as an artifact of the sampling method we used. Fortunately, there is a simple solution to this problem: drop the treatment group motions arising from cases

FIGURE 2: SAMPLING SCHEME, SHOWING ELIGIBILITY OF MOTIONS FOR 6 MONTH LIST BY CASE & MOTION FILING DATE



B. Methods

While virtually all of our results are visible in graphic form, we also rely on statistical analyses. Here, we explain and justify the approaches we use.

As is well-known in empirical social science, standard regression analysis is ultimately only correlational: it can establish that two variables are “associated” with each other, but it cannot really shed much light on whether X causes Y, Y causes X, or some third variable Z causes both X and Y. To better address these problems of causal inference, statisticians have developed several advances on standard methods, and we deploy two of them in our analysis. Both strengthen our claim that the List has truly causal effects on various aspects of judicial behavior.

1. Difference-in-Differences

One way to fortify a causal claim is to use a “difference-in-differences” study design. In our context, the idea is that we are not just testing whether there are regular spikes in the volume of case closures in List weeks. To be sure, that is evidence of a *correlation* between List dates and case closures (as in Figure 1), but it does not establish that the List is what *causes* these spikes. Perhaps the peaks in closures are being driven by something else that regularly occurs at these times. For example, maybe school vacations happen to occur in late March and late September every year, and it is really these

---

filed in November, and drop the control group motions filed in April. We re-analyzed all the motions data in the paper using this correction, and found no significant changes in the results. One area in which the results were affected was in the covariate balance results, where we had one more covariate—the difference in days between the case filing and the summary judgment motion being filed, which for some bandwidths is more unbalanced when we rely on the sampling strategy proposed by Professor Hubbard.

vacations that are driving the bunching we observe. If so, then attributing the difference in case closures in List weeks to the List itself would be a mistake.<sup>89</sup>

To strengthen our assertion that the List actually *causes* bunching in closures, we rely on the availability of pre-List case data.<sup>90</sup> Instead of measuring the effects of the List by the difference between closures in week 13 and other weeks, we compare *differences* between closures in week 13 and other weeks in the period before and after the CJRA went into effect. This “double-differencing” eliminates the effects of any potential confounding factors that are constant across the pre/post period. If vacation schedules are constant over time, they cannot explain why List weeks see more case closures relative to other weeks *after* the List was compiled, but not before.

## 2. *Quasi-Random Assignment*

We also make use of a second strategy based on a plausible analog to the way a classic laboratory experiment would measure the effects of the List.

In an ideal world, the first step one would take to measure the effects of the List would be to randomly assign some judges to be “treated” (i.e., subject to the List) and others to serve as untreated “controls.”<sup>91</sup> The claim that the List causes, say, a reduction in case duration, could then be evaluated simply by comparing the durations of the treated and untreated cases. Assuming no factors are correlated with filing date cutoff for List eligibility, which we think is reasonable, any difference in average case duration between the two groups could be attributed to the treatment, since any and all other confounding

---

<sup>89</sup> The existence of bunching conclusively demonstrates that judges are paying attention to the calendar for *some* reason; so even if that reason is not necessarily the Six-Month List, there would still be something important going on.

<sup>90</sup> Our motion data run only from 2011 through 2013, and do not span the pre-CJRA period. Thus, we cannot use difference-in-differences methods with these data.

<sup>91</sup> A better approach might be to randomize at the level of judicial districts (groups of judges), given that the List apparently derives its effectiveness from peer comparisons. This sort of subtlety plays an important role in the design of social experiments, but it is irrelevant for our purposes. Whether experimental data necessarily constitutes the “gold standard” for causal inference is still a contested proposition. For an extended discussion of these topics, see Angus Deaton, *Instruments, Randomization, and Learning about Development*, 48 J. ECON. LITERATURE 424, 426 (2010); James J. Heckman, *Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy*, 48 J. ECON. LITERATURE 356, 357 (2010); Guido W. Imbens, *Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)*, 48 J. ECON. LITERATURE 399, 400 (2010).



factors would—by definition—have been randomized-away at the start through our random assignment of motions to the two groups.

In the real world, of course, it is often impossible or impractical to conduct a randomized controlled trial, and we did not attempt to do so. Instead, we sought a source of variation in exposure to the List that is driven by factors unrelated to the List, and is thus “as good as random” for our purposes. We suggest that such a source is the filing date of a lawsuit or a motion.

To see the basic idea in the context of motions, return to Figure 2. As it shows, the cutoff date for a motion to be eligible for the March 31, 2012 List is August 30, 2011; a motion filed on August 31, 2011 would have been only be 213 days old if it were still open when the March 31, 2012 List was compiled, and hence would not have been old enough to appear on that List. In fact, it would not be eligible until the *next* List, compiled on September 30, 2012, a full 396 days after it was filed. Conversely, a motion filed two days earlier (August 29, 2011) *would* be eligible for the March 31, 2012 List, by virtue of being 215 days old at the List date. This means that two motions are treated differently, despite very small differences in their filing dates. The earlier motion faces a 215-day deadline, while the later motion has 53% longer before it is at risk of appearing on the List. If the deadline actually reduces duration, motions such as the first should have shorter average durations than those such as the second.

Exactly the same logic applies to case filings, although on a different time scale. A case is eligible to appear on a List if it was filed more than three years (about 1,095 days) before that List is compiled. So a case filed on September 29, 2008, would be eligible for the September 30, 2011 List, while a case filed on October 1, 2008, (two days later) would not be eligible until the March 31, 2012 List, six months later.

If the timing of motion or case filings is “as good as random” over short intervals of time, then we can apply the experimental logic above to make inferences from our quasi-experimental data. As long as there are no systematic patterns in the filing of cases or motions inside a relatively narrow window on either side of the cutoff date for List eligibility, a comparison of the “treated” and control groups will yield an accurate assessment of the causal effect of the List. The obvious question then becomes: how plausible is it that motions (or cases) are randomly filed over the time periods in question?

a. *Assessing Randomness in Filing Dates*

Of course, case and motion filing dates are not produced by random number generators, and thus are not truly random. Fortunately, however, our methods do not actually require true random assignment; all we need is filing dates that are not chosen in response to the existence of the List, or to factors correlated with it. We have three kinds of evidence that support this conclusion.

First, there is direct evidence from conversations with informed parties (lawyers, judges, and court administrators). Our identification strategy would be seriously undermined if parties deliberately file complaints or motions on August 29th rather than August 31st in order to take advantage of the stricter deadline that governs items filed before the cutoff. Similar problems would be presented if judges could somehow manipulate filing dates (which is not impossible, at least for motions). But our discussions strongly suggest that parties are not manipulating case or motion filing dates in the shadow of the List. Even though judges are acutely aware of which cases and motions will show up on the next List, most apparently do not have a firm grasp of the subtleties of how List-eligibility is determined.<sup>92</sup> The same appears to be true of the federal bar: none of the lawyers we spoke with had ever heard of anyone

---

<sup>92</sup> We did hear rumors that (some) judges (occasionally) resort to various kinds of administrative manipulations to avoid showing up on the Six-Month List. For example, one lawyer told us that a judge had asked him to withdraw a motion that was about to appear on an upcoming List, and to refile it after the List was compiled. Interview with plaintiff-side lawyer (Oct. 27, 2016) (on file with authors). Here is another example from our research: a summary judgment motion was filed on August 26, 2011, making it eligible for the March 31, 2012 List if not resolved before then. There was a settlement conference, apparently unsuccessful, after which the judge filed an order “resetting” the submission date on the motion to October 12, 2011, thus essentially giving himself a year to resolve the motion without appearing on the List. The motion was ultimately resolved on March 19, 2012, within the six-month time period of the original summary judgment filing. *Ledet v. Treasure Chest Casino*, No. 2:10-CV-04561 (E.D. La. Sept. 15, 2011) (“ORDER GRANTING MOTION to Continue, Resetting Submission Date on MOTION for Summary Judgment. Motion reset for 10/12/2011 on the briefs. Signed . . . on 9/14/11.”).

There are also rumors that judges will sometimes order a case “administratively closed” before a List date, thereby removing it from the official backlog, and then reopen it afterwards. An administrative or “statistical” closure occurs when the clerk marks the case “closed” upon a judicial order, but the case is subject to reopening. (This might occur, for example, if there is a bankruptcy proceeding involving one of the parties that has priority over the primary case.) Such closures can be distinguished, for example, from a dismissal or a dismissal with prejudice which would require a party to refile their case; administrative closures are entirely within judicial control. We tested whether “statistical closures”—there are only about 100,000+ in our data—are more likely in List weeks (or

timing the filing of a motion (and certainly not a complaint) because of the List. While all of them knew about the Six-Month List, most lawyers were also largely ignorant about its mechanics. Most believed that any motion older than six months appeared on the List, which is inaccurate.

Second, we can look at data on the timing of case and motion *filings*. Any signs that there are spikes in filings just before (or after) List eligibility dates would be strong evidence of manipulation. Fortunately, we find no such evidence.

Finally, we can look indirectly at evidence on “covariate balance.” Although we cannot observe every variable of interest, we can assess the similarity of the treatment and control groups across the set of case or motion characteristics that we *do* observe.<sup>93</sup> The idea is that if parties were deliberately manipulating filing dates, we would see evidence of such interventions in the data, in the form of pre-treatment differences between the “treatment” and “control” groups.<sup>94</sup> The two groups should be very similar across all the variables we can observe. If they are not, our identification strategy becomes suspect.

In Table B of the Appendix, we present “covariate balance” tests for the motions sample. The evidence establishes that the treatment and control groups are substantially similar across most of the variables we identify as important and suggests—though it cannot dispositively prove—that there is no manipulation that threatens our empirical strategy.

---

shortly before). We found no such evidence, nor did we find that re-openings of these cases were more likely in periods just after Lists were compiled.

<sup>93</sup> Unfortunately, we cannot test whether the treatment and control groups are identical in *all* dimensions, as would be true by definition if assignment to treatment and control groups were actually random. There are inevitably facts about each case that we cannot observe, so in the absence of actual randomization, we cannot be sure that seemingly identical cases might differ along some dimension that we don’t see.

<sup>94</sup> “Pre-Treatment” refers to any characteristics of subjects that exist before the start of an experiment. With true random assignment, such differences are not a cause for concern because they are guaranteed to be zero on average in sufficiently large samples. But we do not have true experimental data, so it is possible that there might be differences between “treatment” and “control” groups. For example, suppose August-filed motions all came from Employment Discrimination cases while September-filed motions all came from Antitrust cases. That would imply nonrandom differences between the treatment and control groups. Moreover, such differences, being logically prior to the intervention we are studying, could not have been caused by the treatment itself, and therefore would confound our attempt to identify the effects of the List.

b. *Further Measurement Issues*

An additional requirement for our method to yield an accurate assessment of the List's effects is that the List should have no effect on the "untreated" (control) group of cases or motions. Unfortunately, we do not have a truly untreated control, since we can only compare motions subject to a 214-day deadline with those subject to a 397-day deadline.<sup>95</sup>

Fortunately, most summary judgment motions take less than a year to resolve.<sup>96</sup> So the one-year deadline is not likely to be a binding constraint on most decisions. Nevertheless, we stress that our results only capture the effects of a Six-Month List deadline *relative to a twelve-month deadline*, not to no deadline at all. Similarly, it is important to remember that motions continue to be subject to deadlines at six-month intervals. That is, an August-filed motion is subject to both the March and, if it is not resolved, the September deadline, and so on, ad infinitum.

If the List delays resolution of the "untreated" motions, then the *difference* between the treated control and treatment groups will still be an unbiased measure of the List's effect on duration. But it would measure both the *faster* processing of the treated group *and* the *slower* processing of the untreated group. Given the limitations of the data, we are unable to establish whether the effects we observe are reductions in duration for the treated group or increases in duration for the control group. This is an important caveat in interpreting all the results that follow.

### III

#### THE EFFECTS OF THE LIST

This Part empirically demonstrates the effects of the List on the federal judiciary. We discuss three kinds of influences: *bunching* of closures, *duration* effects, and *outcome* effects.

Both motion- and case-level data show that there is a 20% to 30% spike in the number of closures (dispositions) just

---

<sup>95</sup> The literal deadlines are 244 to 214 days for motions filed between August 1 and August 30, inclusive; and 365 to 395 days for motions filed between August 31 and September 30, inclusive. For cases, the deadlines are three years for motions filed in the last week of March or September, versus 3.5 years for motions filed in the first week of April or October.

<sup>96</sup> In our sample, the average summary judgment motion takes about 170 days to resolve—much less than one year. A twenty-year deadline would be the same as no deadline at all, since it presumably would never be binding on any judge's decision.

before the List is compiled. We document a change in judicial behavior when the List was instituted: there is more bunching of case closures after the CJRA went into effect than previously. Moreover, these findings are not merely correlational: we find that List-eligible motions exhibit bunching of closures in List weeks, but plausibly identical motions that are ineligible show no such pattern. Thus, we confirm that the List *causes* judges to bunch their decisions around the deadline.

We also detect a pattern of “secondary bunching” in the motions data. Roughly 20% of the August-filed (treatment group) motions that survived past their initial (March 31, 2012) deadline are not decided until the last three weeks before the *next* List date in September, 2012, and almost 13% are not decided until the last week before the deadline. The most likely explanation for this phenomenon is a “mothballing” effect—motions that are not resolved by their first Six-Month List are apparently ignored until the next deadline draws near. None of the bunching we observe (and especially not the secondary bunching) could plausibly have been intended by the proponents of the CJRA.

Since the explicit purpose of the List was to reduce backlogs and delays, we also investigate its effects on case and motion *durations*. We find only very weak evidence that the three-year case-level deadline reduces time to disposition. Case duration appears to be reduced by about 4%, but only for a very restricted set of cases. Most cases are unaffected by the three-year deadline, which is unsurprising, since virtually all cases terminate long before the three-year constraint becomes binding. Comparing “just-eligible” with “just-ineligible” summary judgment *motions*, however, we find that having a six-month deadline shortens the duration of summary judgment motions that close in less than two years by ten to thirty days (6% to 18%) relative to having a one-year deadline.<sup>97</sup> That number is misleading, however; while the List reduces average motion duration, it has no effect on the duration of most motions, and it may even lengthen the duration of those that survive past their first deadline. As a consequence, the vari-

---

<sup>97</sup> It is important to keep in mind that in what follows, the comparison is between longer and shorter deadlines, not between a six-month deadline and no deadline at all. We measure the effects of the List as the difference between the duration of motions subject to a six-month deadline and those subject to a one-year deadline. We can only identify the *difference* in duration between these two groups; we cannot say whether that difference occurs because the one-year deadline *increases* the duration of motions or the six-month deadline *decreases* it. For further discussion, see *supra* subpart II.B.

ance or dispersion of motion durations is increased—another unintended consequence of the List.

We also find evidence indicative of a List effect on case and motion *outcomes*, or at least on their timing. This finding leads us to worry that the List creates the risk of a negative effect on accuracy. Cases that are decided in the immediate shadow of a List are about 18% (7 percentage points) more likely to favor defendants than those decided at other times. Moreover, List week closures have a 40% higher likelihood of being remanded on appeal than do other cases (albeit from an extremely low base level). Strengthening our confidence that we have captured a causal relationship, neither win rate nor remand rate effects were present before the CJRA was implemented in 1991.

Turning to motion-level data, we show that eligible motions decided in the week before a List is compiled are more likely to be *denied* than at other times. This, too, is a causally robust finding. Furthermore, the List deadline seems to increase slightly the chances that defendants will prevail in their lawsuit when they move for summary judgment, consistent with the case-level data.

Our bottom line is simple: judges respond to the incentive created by the Six-Month List. But the List has had only mixed success in achieving its intended goals, and it has had collateral consequences (win-rate dips, increases in remand rates, lengthened duration for some motions) that were obviously unintended and likely undesirable. We discuss the implications of these findings for policymakers in the final section of this Article.

## A. Bunching

In this section, we demonstrate that judges respond to the List by shifting decisions to the period just before the deadline.

### 1. *Case-Level Bunching*

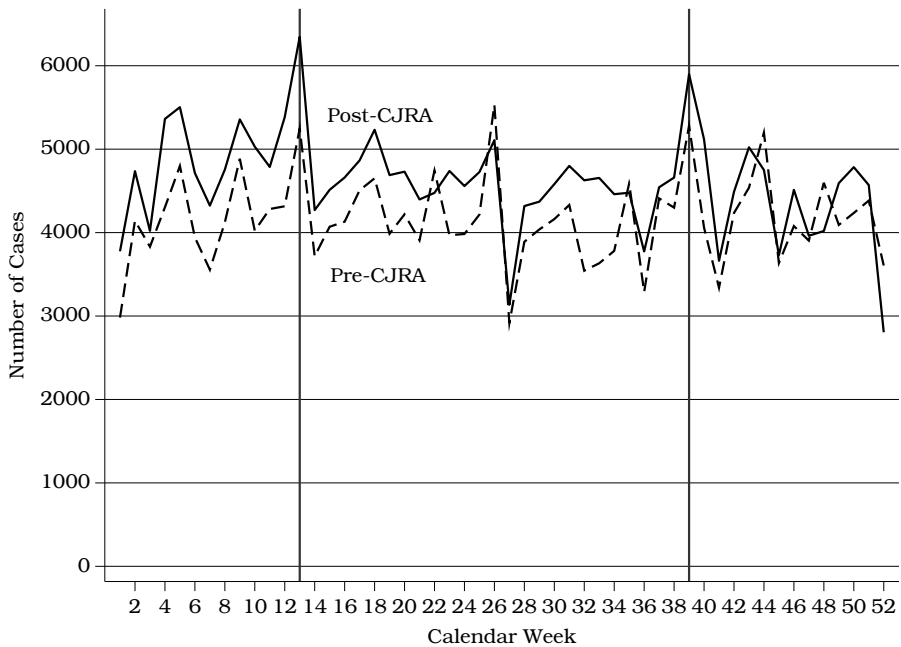
#### a. *Spikes in Volume of Case Closures*

Just before a List is compiled (weeks 13 and 39), the number of case closures increases dramatically, returning to normal volumes immediately afterward. While there are, of course, some spikes in non-List weeks in some years, the pattern of List-week spikes is consistent throughout the period from 1991 to 2017.

We note that besides being List dates, March 31 and September 30 also mark the ends of the first and third calendar quarters. To assess whether the List or some other event oc-

curing at quarterly frequencies (staff turnover or vacations, for example), is causing the spike in closures, we use a difference-in-differences approach.<sup>98</sup> Figure 3 compares the weekly volume of case closures before the Civil Justice Reform Act went into effect (in 1991) and after. If the March and September spikes are driven by some constant quarterly pattern, they should be the same before and after the CJRA was passed. If instead, the spikes are caused by the presence of the List, then bunching in late March and late September should emerge only *after* the CJRA went into effect.

FIGURE 3: AVERAGE WEEKLY CASE CLOSURES, BY CALENDAR WEEK: 1980–1990 & 1991–2017



In fact, some List-week bunching is clearly discernible even before the CJRA went into effect, but the size or extent of bunching gets substantially larger after 1991. The most likely explanation is that while there was no Six-Month List before the CJRA took effect, there were quarterly reporting requirements, at least in some circuits.<sup>99</sup> Thus, there was no “clean” break before and after the passage of the CJRA. The difference

<sup>98</sup> See *supra* subpart II.B.

<sup>99</sup> See, e.g., Geyh, *supra* note 29, at 520–23 (noting that each Circuit’s judicial council took action to alleviate decision-making delay “by calling judges to task for delays, suspending their caseloads, or reshuffling their dockets”).

between pre- and post-CJRA bunching is therefore a conservative estimate of the effects of the Six-Month List, since it includes the influences of any CJRA precursors (proto-Lists). Nevertheless, differences between the pre- and post-CJRA period suggest that the move from a more ad hoc quarterly reporting system to a mandatory biannual deadline substantially changed judicial behavior.<sup>100</sup>

For readers unconvinced by the graphical demonstration, Table 1 presents quantitative evidence on List week bunching. Our basic specification is given in equation (1).

$$Close_t = \beta_0 + \beta_1 Termweek_t + \beta_2 Termweek_t^2 + \beta_3 Week\_13 + \beta_4 Week\_39 + \beta_5 CJRA + e_t \quad (1)$$

Here,  $Close_t$  is the number of cases closing in week  $t$  (running over the 1,956 weeks between January 1, 1980, and June 30, 2017).  $Termweek_t$  and its square are linear and quadratic time trends.  $Week\_13$  and  $Week\_39$  are dummy variables that are 1 in, respectively in the thirteenth and thirty-ninth week of any year (zero otherwise). And  $CJRA$  is a dummy variable that is 0 before 1991 (when the CJRA went into effect and the List was created).<sup>101</sup> Some specifications also interact the  $CJRA$  variable with the  $Week\_13$  and  $Week\_39$  dummies.

<sup>100</sup> The other important pattern that emerges from Figure 3 is the presence of bunching in week 26 (ending July 30). July bunching diminishes after 1991. We attribute this effect to the change in the Federal judiciary's fiscal year, which ended on July 30 until 1992, after which it was moved to September 30. The disappearance of July bunching with the change in the court's fiscal year provides further proof that judges follow the calendar in deciding when to decide cases. Before the CJRA there was also greater bunching right before the courts' fiscal year (which started on July 1); this bunching diminishes after the biannual reporting deadlines are instituted.

<sup>101</sup> Equations in Table 1 are all time series regressions and are estimated with Prais-Winsten correction for AR(1) errors and with robust standard errors. Alternative specifications (not reported) in which we drop (i) cases with unusually long durations, (ii) class action cases, (iii) cases that are not original jurisdiction, and (iv) cases that involve arbitration do not qualitatively alter any of our results. Logged transformations also have no qualitative effects. We also worried about a 'misalignment' between STATA's definition of calendar weeks and List dates, which are defined by specific calendar days (March 31 and September 30). In nonleap years, the thirteenth calendar week runs from March 26 through April 1, extending beyond the List date. The thirty-ninth calendar week starts on September 24 and runs through September 30, so there is no misalignment problem. Conversely, in leap years, the thirteenth calendar week starts on March 25 and ends, appropriately, on March 31; but the thirty-ninth calendar week starts on September 23 and runs only through September 30, omitting the day before the List is compiled. We reran the regressions in Table 1 after reclassifying STATA's calendar weeks to correct for these misalignments (so that the thirteenth week always ended on March 31 and the thirty-ninth week always ended on September 30) by adjusting the length of calendar weeks far from the List dates (weeks 26 and 52). Our results were unchanged by these adjustments.



TABLE 1: OLS REGRESSIONS EXPLAINING VOLUME OF WEEKLY CASE CLOSURES: 1980–2017

	(1)	(2)	(3)	(4)	(5)
	1980–2017	1980–2017	1980–2017	1980–1990	1991–2017
Time	3.25*** (0.487)	6.04*** (0.679)	6.04*** (0.676)	94.87*** (3.936)	1.16 (1.164)
Time <sup>2</sup>	-0.0006*** (0.00012)	-0.0012*** (0.00015)	-0.0012*** (0.00015)	-0.0343*** (0.00151)	-0.0001 (0.00025)
Week 13		1,475*** (122.2)	1,035*** (283.9)	1,100*** (157.1)	1,691*** (117.3)
Week 26	1,007*** (145.5)	981.4*** (140.1)	980.1*** (139.8)	1,508*** (202.9)	588.9*** (86.7)
Week 39		1,249*** (146.8)	1,135*** (269.8)	1,129*** (159.9)	1,423*** (126.1)
Week 13 × CJRA			643.3** (301.7)		
Week 39 × CJRA			163.7 (320.4)		
CJRA		-850.2*** (136.1)	-865.6*** (136.4)		
List week	1,357*** (99.1)				
Constant	472.0 (472.5)	-2,132*** (651.0)	-2,120*** (647.9)	-60,668*** (2,530)	2,524* (1,306)
Observations	1,949	1,949	1,949	571	1,377
Adj R <sup>2</sup>	0.142	0.155	0.156	0.673	0.090
Durbin-Watson	2.080	2.069	2.069	2.049	2.004

Note: All regressions using Prais-Winsten correction for autocorrelation. Robust Standard errors in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

The coefficient estimates in the first row of Table 1 illustrate a modest tendency for the number of case closures to increase over time (about three to four cases per week for the period as a whole), with a small (negative) quadratic term.<sup>102</sup> From Column 1, we see that List weeks (Weeks 13 and 39) have about 1,250 more closures than the average week. That is about 30% more than the average week, and is statistically significant. In Columns 3 and 4, we adopt a difference-in-differences specification. That is, we compare the volume of week 13 and week 39 closures to those in average weeks, and then look at how this difference changes when we move from the pre- to the post-CJRA era (when the List went into effect). The additional List-week cases post-CJRA are measured by the interaction terms labeled *Week\_13x Post-CJRA* and *Week\_39x*

<sup>102</sup> The time trend was much larger—ninety cases per week—in the period between 1980 and 1990 (Column 5), when overall case volumes were rising significantly.

*Post-CJRA*. Using this measure of the effects of the List, we find about 160 to 640 more case closures in List weeks after the CJRA than List weeks before the Act was passed. Column 4 also confirms another dramatic List-like effect. As noted earlier, when the courts' fiscal year was moved from July 30 (week 26) to September 30 (week 39) in 1992, the volume of case closures in week 26 dropped by almost 900 (compare Columns 4 and 5), although it remained statistically significant. Finally, Columns 5 and 6 allow for comparison of the same specification in the pre- and post-CJRA periods. The *Week\_13* and *Week\_39* coefficients are 300–600 cases per week larger after 1991 (while the *Week\_26* coefficient is about 900 cases per week smaller).

The evidence shows unusually large volumes of case closures in weeks 13 and 39. This pattern was apparent even before the Six-Month List was created, so it is unlikely that *all* of the bunching can be attributed to the List per se. We attribute the bunching before 1991 to the quarterly reporting requirements that were a precursor to the List. Excess closures in weeks 13 and 39 grow substantially after the CJRA went into effect in 1991. This pattern shows that judges pay attention to deadlines deciding when to close cases. It also suggests that it is the Six-Month List in particular that drives judges to close cases in weeks 13 and 39.

Evidence from the volume of weekend closures further supports our claim that judges make extra effort to close cases to meet List deadlines. Cases rarely close on weekends, and they do so much less often now than in the past.<sup>103</sup> Nevertheless, we find that weekend closures spike very substantially if a March 31 or September 30 List deadline falls on a Friday, Saturday, or Sunday. Moreover, this weekend spike is only evident for the period after 1990 when the CJRA went into effect. Before 1991, March 31 or September 30 weekends had no more closures than ordinary weekends.<sup>104</sup>

---

<sup>103</sup> We suspect that what appeared to have been weekend closures were due to data entry errors that have been greatly reduced as electronic record keeping has improved.

<sup>104</sup> Results available from authors. We are not sure exactly how these weekend closures are accomplished. It is possible that judges issue orders on a Saturday or Sunday, and the court clerk then docket those orders retroactively on the Monday immediately afterwards, *before* compiling the Six-Month List. Alternatively, staff may work on weekends to enter orders on the day they are actually issued.

b. *No Bunching of Case Filings*

Bunching is being driven by judicial decisions, not by the timing of case or motion *filings*, and we find no evidence—either quantitative or anecdotal—that there is manipulation of case or motion filings by judges or parties.<sup>105</sup>

We find no evidence that List weeks have larger numbers of case *filings* than any other times during the year. Indeed, there appears to be no weekly pattern at all in the volume of case filings, except that week 27 shows a substantial decline vis-à-vis other weeks.<sup>106</sup> In a regression controlling for time trends, there is no List week effect on filing volumes. Neither is there any evidence that bunching of filings in List weeks changed after the CJRA went into effect in 1991. Finally, the fiscal year change in 1992 had no discernible effect on week 26 filings.

These findings are consistent with our interviews of lawyers on both the plaintiff and defense side, who uniformly report that they do not consider the List in timing the filing of cases or motions.<sup>107</sup>

2. *Motion-Level Bunching*

As described earlier, we compare random samples of motions subject to the six-month deadline (the treatment group) with those subject to a twelve-month deadline (the control group) as described earlier.<sup>108</sup> Figures 4A, 4B, 4C, and Table 2 summarize the date-of-disposition data from our sample of summary judgment motions.

---

<sup>105</sup> See discussion *supra* note 92.

<sup>106</sup> Results available from the authors.

<sup>107</sup> We conducted semistructured anonymous background interviews with five lawyers with sophisticated federal court practices at medium to large firms, three on the defense side and two on the plaintiff's side. Interviews with court clerks and judges confirm this finding. Lawyers report that the List only affects their filing of motions when the judge requests a schedule change, although not all of them report such requests on the part of judges.

<sup>108</sup> Motions filed before August 30, 2011, were eligible for inclusion on the March 31, 2012 List if they were not resolved by then. Motions filed after that date were only eligible for the September 30, 2012 List. Recall that in practice, this "six-month" deadline is actually 214 days and the control group deadline is 397 days.

TABLE 2: SUMMARY STATISTICS ON MOTION BUNCHING: 2012

	Aug. 1-29	Aug. 30-Sept. 28
Number Filed	391	367
Number closed by March 31	325	240
% closing by March 31	83.1%	65.4%
% closing in last week of March	5.5%	2.5%
% closing in last 3 weeks of March	14.2%	4.6%
Number open <i>after</i> March 31	66	127
Number closing in last week of September	9	15
% (of those open after March 31)	13.6%	11.8%
Number closing in last 3 weeks of September	16	27
% (of those open after March 31)	24.2%	21.3%

a. *Bunching in Motion Dispositions*

Our analysis shows that, consistent with the case-level data, judges bunch their motion decisions around List weeks. Importantly, we also find that Judges who fail to decide a motion in a timely fashion will sometimes delay their decision until the next deadline, creating what we call “secondary” bunching. Both of these phenomena are causally linked to the List.

FIGURE 4A: SUMMARY JUDGMENT MOTION RESOLUTION DATE, BY CALENDAR WEEK TREATMENT GROUP (MOTIONS FILED ON OR BEFORE AUG. 30, 2011)

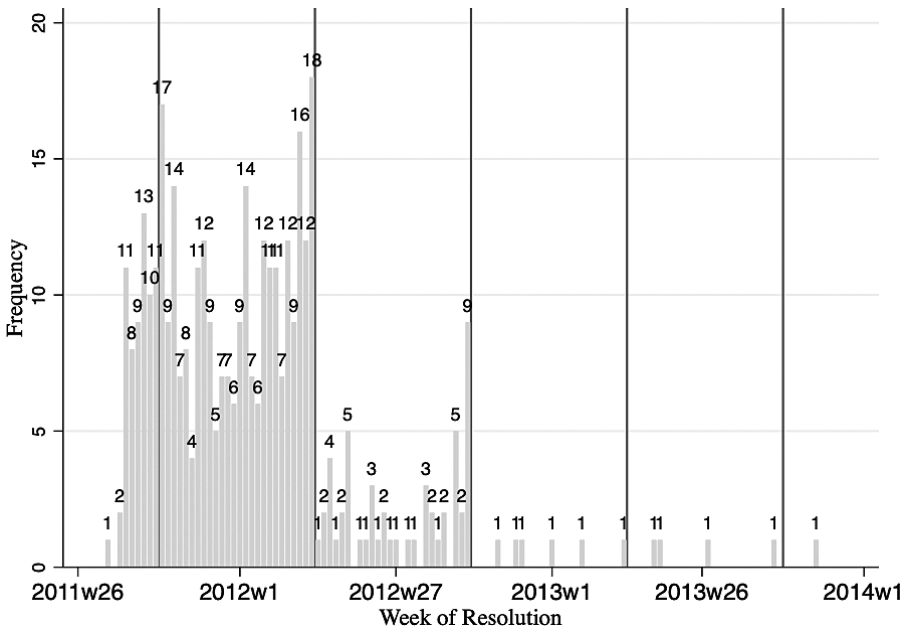


FIGURE 4B: SUMMARY JUDGMENT MOTION ORDER DATE, BY CALENDAR WEEK CONTROL GROUP (MOTIONS FILED AFTER AUG. 30, 2011)

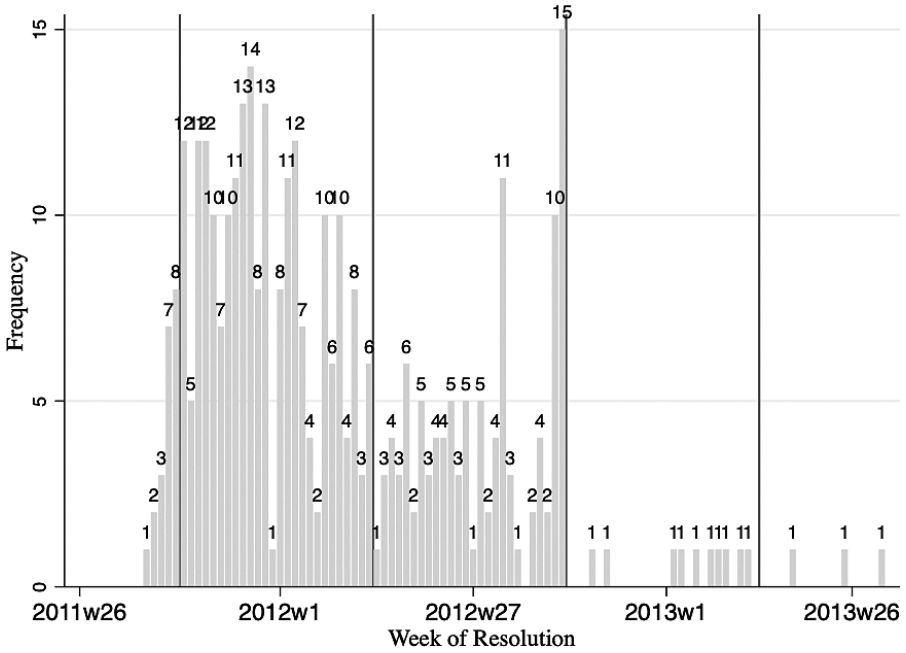
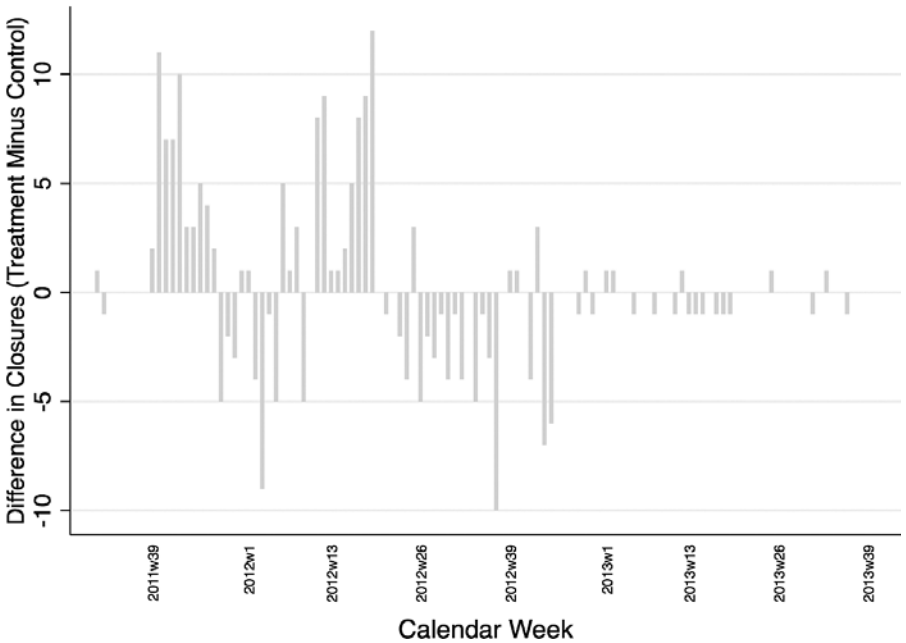


FIGURE 4C: DIFFERENCE IN THE NUMBER OF CLOSURES BETWEEN THE TREATMENT AND CONTROL GROUPS, BY CALENDAR WEEK



First, Figure 4 demonstrates that there is bunching of motion dispositions at List dates—there are obvious spikes in week 13 and week 39 of 2012, just as we found in the case-level data. The pattern is not uniform, however: August-filed (treated) motions exhibit a closure spike just before the March 31, 2012 List is compiled (Figure 4A). That is unsurprising, since they are eligible for inclusion on the March 31 List if they are not resolved by that date. But September-filed motions, which are *not* eligible for the March List, show no substantial uptick in closures until the last weeks of September 2012 (Figure 4B).

Second, Figure 4A shows evidence of what we term “secondary bunching”: the August-filed motions that survive past their first deadline at the end of March, 2012 *also* experience a bunching of closures at the September List date, more than a year after they were filed. Instead of working steadily to close these motions, judges seem to be procrastinating until the next reporting deadline. Meanwhile, September-filed motions close in greater numbers throughout the period between April 1 and September 30, 2012. While the List may speed up disposition of *some* of the August-filed motions (to meet the March deadline), those motions that survive past March 31 apparently get less attention than others, at least until the September deadline draws near. Paradoxically, the List seems to *lengthen* processing time for the motions that fail to close before their first deadline (March 31). We return to this issue in our discussion of List effects on duration.<sup>109</sup>

Third, the List has a causal effect on these phenomena. Figure 4C plots the difference between the number of closures in the two groups—it represents the vertical subtraction of the top two panels. The difference between treatment and control group closures is a measure of the true effects of the List, one that is not subject to confounding by omission of unobservable variables. The bottom line is that facing a six-month deadline is responsible for eighty-five more closures in the last week of March than facing a one-year deadline. Further, 83.1% of August-filed summary judgment motions closed before the March 31, 2012 List date. By contrast, the September-filed motions were much less likely to have been decided before the March List was compiled. Only 65.4% of this group were decided before March 31. Yet these motions were filed very close together in time (within one month). If we focus on the bunching

---

<sup>109</sup> See discussion *infra* subsection III.B.2.

of closures immediately prior to the List date, we see a further contrast between the treatment and control groups in Table 2: 5.5% of August-filed motions closed between March 25 and March 31, while only 2.5% of September motions closed in this interval. Column 1 also reveals secondary bunching—24.2% of the sixty-six treatment group motions that survive past March 31 closed in the last three weeks of September. Judges are not rendering decisions uniformly over time, but instead altering their behavior to decide a greater number of motions at the deadline.

b. *Controlling for Age?*

Figure 4A is based on the actual order date for each motion. On the reasonable assumption that the effect of List-eligibility is the same for all “treated” motions, regardless of the date they were filed, this is appropriate. The secondary bunching we observe for “treated” motions in late September 2012 should not be influenced by when (during August of 2011) the motion was filed. And even if it is, we do not care; what is important here is *when* the motion closed, not how *old* it was when it did so.

The data above are potentially misleading for comparisons between treatment and control groups, however, because they compare order dates for motions of different chronological ages in a way that potentially confounds the List effect that we seek to measure. For example, a motion filed on August 2, 2011, was 242 days old when the March 31, 2012 List was compiled, but a motion filed on September 29, 2011, was only 189 days old. It is not appropriate to contrast the behavior of these two motions as they approach the March 31 deadline, since doing so conflates a potential age effect (older motions may be more likely to close just by virtue of that fact) with a pure List effect (the older motion is List-eligible while the younger one is not).

To remedy this problem, we adopt a technique from the statistical analysis of survival or failure-time data.<sup>110</sup> We “standardize” every motion’s age by assuming that it is filed on

---

<sup>110</sup> On the distinction between “analysis time” (pure duration, in which all observations are considered to start on the same date) and calendar time (based on actual start dates, which may differ across observations), see, e.g., DAVID W. HOSMER, STANLEY LEMESHOW & SUSANNE MAY, *APPLIED SURVIVAL ANALYSIS: REGRESSION MODELING OF TIME-TO-EVENT DATA* 6–7 (2d ed. 2008) (noting that researchers can incorrectly calculate survival time because they do not statistically account for subjects entering studies at different times).

August 30, 2011.<sup>111</sup> We then measure all subsequent dates *relative* to that starting point. Thus, the March 31, 2012 List date is simply day 214 of the study. As Appendix Figures B(1), B(2), and B(3) demonstrate, with one exception, our results are robust to this adjustment. Our secondary bunching finding does go away, since by controlling for motion age *within* the treated group, we lose the connection to the real-time calendar that drives judges to process motions.<sup>112</sup>

In sum, judges close cases and decide motions in substantially greater numbers in List weeks than at other times. These findings are causally robust. If August-filed motions are essentially identical to September-filed motions, as we believe, we can attribute any difference in closure volumes between the two groups to the treatment itself, rather than any other factor—including motion age and any other unknown and unobservable variables. This spike is not just a function of the calendar: motions that are ineligible to appear on a given List exhibit no spike, while those that are eligible do. The List is apparently so substantial a driver of judicial behavior that some motions that cannot be resolved before their first deadline are “mothballed” until just before the next one approaches.

#### B. Duration: Does the List Speed up Case and Motion Processing?

From the first, the Six-Month List was designed to pressure judges to process cases and motions faster than they otherwise would. Does it actually do so? The evidence is equivocal.

We find weak evidence that the three-year, case-level deadline speeds up the processing of a small fraction of cases by a small amount. By contrast, the six-month *motion*-level deadline does reduce the duration of some motions. Compared to motions subject to a one-year deadline, motions subject to a six-month deadline take on average about two weeks (roughly 8%) less to be processed. This effect is not uniform, however. As one might expect, the List does not shorten the duration of the shortest or longest motions—its effects are concentrated exclusively on motions that would take roughly six months to handle in its absence. Moreover, those August-filed motions that survive their first deadline in March apparently take *longer*

---

<sup>111</sup> For example, a motion that was filed on August 15, 2011, and was disposed of on March 28, 2012, had a duration of 226 days. We simply shift both the start and end date of that motion so that it nominally begins on August 30, 2011, and ends 226 days later on April 12, 2012.

<sup>112</sup> Regression results are available from the authors.



to process of than they would in the absence of the List because of “secondary bunching.”

We also find that the motion-level effects carry over, albeit imperfectly, to the case-level. On average, cases with a summary judgment motion filed in August last about forty-one days less than those in which a summary judgment motion is filed in September. The quasi-random assignment to treatment and control group means that this difference cannot be attributed to unobservable confounding variables such as lawyer skill or size of stakes. However, the reduction in case processing time occurs only for those cases that are resolved by the motion. If the motion is not dispositive, the time to resolve the case is unchanged.

1. Case-Level Effects

a. “Durational Bunching”

Cases closing in List weeks are substantially older at “death” than those closing during the rest of the year (by about fifty days for the mean and sixty-five days for the median), but only in the period after 1991. Figures 5A and 5B illustrate this finding, plotting the mean (Figure 5A) and median (Figure 5B) duration of cases by week of closure.

FIGURE 5A: MEAN DURATION OF TERMINATED CASES BY CALENDAR WEEK, BEFORE (1980–1990) AND AFTER (1991–2016) CJRA FOR CASES LESS THAN 1,500 DAYS

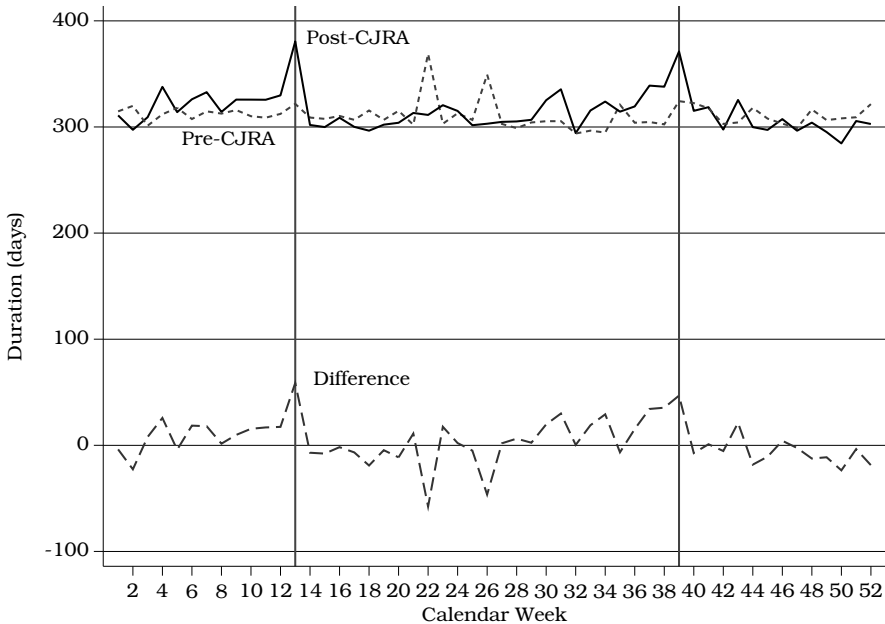
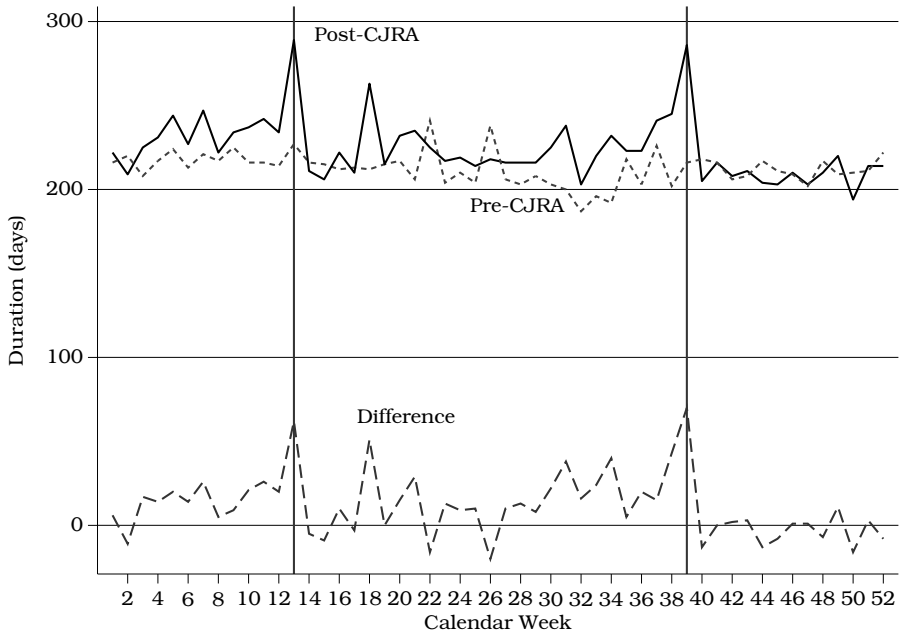


FIGURE 5B: MEDIAN DURATION OF TERMINATED CASES BY CALENDAR WEEK, BEFORE (1980–1990) AND AFTER (1991–2016) CJRA FOR CASES LESS THAN 1,500 DAYS



This is the exactly the pattern we would predict if judges make special efforts to close eligible cases in List weeks. By definition, eligible cases are older than ineligible ones. So when the share of closures that are List-eligible increases, the average age of closed cases must rise.<sup>113</sup>

This finding demonstrates the power of the List. It is hard to think of any other plausible explanation for why List week closures should be substantially older than those at other

<sup>113</sup> Assume that during List weeks in the post-CJRA period, judges pay more attention than usual to closing cases that would appear on the List if not closed. (We focus on cases here, but the same analysis would apply if judges are concentrating on disposing of List-eligible motions.) By definition, eligible cases are those that are more than three years old as of a List date. The average age of these List-eligible cases is necessarily greater than for noneligible cases, since the minimum age among List-eligible cases is greater than the maximum age among cases that are ineligible for the List. Let  $D_t$  be the average age of the List-eligible cases closing at time  $t$ ,  $d_t$  be the average age of the ineligible cases, and  $e_t$  be the proportion of cases closing in week  $t$  that are List-eligible. Then the average duration of all cases closing in week  $t$  is just the weighted sum  $A_t = e_t D + (1 - e_t) d$ .  $A_t$  is obviously increasing in  $e_t$ . (Technically,  $dA/de = D - d > 0$ .) Thus, in weeks when judges decide a greater proportion of List-eligible cases, the average age of closures must go up. The duration spikes (and their pre-post CJRA differences) are all statistically significant in a regression format. Results available from the authors.

times of the year, except that judges are selecting older cases to close at those times.

Unfortunately, knowing that List week closures are older than closures at other times does not shed any light on whether the List actually lowers the duration of cases overall. This is not a straightforward question to answer because we do not know what would have happened to average or median duration in the absence of the List. One might be tempted to use a simple comparison of case durations before and after passage of the CJRA to measure the effects of the List. But that comparison would be misleading: both mean and median case durations have increased since 1991 (by about 16% and 15%, respectively). It would be naive to attribute this increase to the passage of the Act, since many other factors were presumably operating in the background. Accordingly, the before-after comparison does not plausibly isolate the true causal effects of the List. We are able to solve this problem to some extent by exploiting filing dates, as described below.

b. *Using Filing Dates to Test Duration Reduction*

To better address the causation issue, we exploit the timing of case filings, using the quasi-random assignment technique we previously discussed in the Methods section and applied to summary judgment motions. We assume that cases filed in the last week of September (or March) are, on average, no different from those filed in the first week of October (or April). That is, we treat filing date as essentially random across these small intervals of time. We can then think of the two groups of cases (late March versus early April, or late September versus early October) as being like plants that are randomly assigned to receive different amounts of fertilizer. Here, the two groups randomly receive different *deadlines*: the “treatment” group becomes List-eligible after three years, while the “control” group’s deadline is six months longer.<sup>114</sup> Given this technique, we can plausibly assess the causal influence of the List by comparing the duration of cases filed in late March with those filed in early April (or late September versus early October). If facing a tighter List deadline has an effect, the “treated” cases should have shorter durations than those in the control group.

We note that the three-year case-level deadline is a priori unlikely to be effective at shortening overall case durations, for

---

<sup>114</sup> That is, a case filed on March 30, 2012, would be eligible for the March 31, 2015 List, while a case filed three days later, on April 2, 2012, is not List-eligible until September 30, 2012, 183 days later.

several reasons. First, the intervention is not very powerful: the treatment group's deadline is only 14% shorter than that of the control group, so we would not expect it to have a large effect on case duration.<sup>115</sup> Second, the three-year deadline is irrelevant to most cases: 92% of them close in fewer than two years, so the three-year constraint is not even close to binding for this group. Third, the cases that last for three years or more are probably unusually complex, and thus difficult for a judge to close even if she wanted to do so. Fourth, our discussions with court personnel suggest that judges do not take the case-level List nearly as seriously as they do the motions-level List, in part for the reasons just mentioned. So the peer pressure effects exerted by the three-year deadline are probably attenuated as well.

With these caveats in mind, Table 3 and Figure 6 present our analytic results.<sup>116</sup> Each pair of bars in Figures 6A and 6B compares mean or median durations for different samples of treated and untreated cases. The difference between the treatment and control groups is the treatment effect—the reduction in duration attributable to the three-year case-level deadline.

---

<sup>115</sup> The treatment group's deadline is thirty-six months versus forty-two months for the control group.

<sup>116</sup> We also used this same technique in a regression framework in which we measured the effects of the treatment on case duration, after controlling for three-digit nature of suit, circuit, and time trends. The results (available on request from the authors) are virtually identical to those presented here.

TABLE 3: DURATION IN DAYS FOR CASES, FOR TREATMENT AND CONTROL GROUPS, USING VARIOUS SAMPLE DEFINITIONS: 1980–2013

	<b>Treatment</b>	<b>Control</b>	<b>Difference</b>	<b>% Difference</b>
1. Full Sample				
N	318,929	316,737		
Mean	356.8	347.3	9.5	2.7%
SD	404.7	396.4		
Median	227	221	6.0	2.6%
2. Drop Bankruptcy & Social Security				
N	294,153	292,425		
Mean	357.6	347.3	10.3	2.9%
SD	415.0	406.2		3.2%
Median	218	211	7.0	
3. & Drop Duration > 4 years				
N	285,752	285,702		
Mean	310.5	308.4	2.1	0.7%
SD	301.4	305.8		
Median	209	204	5.0	2.4%
4. & Drop Filed Before 1991				
N	196,793	200,147		
Mean	308.3	304.1	4.2	1.4%
SD	299.1	304.3		
Median	210	203	7.0	3.3%
5. & Drop Duration < 2.5 years				
N	12,178	12,179		
Mean	1,125.2	1,166.5	-41.3	-3.7%
SD	149.4	155.6		
Median	1,094	1,158	-64.0	-5.5%
<p>Note: Initially, the treatment group consists of all cases filed in the last week of March or September of any year between 1980 and 2013 inclusive. The control group is all cases filed in the first week of April or October over the same period. Restrictions in each panel apply to all subsequent panels. All differences in means are statistically significant at the 0.01 level using a 2-sample t-test. All differences in medians are statistically significant at the 0.01 level using a chi-squared test</p>				

FIGURE 6A: MEAN DURATION OF CASES IN TREATMENT AND CONTROL GROUPS, FOR VARIOUS SAMPLE DEFINITIONS

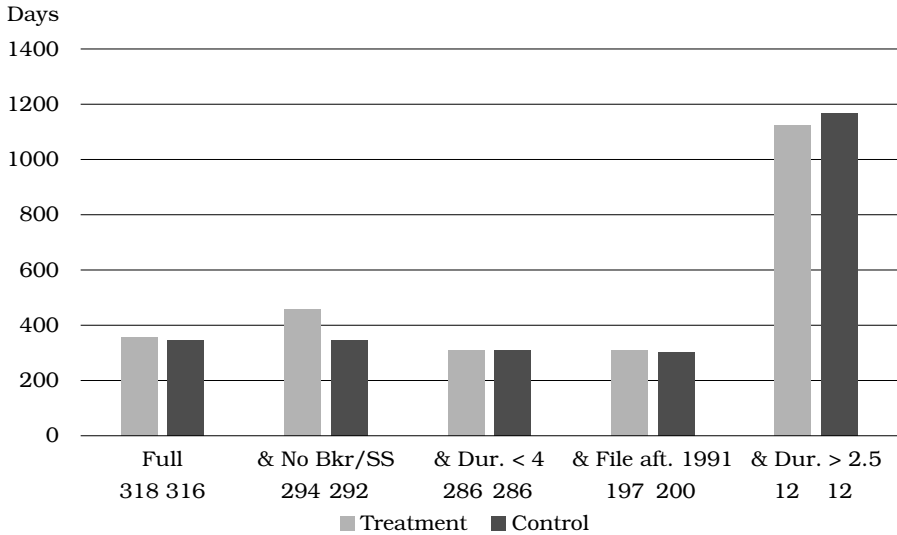
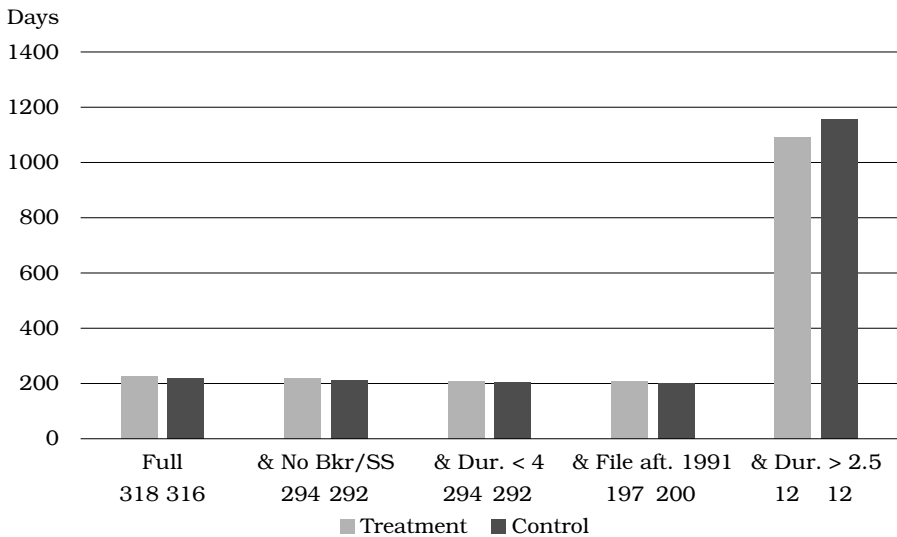


FIGURE 6B: MEDIAN DURATION OF CASES IN TREATMENT AND CONTROL GROUPS, FOR VARIOUS SAMPLE DEFINITIONS



Note: Numbers on the x-axis are sample sizes (in thousands) for the Treatment and Control groups, respectively. For further description, see notes to previous Table.

As we move to the right in Figures 6A and 6B (or down in Table 3), we narrow the sample to cases that are *a priori* increasingly likely to be influenced by the three-year deadline. For example, the second set of bars drops bankruptcy and

social security cases, which have the same List compilation dates but different timing requirements for eligibility. The third set of bars continues by dropping cases with durations over four years, since there should be no difference between the treatment and control groups for these cases: all are List-eligible. We then drop cases filed before the CJRA went into effect in 1991. And finally, we drop cases with durations of less than 2.5 years, since this group would not likely be affected by a three-year deadline that would not pose a binding constraint on their closure date.

In the full sample, we see that the treated cases are actually six to nine days *longer* than the control cases, about 2.7% for both mean and median durations.<sup>117</sup> As we narrow the sample to cases that seem most likely to be affected by the three-year deadline, we see virtually no change in the difference in duration between the treated and control groups. Only when we focus on post-1991 cases that last more than 2.5 years (but fewer than four years) in the final set of bars do we find a small (forty-one days, or 3.7% to 5.5%) difference in favor of the treated group.

In sum, the three-year deadline may shorten case duration by a small amount for a small group of cases, constituting less than 4% of all cases filed. There is no statistically detectable effect for most cases.

## 2. *Motions*

In contrast with case duration, we conclude that the List does reduce the time for the processing of motions, on average by twenty to thirty days.

### a. *Duration-Reducing Effects of the List*

Our analysis shows that on average, the List deadline for motions reduces the duration of motions filed in August relative to those filed in September. In the treatment group, the mean of time to disposition of the motion is 172.3 days whereas the comparable figure for the control group is 188.2 days. Although that difference is not statistically significant at conventional levels, it is close. When twelve outlier motions longer than two years are excluded, the difference between the treatment and control group becomes statistically significant and decreases from 15.9 to 13.0 days. This difference is substan-

---

<sup>117</sup> Because of the large sample sizes involved, this and all subsequent differences in both means and medians are statistically significant at the 1% level.

tively meaningful; treatment group motions last 157.9 days compared to a mean duration of 170.9 days in the control group.

To begin with, as illustrated in Table 4, when we use our entire data set of 758 motions (panel A), the effect of the List on motion duration is borderline in terms of reaching conventional levels of statistical significance ( $p = 0.11$ ). This data includes twelve outlier motions (or 1.6% of our sample), however, and dropping them is appropriate, since a summary judgment motion which takes longer than two years to decide is a clear outlier. When we do so (panel B), our analysis shows a consistent shortening of motion durations in the treatment group relative to the control group at most bandwidths. Table D(2) in the Appendix demonstrates this pattern. Depending on the specification, the effect size ranges from twenty to thirty extra days for the control group relative to the treatment group. In other words, a motion filed in late August is likely to be decided twenty to thirty days faster than one filed in early September, even if the filings are only days apart. While this is a wide range, even the bottom end of the range is substantively significant, given an average disposition time of about 160 days.

TABLE 4: THE EFFECT OF THE LIST ON MOTION DURATION:  
DURATION OF MOTIONS FILED IN AUGUST AND  
SEPTEMBER, 2011, IN DAYS

A. All motions					
	Number of Obs.	Mean Duration	Std. Dev.	Min	Max
Control Group (Sept. Filed)	367	188.2	189.5	0	1,609
Treatment Group (Aug. Filed)	391	172.3	169.2	0	1,789
Difference		15.9 days			
		$p = 0.112$			
B. Duration $\leq 2$ Years Only					
Control Group (Sept. Filed)	361	170.9	131.9	0	685
Treatment Group (Aug. Filed)	385	157.9	117.3	0	693
Difference		13.0 days			
		$p = 0.079^*$			

Note: Reported p-values are from a one-tailed test. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .



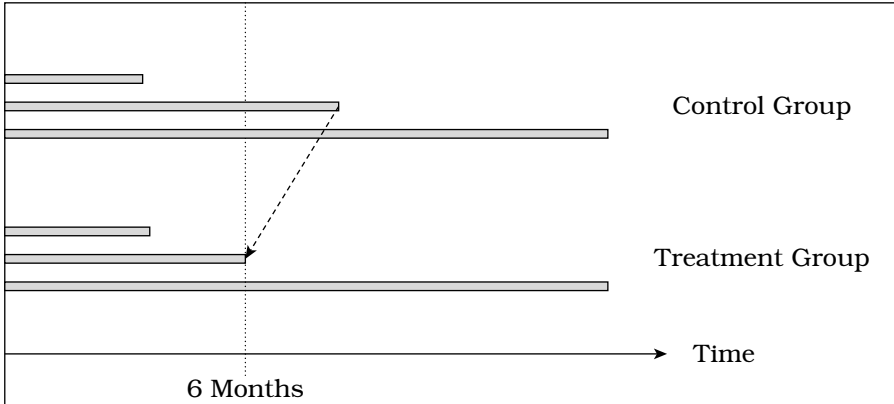
b. Which *Motions are Processed More Rapidly?*

Our theory predicts more than just that the List's motion deadline should *generally* speed up motion processing. Rather, its effects should not be uniform across all motions. For example, a six-month deadline should not affect the handling of a motion that would take almost no time to resolve. It would also be surprising if the desire to avoid getting on the List could induce a judge to cut five months (45%) from the processing time for a motion that would otherwise have taken eleven months to decide. So substantial a time saving is probably infeasible. Instead, we expect the List to have its greatest influence on motions that would otherwise take roughly 214 days to process in its absence. These are the motions that could feasibly be sped up so as to close them before the List is compiled, and it is on these motions that a judge eager to avoid the List would want to focus his or her efforts.

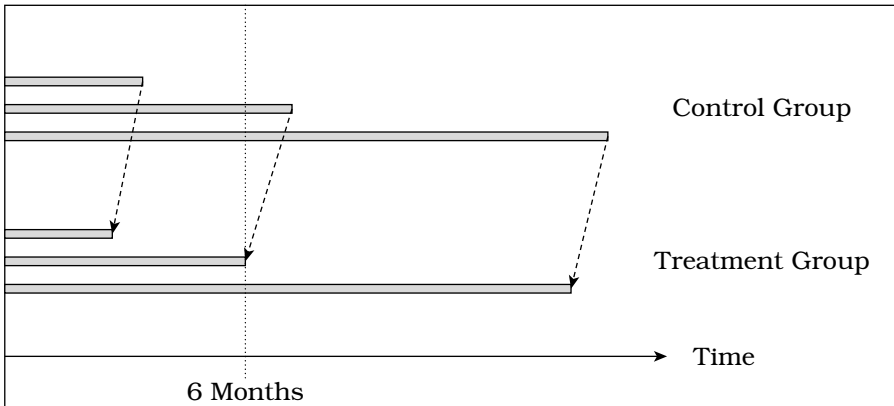
Figure 7 illustrates this argument, using a simplified hypothetical group of motions. In both panels, the treatment group motions have a shorter *median* duration than the control group. But the effects of the treatment are not distributed in the same way in the two panels. In the first, the longest and shortest motions are unaffected by the List; it is only the middle motion that is shortened, from just over to just under six months. In the second panel, by contrast, *all* motions are uniformly shortened by a small amount. Panel A is consistent with our theory of how the List should operate; panel B is not.

FIGURE 7: HYPOTHETICAL LIST EFFECTS FOR HETEROGENEOUS VS. UNIFORM DURATION REDUCTION

A. REDUCTION IN DURATION FOR THE MEDIAN MOTION, BUT NO REDUCTION FOR OTHER DURATIONS.



B. REDUCTION IN DURATION FOR ALL DURATIONS



In Figure 8, we look not at the effect of the List on the duration of the *average* motion, but instead at various percentiles of the duration distribution, plotting the empirical cumulative distribution function for the treatment and control groups.<sup>118</sup> The horizontal difference between the two curves is the treatment effect we are trying to assess.

<sup>118</sup> Despite its intimidating name, the ECDF is conceptually simple: it is just the share of all motions that have a duration that is less than or equal to a given length. So, for example, in our sample, roughly one half of all motions have a duration of 140 days or less, as indicated in Figure 8.

FIGURE 8: EMPIRICAL CUMULATIVE DISTRIBUTION FUNCTION FOR MOTION DURATION, BY TREATMENT VS. CONTROL (MOTIONS LESS THAN OR EQUAL TO 2 YEARS)

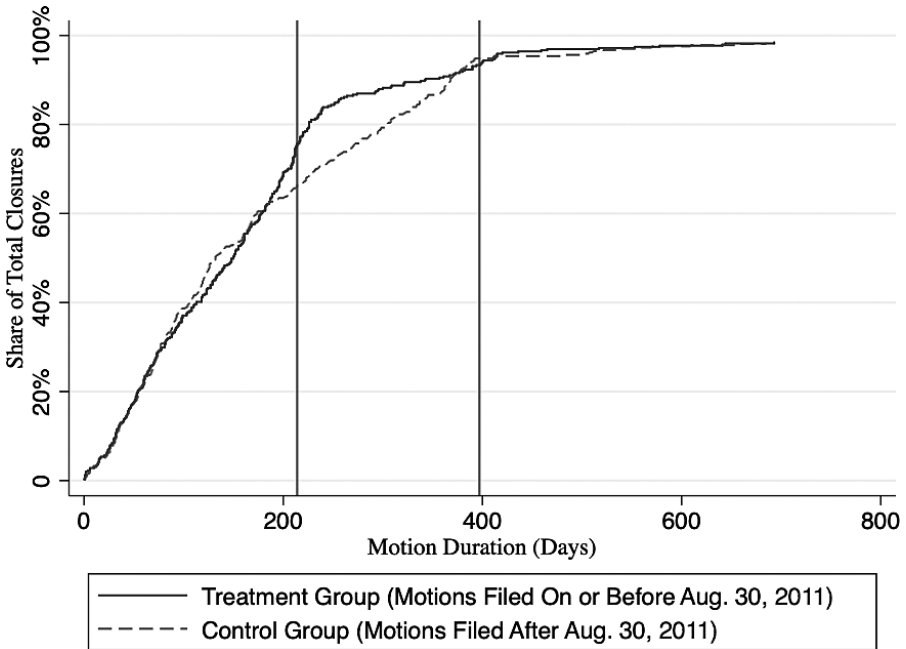


Figure 8 strongly supports our theory of how the List works. The duration of short-duration motions looks virtually identical, whether they are from the control or treatment group. The reduction in motion-processing time caused by the List first becomes evident for motions that take roughly 175 days.<sup>119</sup> These are the “Goldilocks” motions that are just long enough that the List deadline might be relevant, yet just short enough that they could plausibly have been “sped up” so as to avoid making it onto the List. And after diverging at around 175 days, the two curves in Panel B begin to converge again after about 275 days. Assuming that judges cannot speed up the processing of a motion by more than a few months, this makes sense. Any motion that would take nine months or longer to process in the *absence* of the List cannot be shortened by enough to put it on the “good” side of the 214-day deadline. And the distinction between treatment and control groups vanishes after 397 days (September 30, 2012), as it should, since all surviving motions are eligible for all subsequent Lists after that point.

<sup>119</sup> Using a Kolmogorov-Smirnov test, the two distributions are statistically distinguishable from one another at the  $p = 0.001$  level.

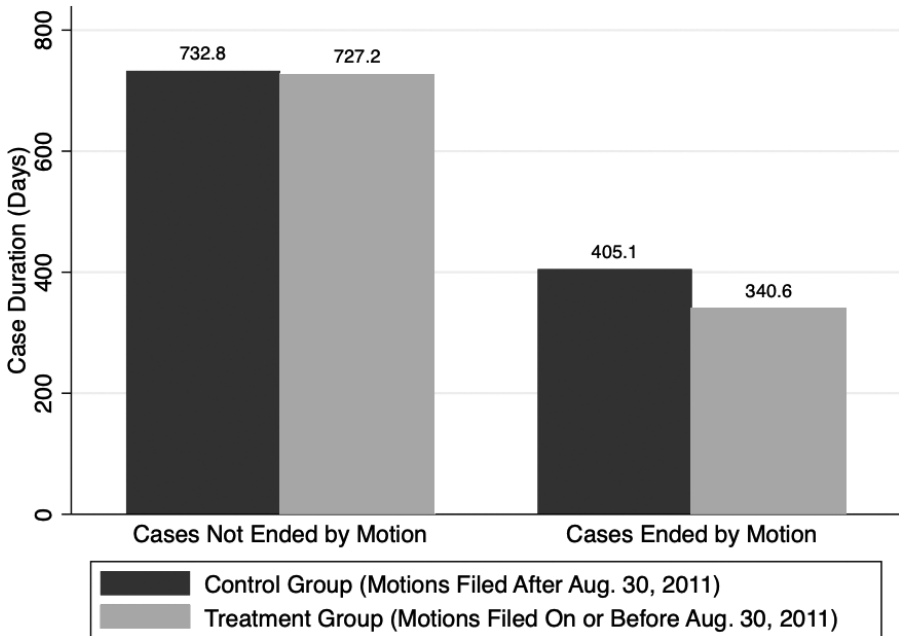
c. *Does Reducing Motion Processing Time Speed up Case Resolution Overall?*

Does processing a motion faster actually reduce the time it takes to close the *case* in which that motion was filed? One possibility is that even if the List cuts motion processing time by twenty to thirty days, any time saved is subsequently dissipated by the parties or the judge, leaving no overall reduction in case-level duration. Alternatively, perhaps the speed-up in deciding the summary judgment motion shifts the timing of all subsequent decisions by that same amount, so that none of the time saved is crowded-out by other activities and the case closes a full twenty to thirty days earlier.

In fact, we find that the six-month motions-level deadline *does* speed up case processing, but only for a subset of the cases. The average case in the treated group (summary judgment motion filed in August) has a duration of 470 days (median of 371 days). By contrast, the average case in which a summary judgment motion is filed in September lasts for 505 days (median of 438 days). So the average *case* is shortened by about thirty-five days as a result of the six-month *motion* deadline. (The median case is shortened by almost twice as long, sixty-seven days.)

But as we have seen elsewhere, the average figure conceals considerable heterogeneity. Figure 9 illustrates. If we divide the cases into those that end with the resolution of the summary judgment motion and those that do not, we find that essentially all of the savings in processing time occurs for the first group. Control group cases in which the summary judgment motion is granted or the case is dismissed (thus disposing of the entire case) last an average of 405 days; the similar figure for the treatment group is 341 days, a 64-day difference. Of course, the cases that are *not* resolved on the summary judgment motion last much longer. But there is virtually no difference between the treatment and control group for these cases (733 days for the control and 727 days for the treatment group, a difference of only six days or 0.8 percent).

FIGURE 9: CASE DURATIONS IN MOTIONS DATASET,  
BY METHOD OF CASE DISPOSITION



Note: Cases ended by motion means a disposition granting summary judgment or dismissing the motion took place and the case terminated within 10 days of the decision. We obtain similar results when examining whether the case terminates on the same day or within 3 days of the decision. Results are available from the authors.

To briefly summarize: we find that the three-year case-level deadline has virtually no effect on case processing time. The six-month motions-level deadline does reduce the length of time that it takes to process a motion: that is, motions filed in August are resolved about twenty to thirty days faster than those filed in September. And finally, the time saved in *motion*-processing translates into a shorter *case*-processing time, but only for those cases for which the summary judgment motion is dispositive. Cases that do not end when the summary judgment motion is resolved have the same duration, regardless of whether the motion was subject to a 214-day or 397-day deadline.

#### d. Generalizing the Results

Because List eligibility is based on the date a motion is filed, some motions are exposed to shorter deadlines than others. For example, the average August-filed motion faces a 228-day deadline, while the average motion filed in September

faces a 382-day deadline.<sup>120</sup> Our estimated List effect is based on this difference in exposure to deadline pressures, but how does it generalize to motions filed at other times of the year?

To answer that question, consider Figure 10, which plots the distance to the nearest relevant deadline for every day of the year (dashed line, left axis), and the fraction of all motions that last longer than the relevant deadline (solid line, right axis). There are sharp spikes in the deadline on March 1 and August 30, corresponding to the statutorily required 214-day offset from the next List date. The fraction of all motions that take longer than the deadline—a measure of the List's stringency for motions filed on that date—exhibits precisely the reverse pattern. Only about 5% of motions take longer than the longest deadline (397 days), while almost 30% of motions have a duration longer than the shortest deadline (214 days).

Importantly, the typical motion (filed on a random day in the year) faces a 306-day deadline. That constraint is clearly not binding for the median motion, which lasts only about 170 days.<sup>121</sup> In fact, only about 16% of the motions in our sample last 300 days or more, and it is only these motions that should be affected by the typical deadline. This suggests as a back-of-the-envelope calculation that the List reduces average duration for all motions (whenever filed) by about eleven to twenty-one days.<sup>122</sup>

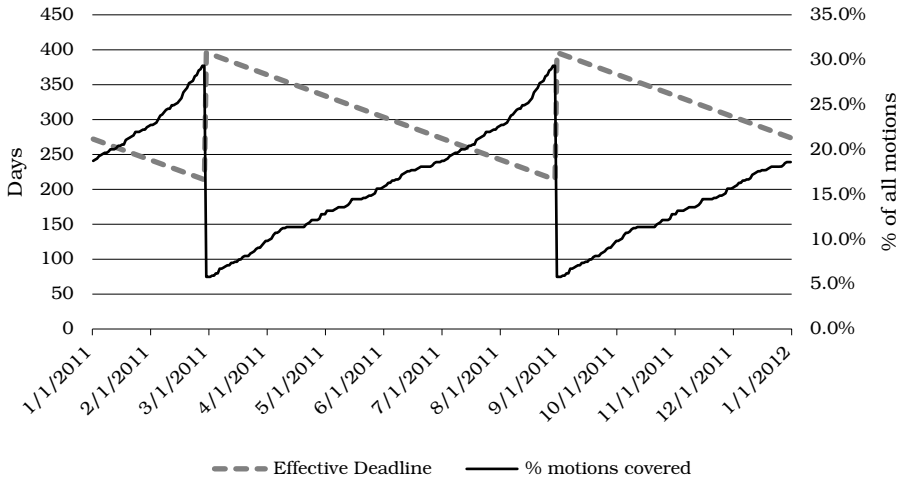
---

<sup>120</sup> The average August-filed motion is filed on August 15, 228 days from the March 31 deadline. The calculation would apply to February-filed motions that are eligible for the September 30 deadline.

<sup>121</sup> Axiomatically, a nonbinding deadline should not speed up processing of a motion that would have been completed before the deadline even in its absence. Giving someone a one-year deadline to take out the garbage should not make them work faster if they would ordinarily take only a week to accomplish this task—the constraint imposed by the deadline is irrelevant.

<sup>122</sup> If a 228-day deadline constrains 30% of all motions and reduces duration by twenty to forty days, then a 300-day deadline that constrains only 16% of all motions should reduce duration by roughly eleven to twenty-one days. That is  $16\%/30\% = x/20$  implies  $x = 11$ .

FIGURE 10: EFFECTIVE DEADLINE (DISTANCE TO NEXT LIST DATE) FOR MOTIONS, AND PROPORTION OF ALL MOTIONS WHOSE DURATION EXCEEDS THAT DEADLINE, BY DAY OF FILING



Note: Left axis shows the time between filing date and the nearest relevant List date. The right axis shows the share of all motions that take longer than this.

### C. Outcomes: Does the List Change Decisions?

Any List effect on outcomes is not only an unintended consequence, but a disturbing one.<sup>123</sup> We begin by noting that our analysis reveals no List effect with respect to case *types*—that is, judges are not deciding certain types of cases more frequently at the deadline. The types of cases that close during List weeks are virtually identical to those closing at other times. For the period as a whole (1980 to 2017) List week closures are distributed across Nature of Suit categories in more or less the same way as cases that close at other times.<sup>124</sup> This is also

<sup>123</sup> As we discuss in the conclusion, some reduction in accuracy might be desirable if it purchased a significant saving in disposition time. But if the List induces systematic *bias*—that is, if it consistently favored one side or the other—that would clearly be inappropriate.

<sup>124</sup> Given the large numbers involved, the data do allow us to reject the hypothesis that the distribution across case types is independent of whether the case closes in a List week vs. Non List week ( $\chi^2(9)=4900$ ,  $p=0.000$ ). But the proportions of each case type differ only in the first decimal place. We also looked at whether the composition of closed cases by basis of jurisdiction (U.S. Plaintiff, U.S. Defendant, Federal Question, Diversity, and Territorial) varied between List- and non-List weeks. We found only minimal differences, either before or after the CJRA went into effect.

The distribution of case types in List- versus Non-List weeks does diverge somewhat after the CJRA passed, however. We capture dissimilarity of distributions of case types using the Duncan index. (The Duncan index of similarity between List and Non-List weeks is computed as  $1/2\sum_i |S_{in} - S_{il}|$ , where  $S_{in} = N_{in}/N_n$

true at the motion level, although our sample size is too small to be entirely confident about the distribution of motions.

At the case level, we find that the plaintiff win rate drops substantially in List weeks. We also find that the appellate remand rate rises for cases decided at the district court level in List weeks, suggesting that judges may be making more mistakes in response to the List than at other times. Motion outcomes also change due to the List. The treatment group had a greater proportion of motions decided in List weeks that were denials than did the control group. This means that judges are denying more motions for summary judgment in response to the List deadline.

Our quasi-experimental design also allows us to detect some further differences between motions subject to shorter versus longer deadlines. A shorter deadline apparently increases the chance that a defendant-filed summary judgment motion will result in a defendant "win" relative to motions facing a longer deadline.<sup>125</sup> We discuss our findings and what explains them in greater detail below.

### 1. *Case-Level Evidence: The Plaintiff Win Rate*

There is no obvious reason why the Six-Month List should have any effect on case or motion outcomes. None of the theoretical models of win rates of which we are aware predict that the win rate would respond to an externally imposed deadline, and we find it difficult to come up with a compelling story about why this should be true.<sup>126</sup> Nevertheless, the clear pattern of

---

= share of case type  $i$  in total case closures in non-List weeks; and  $S_{il} = N_{il}/N_l$ , the share of case type  $i$  in total case closures during List weeks. The post-CJRA Duncan index of 5% means that 1/20th of all cases would have to be reallocated across case types in order for the List-week distribution to match that of Non-List weeks.) By that measure, List weeks are (slightly) more different from non-List weeks in the post-CJRA period than they were before: the Duncan index increases very modestly from 1.5% to 5.5%. In sum, there are changes in the distribution across case types between List weeks and others, but they are small, and the evidence that certain kinds of cases are being selected for List-week closure is weak.

<sup>125</sup> We define a defendant "win" as either a grant of that party's summary judgment motion or the dismissal of the case. This definition is reasonable, given that a substantial fraction of summary judgment motions are filed as a "motion to dismiss, or in the alternative, for summary judgment," giving the judge a choice of how to resolve them.

<sup>126</sup> See generally Lucian Arye Bebchuk, *Litigation and Settlement Under Imperfect Information*, 15 RAND J. ECON. 404 (1984) (proposing model based on asymmetric information and rational bargaining between parties); George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984) (proposing model of dispute selection based on party optimism and uncertainty about decision standard); Steven Shavell, *Any Frequency of Plaintiff Victory at*



bunching of case closures suggests that it is worth looking at whether the List has a systematic effect on plaintiff win rates (hereinafter, “win rates”).<sup>127</sup> We find that it does, as is clearly discernible in the visual evidence in Figure 10, and the regression results in Table 5.

---

*Trial is Possible*, 25 J. LEGAL STUD. 493 (1996) (demonstrating that plausible models of the win rate can generate a wide range of predictions depending on assumptions about information structures and bargaining processes).

<sup>127</sup> We note that some of the Administrative Office “outcome” data are known to be unreliable. See, e.g., John J. Donohue III & Peter Siegelman, *Law and Macroeconomics: Employment Discrimination Litigation Over the Business Cycle*, 66 S. CAL. L. REV. 709, 760 n.98 (1993) (noting that the AO had incorrectly coded award sizes, resulting in the misrepresentation of certain awards by millions of dollars); Gillian K. Hadfield, *Where Have All the Trials Gone? Settlements, Nontrial Adjudications, and Statistical Artifacts in the Changing Disposition of Federal Civil Cases*, 1 J. EMPIRICAL LEGAL STUD. 705, 713–14 (2004) (noting that problems in disposition codes distort statistical results).

As far as we know, however, the accuracy of the *Judgment For* variable has not been called into question, and indeed, the only study we know of that discusses this issue concludes that the variable is usually entered correctly. See Theodore Eisenberg & Margo Schlanger, *The Reliability of the Administrative Office of the U.S. Courts Database: An Initial Empirical Analysis*, 78 NOTRE DAME L. REV. 1455, 1460 (2003). That study checked AO records against PACER docket sheets for a sample of tort and inmate civil rights cases that closed in Fiscal Year 1993. The authors concluded that “the AO data are very accurate when they report a judgment for plaintiff or defendant.” *Id.* Moreover, it does not seem likely that errors in entering or reporting outcomes would differ between List weeks and other times of the year. And even if they did, it is hard to see why they would systematically favor defendants.

FIGURE 11: PLAINTIFF WIN RATE IN ADJUDICATED CASES,  
BY CALENDAR WEEK, BEFORE AND AFTER CJRA

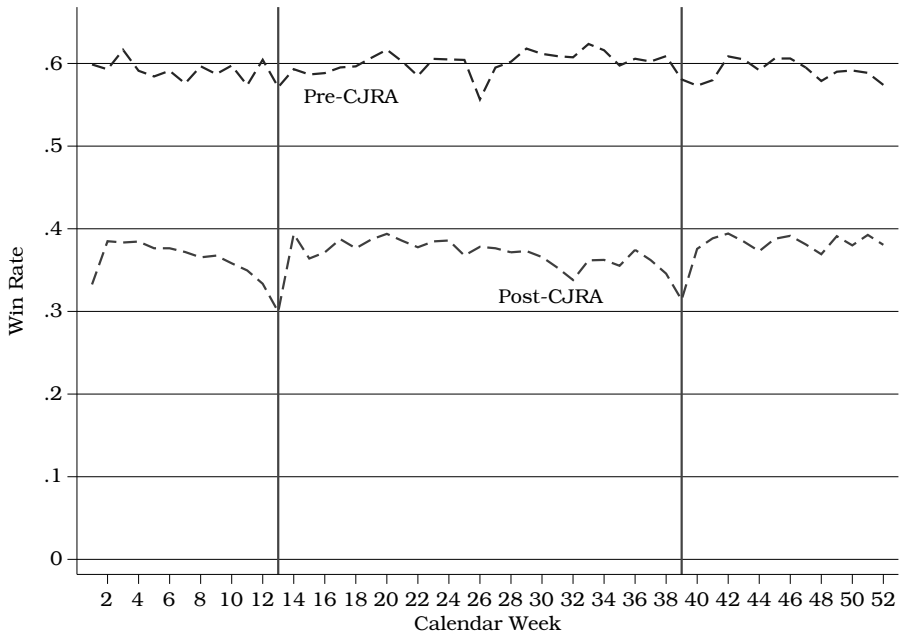


Figure 11 shows a drop in the win rate during List weeks, but only after 1991. In the period after the CJRA was passed (bottom line), the average win rate for non-List weeks is 37.3%. It drops to 30.6% in List weeks (an 18% fall, or more than three standard deviations), and then rebounds completely in the weeks immediately following. Moreover, the win rate stayed virtually constant during List weeks before 1991 (top line),<sup>128</sup> although there was a modest drop at week 26 (the end of the AO fiscal year) in the pre-CJRA era.

Table 5 confirms that these results are statistically significant, even after controlling for the underlying adjudication rate, nature of suit, and time trends.<sup>129</sup> Depending on the specification, post-CJRA List weeks have a 1 to 3 percentage point lower win rate than other weeks.

<sup>128</sup> Another significant fact that emerges from Figure 11 is that win rates were dramatically higher before 1991 than since. We do not discuss this intriguing issue here, but for an extended discussion, see Alexandra D. Lahav & Peter Siegelman, *The Curious Incident of the Falling Win Rate: Individual vs System-Level Justification and the Rule of Law*, 52 U. CAL. DAVIS L. REV. 1371 (2019) (exploring reasons for drop in win rates after 1985).

<sup>129</sup> We measure the adjudication rate as the share of all cases in quarter  $t$  and nature of suit category  $j$  that ultimately end in a judgment for either plaintiff or defendant.

TABLE 5: EXPLAINING THE PROBABILITY OF A PLAINTIFF WIN IN ADJUDICATED CASES (OLS REGRESSIONS)

	(1)	(2)	(3)	(4)
	1980–2017	1980–2017	1980–2017	Filed bef. 2015
Adjudication Rate	0.921*** (0.0019)	-0.025*** (0.0041)	-0.025 (0.0207)	-0.022*** (0.0042)
Time	0.000199*** (6.67e-06)	-6.36e-05*** (5.22e-06)	-6.36e-05*** (1.32e-05)	-5.16e-05*** (5.37e-06)
Time <sup>2</sup>	-4.84e-08*** (1.48e-09)	8.77e-09*** (1.15e-09)	8.77e-09*** (2.62e-09)	5.60e-09*** (1.19e-09)
List Week	-0.0165*** (0.00238)	-0.0032* (0.00171)	-0.0032 (0.00368)	-0.0031* (0.00171)
CJRA Dummy	-0.137*** (0.00147)	-0.037*** (0.00105)	-0.037*** (0.00219)	-0.037*** (0.00105)
CJRA × List Week	-0.0327*** (0.00284)	-0.0234*** (0.00208)	-0.0234*** (0.00455)	-0.0231*** (0.00208)
Constant	0.021*** (0.0064)	0.789*** (0.0063)	0.789*** (0.0190)	0.776*** (0.0065)
Additional Controls?	N	Y	Y	Y
Cluster?	N	N	Y	N
Observations	2,285,355	2,285,355	2,285,355	2,254,371
Adjusted R-squared	0.134	0.558	0.558	0.558

Notes: Dependent variable is 1 if the plaintiff prevails, 0 otherwise. The Adjudication Rate is the filing quarter × nature of suit rate at which cases are adjudicated (listed as won by either plaintiff or defendant). Additional controls are for Circuit, basis of jurisdiction, and 3-digit Nature of Suit. dummies. Robust standard errors in parentheses: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Col. 3: Clusters on Termination Week. Col. 4: Omits cases filed after 2015.

It is possible to further disaggregate List week effects on win rates by nature of suit and by the AO's basis for jurisdiction codes. We do not show these results, however, because we found them largely uninformative. In sum, all four jurisdictional bases (U.S. Plaintiff, Diversity, Federal Question, and U.S. Defendant) experienced List-week dips in the post-CJRA era, although the dips were much stronger for Diversity and Federal Question cases than the other two types.

And among broad nature of suit categories, there were large List week dips for Labor, Tort/Property and Administrative categories, and no dips for Social Security, Prisoner, or Civil Rights cases. All of these dips were present only after 1990.

## 2. Case-Level Evidence: The Remand Rate

Since writing a more thorough or comprehensive opinion presumably takes more time than writing a less-careful one, judges who have no “spare” time inevitably sacrifice decisional

“quality” to achieve a larger volume (quantity) of dispositions; the more matters a judge decides in a given amount of time, the less thorough he or she can be in deciding each. The Six-Month List should make time constraints more pressing—at least at some times of the year—since it encourages judges to work harder at the extensive margin (closing more cases and motions) in the weeks before it is compiled. One might therefore expect decision quality to decline in these periods.<sup>130</sup>

While negative List effects on decisional quality are theoretically plausible, identifying them empirically is challenging because quality is notoriously difficult to measure. The AO data does, however, allow us to determine whether a closed case ever reopens, and if so, for what reason. We suggest that one plausible proxy for decision quality is whether an adjudicated case subsequently reappears on remand from an appellate court.<sup>131</sup> Of course, remands are a function of many factors: the losing party must appeal, the appellant must prevail (at least in part), the court must order a “do-over,” and the case must not settle. It is clear that the remand rate (the probability that a given case will reappear on remand after it has initially closed) is thus a very noisy measure of case quality, but it is the best we can construct from the AO data.<sup>132</sup> The presence of

---

<sup>130</sup> In a companion paper, we develop a simple model of how a judge who discounts future rewards and penalties would rationally respond to periodic assessment of his or her backlog. See Miceli et al., *supra* note 81. Our model predicts both an increase in the volume of cases closed and a decline in decisional “quality” just before a List is compiled. In a somewhat similar vein, Liebman and Mahoney find that when Federal agencies rush to spend remaining funds in the last week of the year, the “quality” of purchases falls. We discuss the problems that arise from rewarding quantity (without comparable incentives for quality) in Part IV. Liebman & Mahoney, *supra* note 76.

<sup>131</sup> To look at remands by date of original decision, we find instances in which the same combination of Judicial District, Office, Docket Number, and Party Name reappear in the data. We then code the *first* appearance of a case as subsequently remanded if the case reappears after it is closed, and its Origin code in its second appearance is given as “Remanded from court of appeals” (AO Origin code “3”). A reversal on appeal would be even more compelling evidence of a low-quality opinion than a remand. And there are roughly twelve times more reversals than remands. See *Table B-5. U.S. Courts of Appeals—Decisions in Cases Terminated on the Merits, by Circuit and Nature of Proceeding, During the 12-Month Period Ending March 31, 2017*, U.S. COURTS, [http://www.uscourts.gov/sites/default/files/data\\_tables/fjcs\\_b5\\_0331.2017.pdf](http://www.uscourts.gov/sites/default/files/data_tables/fjcs_b5_0331.2017.pdf) [<https://perma.cc/D62J-BVAW>] (last visited Nov. 25, 2017). But looking at reversals by date of district court decision would require matching appellate and district court datasets, which is beyond the scope of this paper.

<sup>132</sup> As an alternative, we also considered a broader indicator of disposition quality: whether a case “returns” as either a remand or a “reopening.” There are roughly six times as many reopenings as there are remands on appeal, but there is no evidence of List week peaks in the reopenings data.

intervening variables (such as whether the losing party decides to appeal) will inevitably attenuate any relationship between decisional quality and the remand rate. Any relationship that emerges despite these measurement problems is, a fortiori, all the more compelling.

Table 6 lays out some summary statistics. The volume of remands is extremely low—only about 6,300 (0.27% of all adjudications) for the period 1980 to 2016 as a whole.<sup>133</sup> The overall remand rate was actually slightly higher in the pre-CJRA period than after 1991 (0.29% versus 0.25%). But before 1991, the remand rate was identical between cases adjudicated in List weeks and at other times. After 1990, however, List week closures had a 40% higher chance of being remanded than those closing during the rest of the year (0.35% versus 0.25%). While very small in magnitude, the post-CJRA List week differences are statistically significant ( $\chi^2$  (1 d.f.) = 34.05,  $p = 0.00$ ).

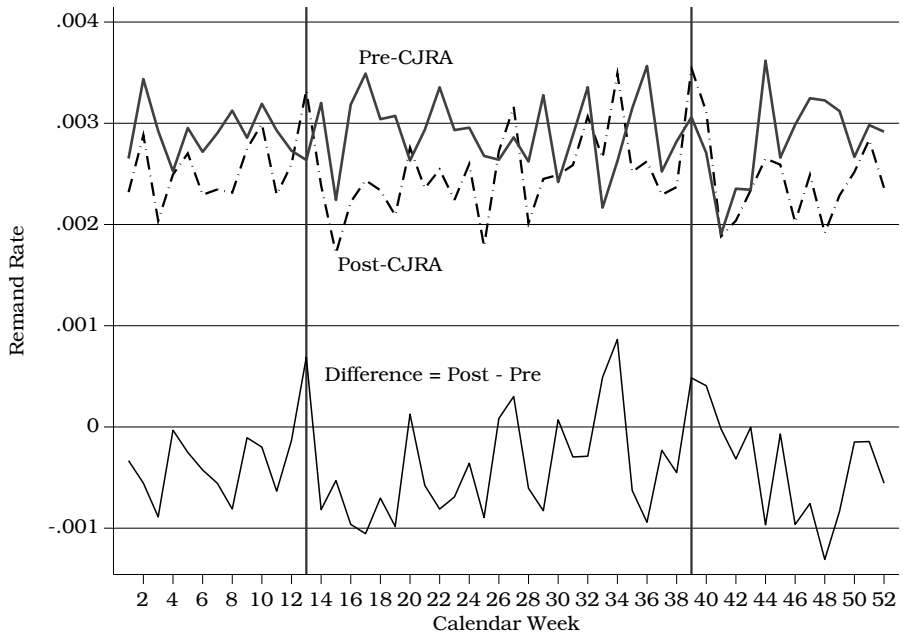
TABLE 6: NUMBER OF REMANDS OF ADJUDICATED CASES CLOSING IN LIST WEEKS VS OTHER WEEKS: 1980–1990 AND 1991–2016

	Pre CJRA: 1980–1990			Post-CJRA: 1991–2016			Row Total
	Non-List Weeks	List Weeks	Total	Non-List Weeks	List Weeks	Total	
Never Remanded	800,460	40,671	841,131	1,416,831	92,320	1,509,151	2,350,282
Ever Remanded	2,343	117	2,460	3,522	322	3,844	6,304
Total	802,803	40,788	843,591	1,420,353	92,642	1,512,995	2,356,586
Remand Rate	0.29%	0.29%	0.29%	0.25%	0.35%	0.25%	0.27%
	$\chi^2(1) = 0.03, p = 0.855$			$\chi^2(1) = 34.05, p = 0.000$			

Figure 12 plots the remand rate by calendar week for the periods before and after the Six-Month List went into effect, as well as the difference in rates (for the same calendar week) between the pre- and post-CJRA periods. Despite a considerable amount of noise, there are obvious spikes at List weeks, confirming the analysis in Table 6. We do not have an explanation for the other apparent spikes (e.g., in week 34).

<sup>133</sup> This seems low, but it is not implausible based on data from the Appellate Courts. For example, in the twelve months ending March 31, 2017, appellate data list only 306 civil cases remanded to district courts. See Table B-5, *U.S. Courts of Appeals—Decisions in Cases Terminated on the Merits, by Circuit and Nature of Proceeding, During the 12-Month Period Ending March 31, 2017*, *supra* note 131. The AO data list 337 remands from appellate decisions during this period, a difference of about 10%. The difference might be accounted for by lags between remands by an appellate court and redocketing by the trial court.

FIGURE 12: REMAND RATE IN ADJUDICATED CASES, BY WEEK OF CLOSURE, 1980–1990 & 1991–2017



Finally, Table 7 provides additional confirmatory evidence. Even after controlling for time trends, and circuit and nature-of-suit fixed effects, there is clear evidence that cases adjudicated in List weeks are more likely to be remanded than those closing at other times. This is only true in the post-CJRA period, however; there are no week 13 or week 39 effects on the remand rate in the period before the List was in effect.

TABLE 7: EXPLAINING REMANDS IN ADJUDICATED CASES:  
OLS REGRESSIONS

	(1)	(2)	(3)	(4)
	1980–2017		1980–1990	1991–2017
Time	-6.09e-07 (8.13e-07)	-6.17e-07 (8.13e-07)	-9.81e-06 (7.61e-06)	1.02e-05*** (1.32e-06)
Time <sup>2</sup>	-1.66e-10 (1.77e-10)	-1.66e-10 (1.77e-10)	2.65e-09 (2.80e-09)	-2.49e-09*** (2.80e-10)
Week 13	0.000457** (0.000216)	-0.000373 (0.000357)	-0.000384 (0.000357)	0.000852*** (0.000269)
Week 39	0.000732*** (0.000230)	0.000162 (0.000398)	0.000149 (0.000398)	0.00101*** (0.000282)
Post-CJRA	-0.000130 (0.000166)	-0.000180 (0.000167)		
Post-CJRA × Week 13		0.00120*** (0.000447)		
Post-CJRA × Week 39		0.000828** (0.000488)		
Constant	0.0039*** (0.00086)	0.0039*** (0.00086)	0.0100** (0.00515)	-0.0078*** (0.00155)
Dummies for Circuit and Nature of Suit	Yes	Yes	Yes	Yes
Observations	2,356,583	2,356,583	843,588	1,512,995
Adjusted R-squared	0.002	0.002	0.003	0.002

Note: The dependent variable is 1 if an adjudicated case that initially terminated in a given week subsequently reappears on remanded (zero otherwise). Week 13 and Week 39 are dummy variables that are 1 if the case (initially) terminated in week 13 or 39 of any year (zero otherwise). Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

We underscore that the effects we observe are very small in size, but they are statistically significant at the 0.01 level.

### 3. *Motion Evidence: Grant/Denial Rates and Other Dispositions*

To examine the impact of the List on motion outcomes, we start by considering whether there are differences in outcomes between the treatment (August-filed) and control (September-filed) motions. A judge has roughly 150 days longer to decide the September motions than the August motions before appearing on the List. Having more time could change motion outcomes if it leads to more accurate determinations. Overall, we find there are no statistically distinguishable differences between the treatment and control groups for grants, denials, partial grants or partial denials, dismissals, settlements, or

withdrawals.<sup>134</sup> However, we do find significant effects for settlement of plaintiff-filed motions and an increase in defendant wins (that is the combined categories of dismissals and grants in defendant-filed summary judgment motions).

Proximity to the List deadline could also affect motion dispositions, so we test whether there are timing effects of the List. For example, given the bunching of decisions near List dates that we have already identified, a higher volume of dispositions at these times might lead to rushed decision-making and more errors.

#### a. *Overall Results*

We do not have sufficient statistical power to look at all the dispositions in our data,<sup>135</sup> so we focus on grants and denials, for which there are enough observations.

Figures 13A and 13B plot the frequency of granted and denied summary judgment motions by List week. Immediately prior to the March 31 List (weeks 10 to 12), the grant rate (that is, grants as a proportion of all motion dispositions, pooling the treatment and control groups) ranges from 69.2% to 81.8%. But in week 13, the grant rate drops precipitously to only 43.5% (pooling the treatment and control groups). Conversely, the denial rate ranges from 15.0% to 36.4% from weeks 10 to 12, and then climbs to 47.8% in week 13 for the treatment and control groups.<sup>136</sup>

---

<sup>134</sup> One case in our sample was transferred. As a result, we exclude that disposition from our analysis.

<sup>135</sup> The possible outcomes we recorded are: grant, denial, partial grant/denial, dismissal on other grounds, settlement, withdrawal. Unlike the case-level AO data, there is no official code for who “won” a motion. We think it reasonable to define a defendant win as a (defendant-filed) motion that is granted or a dismissal of the case on grounds other than the disposition of the summary judgment motion.

<sup>136</sup> The trends are even more striking when one examines the treatment group exclusively, with the grant rate ranging from 63.6% to 77.8% in weeks 10 to 12 but dropping to 33.3% in week 13. The denial rate for the treatment group ranges from 6.3% to 13.6% in weeks 10 to 12, and spikes to 25.0% in week 13.



FIGURE 13A: SUMMARY JUDGMENT GRANTS AND DENIALS, BY CALENDAR WEEK TREATMENT GROUP

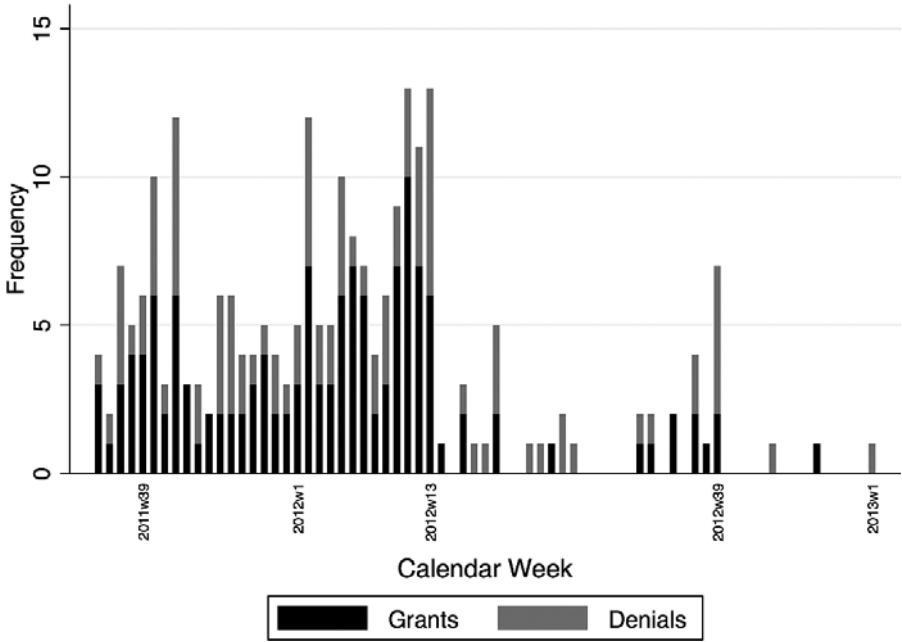
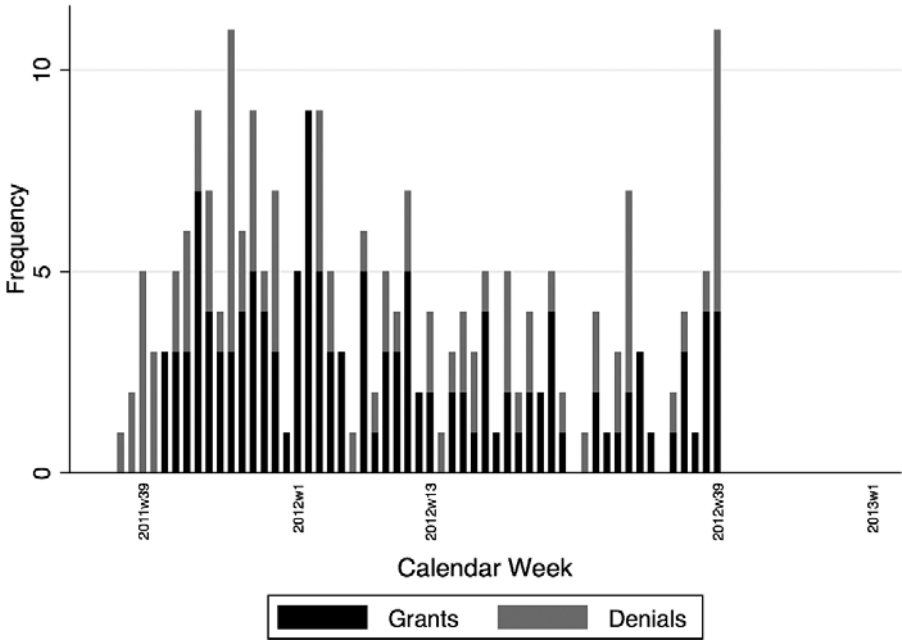


FIGURE 13B: SUMMARY JUDGMENT GRANTS AND DENIALS, BY CALENDAR WEEK CONTROL GROUP



Our regression analysis confirms that the List has an effect on grants, denials, and dismissals. We compared these out-

comes in List weeks with non-List weeks within the treatment and control groups, as well as comparing the effect across the groups. We found an increase in grants of summary judgment motions in the weeks leading up to the List weeks in both treatment and control groups. Table 8 shows an increase in the treatment group in weeks 11 to 13 that ranged from about four to five additional granted motions relative to other List-eligible weeks.

TABLE 8: THE EFFECT OF THE LIST ON MOTIONS GRANTED:  
OLS REGRESSIONS

	(1) Treatment	(2) Control	(3) Difference
Time	-0.167 (0.115)	-0.089 (0.084)	-0.086 (0.062)
Time <sup>2</sup>	2.87e-05 (2.02e-05)	1.49e-05 (1.48e-05)	1.52e-05 (1.08e-05)
Week 11	4.966*** (0.367)	2.918*** (0.274)	3.176*** (0.141)
Week 12	3.521*** (0.571)	0.644 (0.459)	4.057*** (0.238)
Week 13	3.822*** (0.686)	1.301** (0.604)	3.279*** (0.406)
Week 37	1.906*** (0.114)	-1.041*** (0.238)	2.321*** (0.088)
Week 38	0.877*** (0.176)	2.698*** (0.400)	-2.425*** (0.114)
Week 39	1.907*** (0.207)	3.342*** (0.529)	-1.779*** (0.110)
Constant	242.2* (163.4)	133.4 (119.7)	120.6 (87.6)
Observations	281	281	281
R-squared	0.222	0.176	0.137

Note: The dependent variable is the number of motions granted in a week. Week 11, 12, 13, 37, 38 and 39 are dummy variables that are 1 if the motion was granted in those weeks of 2012. All regressions use Prais-Winsten correction for autocorrelation. Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

A similar uptick of some one to two grants per week took place in the treatment group weeks prior to the week 39 deadline. Although the control group experienced a drop of approximately one grant in week 37, as expected, it had an increase grants (approximately three) that was higher relative to the

treatment group in weeks 38 and 39. When we examine denials in Table 9, we find an increase in denials during weeks 11 and 12 (an estimated respective increase of 1.87 and 3.23 denials in those weeks), but the estimate jumps to 6.5 denials in the List week. Although the control group has small movements in denials ranging from -0.666 in week 37 and 0.481 in week 38, we see a similar uptick to 6.6 denials in week 39, the last week prior to the deadline. The trends suggest that judges are clearing more labor-intensive grants prior to working on denials, which spike at the List week. The differences within the treatment and control group in terms of the numbers of grants and denials substantiate these underlying trends, and are statistically significant at conventional levels.

TABLE 9: THE EFFECT OF THE LIST ON MOTIONS DENIED:  
OLS REGRESSIONS

	(1)	(2)	(3)
	Treatment	Control	Difference
Time	-0.118** (0.0596)	-0.093* (0.0528)	-0.025 (0.0341)
Time <sup>2</sup>	2.02e-05** (1.05e-05)	1.59e-05 (9.28e-06)	4.38e-06 (5.98e-06)
Week 11	1.865*** (0.112)	1.192*** (0.111)	1.070*** (0.132)
Week 12	3.227*** (0.172)	-0.761*** (0.142)	3.993*** (0.120)
Week 13	6.502*** (0.316)	1.198*** (0.168)	4.914*** (0.168)
Week 37	1.688*** (0.0746)	-0.666*** (0.0817)	1.930*** (0.0954)
Week 38	-0.383*** (0.117)	0.481*** (0.121)	-0.985*** (0.086)
Week 39	4.693*** (0.219)	6.636*** (0.223)	-1.990*** (0.093)
Constant	170.8** (84.72)	136.5* (75.12)	34.3 (48.55)
Observations	281	281	281
R-squared	0.377	0.284	0.189

Note: The dependent variable is the number of motions denied in a week. Week 11, 12, 13, 37, 38, and 39 are dummy variables that are 1 if the motion was denied in those weeks in 2012. All regressions use Prais-Winsten correction for autocorrelation. Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Our regressions also show that as the September List date approaches, judges prioritize motions that were already on the March List. Three weeks prior to the September List, judges process motions that were already on the List *before* clearing first-time eligible motions. This suggests that judges may see a marginal penalty for being on the List more than once for the same motion.

Judges also wait to dismiss cases until the List week, delaying dismissal decisions longer than summary judgment decisions.<sup>137</sup> Table 10 shows the treatment group has two more dismissals in week 13 relative to other weeks and almost two more dismissals in week 39. The control group, by contrast, has one fewer dismissal each in weeks 11 and 12, no statistically substantively notable difference in week 13, and almost one more dismissal in week 39. The treatment effect for week 13 is almost two more dismissals for the treatment group and almost one more in week 39. These differences, although statistically significant, are modest in substantive terms.

---

<sup>137</sup> Each of these decisions is spurred by a summary judgment motion being filed. The decisions in our motion dataset that end in a dismissal are those in which the party filed a motion to dismiss or for summary judgment in the alternative and the judge treated the motion as a motion to dismiss.

TABLE 10: THE EFFECT OF THE LIST ON DISMISSALS:  
OLS REGRESSIONS

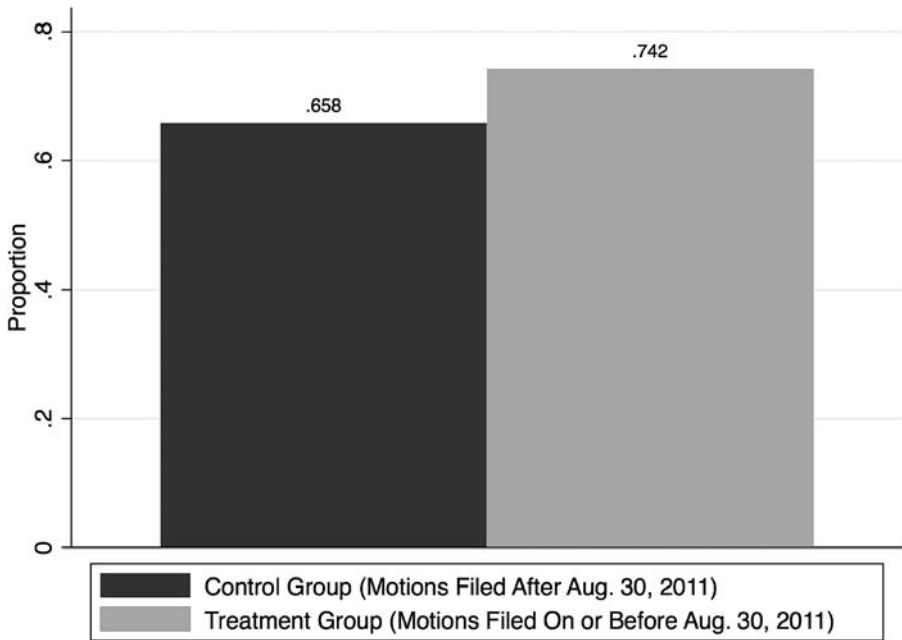
	(1)	(2)	(3)
	Treatment	Control	Difference
Time	-0.185** (0.0869)	-0.099** (0.0500)	-0.097** (0.0393)
Time <sup>2</sup>	3.21e-05** (1.53e-05)	1.71e-05* (8.79e-06)	1.70e-05** (6.91e-06)
Week 11	1.763*** (0.132)	-0.365*** (0.101)	1.766*** (0.124)
Week 12	0.693*** (0.226)	-0.444*** (0.142)	0.781*** (0.117)
Week 13	2.766*** (0.323)	0.641*** (0.238)	1.776*** (0.121)
Week 37	0.840*** (0.0873)	0.734*** (0.0685)	-0.122 (0.0817)
Week 38	0.794*** (0.150)	0.676*** (0.0974)	-0.113 (0.0766)
Week 39	1.843*** (0.215)	0.738*** (0.171)	0.886*** (0.079)
Constant	265.5** (123.8)	143.7** (71.2)	138.2** (56.0)
Observations	281	281	281
R-squared	0.143	0.089	0.089

Note: The dependent variable is the number of dismissals in a week. Week 13 and Week 39 are dummy variables that are 1 if there was a dismissal in week 13 or 39 of 2012. All regressions use Prais-Winsten correction for autocorrelation. Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

b. *Interaction of Treatment and Party Filing Motion*

When we examine outcomes by the party that filed the motion we find that the List has one effect, which is summarized in Figure 14.

FIGURE 14: THE EFFECT OF THE LIST ON GRANTS AND SETTLEMENTS IN DEFENDANT-FILED MOTIONS



Being subject to a shorter deadline has a positive effect on defendant success, as the set of bars in Figure 14 indicates. Summary judgment motions brought by defendants are more likely to be granted (or to end with the case being dismissed) in the treatment group than the control group. There is an increase in combined grants and dismissals for motions brought by defendants from 65.8% in the control group to 74.2% in the treatment group. This increase of 8.4 percentage points is a 12.8% increase in grants plus dismissals,<sup>138</sup> and is consistent with our earlier results at the case level.

We discuss the interpretation of these findings in the next section.

#### 4. Discussion

Three interrelated puzzles emerge from our analysis of outcome effects. First, why are there pronounced drops in plaintiff win rates for cases decided during List weeks? Second, why are there increases in summary judgment denial rates for mo-

<sup>138</sup> We note that when we look at either grants or dismissals separately, we do not find that there is a List effect. But many summary judgment motions in our data are also filed as motions to dismiss in the alternative, and since both of these dispositions constitute a defendant “win,” it is sensible to consider them together.

tions decided in these periods? And finally, why does the List affect (some) outcomes in the treatment group of motions differently from the control group? We discuss each of these issues in turn, noting in advance that we will be speculating, given the absence of any strong evidence.

a. *List Week Drops in Win Rates at the Case Level*

There are two stories one might tell to explain drops in win rates. One focuses on the selection of cases and the timing of decisions: perhaps judges are simply moving cases that would have been defendant wins at other times to List weeks (time shifting). That is, instead of deciding a case in favor of the defendant in April or May, judges move some of those decisions to the last week of March.<sup>139</sup> An alternative is that decisions are not being shifted across time, but that the higher volume of List-week decisions directly causes a drop in the plaintiff win rate—cases are being decided differently, not simply at different times.

The time-shifting story requires a plausible explanation for why judges would want to disproportionately move defendant wins to List weeks. Explicit selection on the basis of which party is going to win strikes us highly unlikely—we find it hard to imagine that judges would deliberately set out to move “defendant wins” to the last week of March or September. More plausibly, judges might choose to shift cases based on something that is correlated with the winning party. For example, judges might choose to put-off deciding “difficult” cases until the List deadline looms. We posit that it takes more effort for a judge to grant a summary judgment motion than to deny one, both because of the burden on the movant and the fact that a grant can be dispositive and presumably requires special justification.<sup>140</sup> As discussed in the next subsection, this is why we think denials of summary judgment motions increase at List weeks. Indeed, we demonstrate in Appendix 2 that a time-shifting model that is consistent with the observed facts requires that judges move “easy” cases to List weeks, rather than difficult ones.<sup>141</sup> Still, moving easy cases to List weeks does not strike us as psychologically compelling: procrastination usu-

---

<sup>139</sup> The effect is to raise the win rate in the last week of March, but to lower win rates at other times, leaving the overall win rate unchanged when measured over a suitably long interval of time.

<sup>140</sup> Interviews with two anonymous federal judges confirm this assumption.

<sup>141</sup> Defendants win roughly 60% of all adjudicated cases. See Table 5, *supra* section III.C.1. “Difficult” cases are, by definition, the 50/50 close calls. Together, these premises mandate that if defendants win a greater share of the List

ally entails delaying the most difficult work, not the easiest decisions.

Alternatively, perhaps win rates drop in List weeks because when judges have to make many decisions over a short period, they are more likely to make mistakes (or otherwise change their behavior). In this account, the List *directly* lowers win rates.

But this explanation faces a major objection: why would more errors disproportionately favor defendants? After all, random errors should not move the win rate at all. One possible story relies on asymmetries in lawyering quality.<sup>142</sup> Suppose that plaintiffs' lawyers are on average of worse quality than defendants' lawyers. When judges have ample time to decide a case, they can afford to make the effort needed to compensate for shortcomings in a plaintiff's brief by doing their own research into the facts and the law. But when trying to decide many matters in a short span of time—as during List weeks—judges cannot spare the time or effort to compensate for plaintiffs' bad briefs with their own research, and are thus more likely to side with defendants. We stress that we have no evidence that directly supports this story, but we think it is at least facially plausible, and it can provide an explanation for why win rates would fall in List weeks.

#### b. *List Week Spikes in Motion Denials*

One explanation for why denials spike during List weeks again centers on time shifting. This story is supported by our finding that starting several weeks before the List is compiled, motion denial rates decrease. They then rise sharply in the last week before the deadline.

Perhaps judges are shifting grants to the second and third weeks of March, saving their denials for the final week before the List is compiled. That could make sense if justifying denials involves less work, and judges know they will have increased workloads right before the deadline. We noted above,

---

week decisions than at other times, it must be because there are more "easy" cases being disposed of at these times, as Appendix 2 demonstrates.

<sup>142</sup> In a qualitative study of briefing quality, Scott Moss finds that plaintiffs' briefs are of lower quality than defendants' briefs in employment litigation. See Scott A. Moss, *(In)competence in Appellate and District Court Brief Writing on Rule 12 and 56 Motions*, 57 N.Y. L. SCH. L. REV. 841, 842 (2012–2013) (summarizing findings); Scott A. Moss, *Bad Briefs, Bad Law, Bad Markets: Documenting the Poor Quality of Plaintiffs' Briefs, Its Impact on the Law, and the Market Failure It Reflects*, 63 EMORY L.J. 59, 63–64 (2013).



however, that this explanation is not psychologically compelling to us.

As with cases, a second possibility is that when judges are working hard to meet the List deadline, they make hasty decisions that disproportionately run in a particular direction, even though they would have reached a different result on some of these motions if they were decided things at another point in time. That is, perhaps judges are denying motions for summary judgment near the List week because this is easier to do quickly even if, upon more reflection, they might have granted the motion. This is consistent with common sense intuition that one is more likely to make mistakes when one is in a rush to meet a deadline, as well as studies of professionals in other contexts.<sup>143</sup>

The main problem with either of these explanations is that they are hard to square with the case-level results. Given that most summary judgment motions are filed by defendants (in our sample, 60%), and List weeks see a spike in *denials*, the plaintiff win rate during List weeks should be *higher*, not lower. Our findings are inconsistent with this logic. Perhaps the cause of the case-level finding lies elsewhere; there are many ways that a case can end other than a motion for summary judgment, and it may be that the case-level win rate drops are to be explained by some other dispositive decision which judges are pressured to make at List weeks.

### c. *Treatment Effects on Motion Grants*

We also find that treated (August-filed) summary judgment motions have a greater chance of being granted than those in the control group (those filed in September). We admit that we lack a compelling explanation for this result. Perhaps judges subject to more stringent deadlines make greater use of rough-and-ready heuristics to decide cases.<sup>144</sup> Under pressure to make more decisions in a shorter time period, judges might rely more heavily on what Daniel Kahneman calls “System 1” (intuitive, automatic, unconscious) reasoning and less on “System 2” (slow, conscious, analytic) reasoning to make decisions.<sup>145</sup> If for some reason System 1 is predisposed to disfavor plaintiffs,

---

<sup>143</sup> See *supra* notes 78 and 79 (describing studies demonstrating reduced quality under deadline pressure).

<sup>144</sup> Cf. DANIEL KAHNEMAN, THINKING, FAST AND SLOW 7 (2011) (explaining how heuristics can cause predictable biases in decision making).

<sup>145</sup> A major theme of Kahneman’s work is not just that we make decisions in two very different ways, but that we often believe we are using System 2 when we are actually using System 1. See *id.* at 13.

then greater reliance on intuitive reasoning could lead to more defendant-favorable decisions.

We are skeptical of this story, however. For one thing, it is hard to see why intuitive reasoning should favor one side over another. Second, it takes more judicial effort to grant a summary judgment motion, thus such a motion is more likely to trigger System 2 thinking.<sup>146</sup> Third, the category of “plaintiffs” include not just tort victims or workers alleging employment discrimination, but also automobile companies suing suppliers for breaches of contract and banks suing customers over loans in default. If intuitive thinking creates biases, one would think that they would cut in favor of sociological categories, not strictly legal ones. And yet, we did not find differences in treatment effects across Nature of Suit categories—all plaintiff types were more-or-less equally disadvantaged in August-filed motions (vis-a-vis those filed in September).

Establishing the List’s direct effects on outcomes is challenging and our evidence is equivocal. Nevertheless, what emerges from piecing together the mosaic of the available clues is a realistic possibility that the List increases defendant’s success rate.<sup>147</sup> This evidence is strong enough to raise concerns about the effect of the List on decisional accuracy, and therefore on justice. First, the shorter deadline faced by the treatment group is linked to a lower plaintiff win rate (measured as the combination of grants of summary judgment and case dismissals). There is no good reason why this should be true. Second, although the causal inference is weaker, there is clear evidence that plaintiff win rates fall substantially in cases that close during List weeks. Third, we find a higher appellate remand rate for List week terminations (albeit from a very low base level), which suggests that cases decided in List weeks are subject to more errors. Finally, there is empirical evidence

---

<sup>146</sup> The reason for this is the standard that the moving party must meet, which is that there is no material fact in dispute and the moving party is entitled to judgment as a matter of law. FED. R. CIV. P. 56.

<sup>147</sup> There is precedent for such an approach. See, e.g., James J. Heckman & Burton Singer, *Abducting Economics*, 107 AM. ECON. REV.: PAPERS & PROC. 298, 298–99 (2017) (“The abductive model for learning from data follows more closely the methods of Sherlock Holmes than those of textbook econometrics. The Sherlock Holmes approach uses many different kinds of clues of varying trustworthiness, weights them, puts them together, and tells a plausible story of the ensemble.”); cf. David E. Pozen, *The Mosaic Theory, National Security, and the Freedom of Information Act*, 115 YALE L.J. 628, 630 (2005) (“The ‘mosaic theory’ describes a basic precept of intelligence gathering: Disparate items of information, though individually of limited or no utility to their possessor, can take on added significance when combined with other items of information.”).

from other social scientific studies that delayed and/or bunched decisions tend to be of poorer average quality; this conclusion also emerges from our theoretical model of the quantity/quality tradeoff engendered by the List.<sup>148</sup> These findings raise concerns that the risk of judicial error or bias increases for decisions made immediately under time pressure from the List.

A heightened risk of error in List weeks is important for considering the trade-off between decisional accuracy and speed, which ought to be central to policymakers. We discuss that trade-off, and potential solutions, in the next Part.

#### IV

#### THE SIX-MONTH LIST AS PUBLIC POLICY: DESIGNING GOOD INCENTIVES

Other things equal, we tend to think that faster disposition of cases and motions is a good thing.<sup>149</sup> But other things are almost certainly not equal. Combining all of our evidence leads us to conclude that the List induces judges to treat similarly situated motions differently; and we suspect that it increases the risk that judges will make errors (especially in List weeks), highlighting the tension between speed and accuracy in adjudication. A simple economic model shows that judges rationally responding to periodic deadlines will increase effort to close cases as the shadow of the deadline begins to loom, but will decrease effort spent on decisional quality.<sup>150</sup>

In this Part, we offer a conjectural cost benefit analysis of the Six-Month List, and conclude that it ought to be abolished. Recognizing that the List is mandated by an act of Congress and is not easily repealed, we go on to suggest an alternative that the courts might adopt until Congress corrects the problem.

---

<sup>148</sup> See *supra* notes 70–76. The Liebman and Mahoney study is particularly compelling in this respect because they have an independent and objective quality measure and show that it falls substantially for projects that are decided-on as deadlines loom.

<sup>149</sup> The analysis of delay is, however, highly complex, and we appreciate that ad hoc theorizing can be hazardous. Valuing faster case-processing is difficult. Moreover, it may be that litigants and society at large care about more than just the average processing time. To that point, it appears that the standard deviation of case duration did not change after the CJRA was passed: it was 438.6 days before and 438.9 days afterwards.

<sup>150</sup> Miceli, Segerson & Siegelman, *supra* note 81 and sources cited in notes 70–76 (showing a quality/quantity tradeoff with deadline pressure in other fields).

### A. Cost Benefit Analysis

Incentive theory has long recognized that speed/quality tradeoffs are a real possibility. A complex task such as resolving litigated cases necessarily has many different dimensions or attributes: we would like cases to be decided quickly, but also fairly and accurately.<sup>151</sup> Theory suggest that when one job component (speed) is relatively easy to measure while another (quality) is unquantifiable, rewarding only what can be measured will distort decision making: actors will respond to incentives by focusing excessively on what is rewarded at the expense of what is not. In such situations, the best approach may actually require abandoning incentives altogether.<sup>152</sup> In the case of judges, it may be better not to reward *anything* than to incentivize speed alone, thereby potentially diverting effort away from quality or other desiderata.<sup>153</sup>

In a world of scarce resources, every policy change creates gains and losses. A normative evaluation of the Six-Month List requires us to assess what benefits it generates and what costs it imposes. Although we do not have hard data on most of the key variables involved, we can use our findings to offer a speculative or “conceptual” cost/benefit analysis, as outlined in Table 11.

---

<sup>151</sup> Zuckerman proposes an accelerated procedure for use in cases where speed is more important (relative to accuracy) than is typical. A.A.S. Zuckerman, *Quality and Economy in Civil Procedure—The Case for Commuting Correct Judgments for Timely Judgments*, 14 OXFORD J. LEGAL STUD. 353 (1994)

<sup>152</sup> Holmstrom, *supra* note 2 (concluding that in settings where quality is not measurable, optimal policy may require no incentive at all). This resonates strongly with Judge Posner’s assessment that judges are not subject to incentives because so much of what we want them to do is unmeasurable. Offering incentives for only the measurable part of the job would likely create distortions that are worse than the problem they were designed to solve. Posner, *supra* note 18.

<sup>153</sup> In the context of Social Security claims, Daniel Ho notes that “[f]ixation on completion rates [case closures as a means of evaluating Administrative Law Judges] can have perverse effects, so it is important for agencies to develop more fine-grained measures of quality.” Ho, *supra* note 69, at 87. In the case of judges, “fine-grained measures of quality” are obviously very difficult to develop.

TABLE 11: COSTS AND BENEFITS OF SIX MONTH LIST DEADLINES<sup>154</sup>

Costs	Benefits
3 Year Case Deadline: None.	3 Year Case Deadline: None.
6 Month Motions Deadline: Administration (Clerk Time, etc.); Error/Bias Risk at List Weeks; Heterogeneity or Unfairness; Diversion from Other Activities.	6 Month Motions deadline: 10–40 days saved in motion processing in 12% of cases.

### 1. Costs

The List potentially imposes four types of costs: (1) administrative costs, (2) risk of error, (3) heterogeneity in judicial administration, and (4) diversion from other activities.

Administrative costs include the time that clerks and court staff have to spend in an effort to comply with the List and to remind judges of the deadlines, as well as the time judges spend reorganizing their schedule to avoid getting on the List. Presumably the variable costs of compiling the List itself are relatively small, since much of this work is embodied in software that has already been written. The volume of judicial decisions spikes during the weeks before a List is compiled, and clerks, staff, and judges tell us that they often work extraordinarily long hours during these periods in an effort to clear their dockets. So there are presumably “overtime” costs as well.<sup>155</sup>

We cannot conclusively demonstrate that the List generates more erroneous decisions. Nevertheless, we think this is a risk that needs to be taken seriously.<sup>156</sup> Outcome differences between some subsets of treatment and control motions are a cause for concern. So are drops in plaintiff win rates in cases closing during List weeks. Almost by definition, rushed work is more likely to be error-prone, as our theoretical model of judicial decision-making suggests.<sup>157</sup> Finally, increases in remand

---

<sup>154</sup> In 2016, cases ending with a “judgment on motion before trial” constituted 11.85 percent of the 276,116 closures recorded in the AO data. For the period 1980–2016, that proportion was 13.8 percent.

<sup>155</sup> If we believe that the marginal cost of effort is increasing (it is more painful to work the 12th hour than the 11th or 10th), then these overtime hours should count as an additional cost of the List, relative to a world where work is spaced-out more evenly.

<sup>156</sup> For a theoretical treatment of these issues, see generally Louis Kaplow, *The Value of Accuracy in Adjudication: An Economic Analysis*, 23 J. LEGAL STUD. 307 (1994) (examining the value of accuracy in adjudication and its tradeoffs).

<sup>157</sup> See *id.*

rates for cases decided in List weeks also raise concerns about decisional accuracy, as does evidence from other contexts.<sup>158</sup> If adjudication errors are relatively small and symmetric, a drop in accuracy might be a price worth paying for faster processing time;<sup>159</sup> if errors are asymmetric (biased in favor of one side), the tradeoff becomes untenable. In any case, we think an open debate about this tradeoff is required. No such discussion occurred when the CJRA was adopted, or since.

Another cost, albeit a subtle one, is the heterogeneity and possible unfairness introduced by the List.<sup>160</sup> As we have pointed out, the effects of the List are highly nonuniform. A motion filed in late August is advantaged relative to one filed only a few days later in early September; a motion that would otherwise take roughly six months to dispose of is advantaged relative to one that would take one or twelve months to process; and cases for which a summary judgment motion is dispositive are advantaged relative to those in which the motion does not resolve the case. Of course, nobody has apparently been aware of these differences (until now). And in many instances, they are more-or-less random, in the sense that it is unclear who they help who will they harm.<sup>161</sup> Still, we think that policy-induced heterogeneity that serves no useful purpose is properly counted as a cost of operating the system, especially if parties are risk-averse.

Finally, we should consider opportunity costs. It is almost inevitable that other matters will be pushed aside in an effort to comply with the List: trials may be delayed; discovery orders may take longer to process; criminal matters may be temporarily set aside; judges may forgo oral argument.<sup>162</sup> We are cer-

---

<sup>158</sup> See *id.*

<sup>159</sup> Zuckerman, *supra* note 151, at 387 (recognizing that there are tradeoffs between timeliness and accuracy, and suggesting that accuracy should not be courts' only objective).

<sup>160</sup> In other contexts, the Supreme Court has recognized the "avoidance of inequitable administration of the laws" as an important principle. *Hanna v. Plumer*, 380 U.S. 460, 468 (1965).

<sup>161</sup> Some argue that randomness itself may not be necessarily unfair. See, e.g., David Lewis, *The Punishment that Leaves Something to Chance*, 18 PHIL. & PUB. AFFAIRS 53, 58, 64-65 (1989) (arguing that randomly adding or subtracting time to all criminal sentences is not unfair).

<sup>162</sup> Displacement is inevitable unless either (i) judges and clerks have slack resources at other times, or (ii) they respond to the List purely by sacrificing leisure and working more hours than they otherwise would, both of which seem unlikely.

We should note that we looked at the data on criminal sentencing hearings and found no evidence that they were being crowded-out in List weeks—there were no drop-offs in hearing volumes during weeks 13 or 39, although the noisiness of the data makes it difficult to detect any pattern that might exist.

tain that the drafters of the CJRA did not give careful consideration to these inevitable consequences of focusing judges' attention on faster processing of motions.

## 2. *Benefits*

On the other side of the ledger, consider the benefits created by the List. The List's six-month deadline does speed up the processing of motions.<sup>163</sup> This time savings amounts to between ten and thirty days for the average motion, with the lower bound being closer to our preferred estimate.

Supposing that there is some welfare gain from reducing the time to resolve a motion, how big is it likely to be?<sup>164</sup> The social value of adjudication stems largely from the precedent it creates and the deterrence value it provides. Faster resolution of a case or motion—assuming no loss in accuracy—speeds up the attainment of these benefits, and hence is valuable.

Even a generous assessment would conclude that the List generates very modest benefits. An oversimplified and conservative (i.e., generous) answer starts by assuming that the List speeds up the resolution of the outcome of interest by thirty days. If we treat this as a discounted cash flow problem, the value of speeding up the receipt of \$1 by thirty days is

$$v = 1 - 1/(1 + r)^{(30/365)},$$

where  $r$  is the appropriate annual interest rate (and interest is compounded daily). Even for a value of  $r$  as high as 20% per year,  $v$  is less than two pennies per dollar, and a rate of 5% per year yields a  $v$  of 0.5 cents per dollar. Of course,  $v$  measures the saving *per dollar of social value created*, and we have no idea how large the social value of adjudication actually is. Still, speeding up the realization of that value by thirty days is unlikely to yield large benefits.

We think the bottom line is pretty clear. Set against the faster disposition of (some) motions, we have uncertain but non-trivial administrative and opportunity costs and the creation of unjustified heterogeneity among cases. Worse still are

---

<sup>163</sup> We find virtually no effect of the three-year case-level deadline, and we therefore ignore it in this discussion.

<sup>164</sup> Private litigants care about which party prevails and the resulting transfer (or not) from defendant to plaintiff. From society's perspective, however, the value of litigation arises from the deterrent effect it has on future conduct and the clarification of the rules that a decision provides. These benefits are doubtless real, but they are extremely difficult to measure. See STEVEN SHAVELL, FOUNDATIONS OF THE ECONOMIC ANALYSIS OF LAW 450–56 (2004). We thank Tom Miceli for clarifying this point.

the risks of errors and/or biases induced by the List. Even after acknowledging the substantial degree of subjectivity involved, we find it hard to see how the gains in processing time could be worth the actual and potential costs.

## B. Complying with the CJRA

Although we recommend eliminating the List, we recognize that this decision is not in the hands of the federal courts. In the absence of its repeal, the CJRA obliges the AO to produce semi-annual reports and requires the AO to publish “the number of motions that have been pending for more than six months” and the name of each case with such a motion.<sup>165</sup> The plain language appears to require something very similar to what the AO currently does.

There is a little wiggle room. The statute permits the AO to set “the standards for categorization or characterization of judicial actions.”<sup>166</sup> This provision has allowed the AO to give judges a grace period of an additional month in motion processing. Although the AO seemed to diverge from the plain language by setting a seven-month deadline, the practice can be justified based on the idea that a motion is not really “pending” until the opposing party has had an opportunity to respond to the movant. This provision also justifies tolling the three-year clock on *qui tam* cases until the government has responded. The AO could probably interpret this provision to permit it to toll the clock for other motions that are more difficult to resolve, but the statute cannot fairly be interpreted to allow the AO to replace the List with a better measure.

Given that the AO has limited flexibility to alter the CJRA’s mandates, we think a second-best option would be to dilute the force of the List by providing more and different information, so that judges would not be so focused on a single measure. Nothing in the statute prohibits the AO from providing additional information beyond what it currently publishes.<sup>167</sup> For the reasons explained below, we suggest that the AO highlight

---

<sup>165</sup> 28 U.S.C. § 476(a)(1) (2012).

<sup>166</sup> 28 U.S.C. § 476(b), 481 (2012).

<sup>167</sup> Average or median case and motion processing times are easily calculated from data the courts already have. Indeed, such data are already published, although only at the district level and not for individual judges. See *Table C-5. U.S. District Courts—Median Time Intervals from Filing to Disposition of Civil Cases Terminated, by District and Method of Disposition, During the 12-Month Period Ending June 30, 2016*, U.S. COURTS, [http://www.uscourts.gov/sites/default/files/data\\_tables/stfj\\_c5\\_630.2016.pdf](http://www.uscourts.gov/sites/default/files/data_tables/stfj_c5_630.2016.pdf) [<https://perma.cc/2VKW-FR3U>] (visited Feb. 16, 2018).

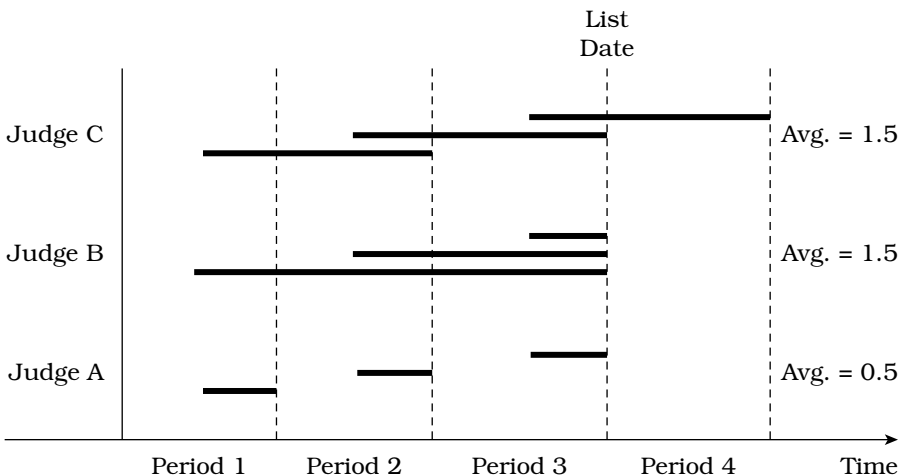


weighted averages in addition to reporting specific motions that have exceeded the deadline. This additional reporting may dilute the incentive to decide motions by the deadline and to mothball motions not so decided until the next deadline, diminishing the heterogeneity of motion processing caused by the List. Inclusion of other measures in the Six-Month List, including a “bench presence” measure, for example, could further dilute the effect of the List.<sup>168</sup>

1. *Problems with the Current Incentives*

In order to design a better incentive, it is first necessary to explain why it is that the List does a poor job of achieving its stated goal of reducing processing times. To understand the problem, consider a simple example involving three judges, A, B, and C. The judges’ workloads are shown in Figure 15. Each receives one new case per period for three periods. Each is subject to a “List” with a three-period window: any case filed in periods 1–3 that is open more than one-half period and is not resolved by the end of period 3 automatically appears on the List.

FIGURE 15: HYPOTHETICAL RELATIONSHIP BETWEEN A BACKLOG LIST AND JUDICIAL WORKLOAD<sup>169</sup>



<sup>168</sup> See William G. Young & Jordan M. Singer, *Bench Presence: Toward a More Comprehensive Model of Federal District Court Productivity*, 118 PENN. ST. L. REV. 55, 58 (2013).

<sup>169</sup> Note: Each judge is assumed to receive one new case in the middle of each period. Any case not resolved at the end of period 3 is put on the List.

Judge A disposes of each of her three cases at the end of the period in which it is filed, so her average duration (time between filing and disposition) is 0.5 periods. By contrast, Judge B disposes of all three of *his* cases at the last possible instant before the List is compiled, giving him an average disposition time of 1.5 periods. Yet, the List mechanism treats A and B identically—neither appears at all, even though B takes three times longer than A. The scheme thus provides no incentive for B to improve his case-processing time, since his current practice of waiting until the last minute to close his backlog already keeps him off the List altogether. And while Judge C has the same average performance as Judge B (a 1.5 period average duration), he *does* show up on the List because his third case is not resolved within the appropriate window.<sup>170</sup> The List thus ends up treating like cases differently and different cases alike—not a good property for any incentive system.

Generalizing from this example, we can see that the Six-Month List as constructed has two significant failings. It does not appropriately track the desired “outcome” variable, even if we assume-away all considerations other than case or motion duration; and it puts too much duration risk on judges. We discuss each of these in turn.

a. *Incentive Mismatch & Excess Heterogeneity*

Rather than pushing judges to work harder or faster *in general*, the List only incentivizes taking less than three years to close a case or 214 days to dispose of an eligible motion. The shadow of the List gives judges a reason to work at closing an ‘eligible’ case or motion in the last few weeks of March or September, but offers little incentive for closing any matter earlier than that. Put differently, judges have no incentive to close any

---

<sup>170</sup> These are not hypothetical concerns. Consider an example from the September 2016, Six-Month List Motions page for Judge Thomas F. Hogan in the D.C. Circuit. The report cites two motions in the case of *Lucas v. District of Columbia*, 214 F. Supp. 3d 1 (2016), meaning that they were not decided in time to avoid appearing on the List. But the accompanying “Notes” field indicates that these motions were “decided after the end of the reporting period” (but presumably before the September List was compiled). Deciding these motions on October 5, 2016, is almost as good as deciding them five days earlier, but Judge Hogan got no official credit for *almost* meeting the deadline. We note that some versions of the published motions List contain a column labeled “CJRA Deadline” which is 214 days from the motion’s filing date. This allows the interested reader to determine not just *that* the motion is on the List, but how much over the deadline it is. The information is not presented in a useful fashion and appears only sporadically. See *CJRA Table 8—Reports of Motions Pending Over Six Months For Period Ending September 30, 2016*, *supra* note 8.

case or motion that is still alive *after* any List until the next List is about to be compiled.

Unless the cost of unresolved matters jumps discretely at six-month intervals—falling to zero in between times—this form of incentive is hard to justify. If the policy objective is to reduce the “typical” duration of filed cases, a different measure would clearly be appropriate.

These flaws are by no means entirely theoretical. Our evidence demonstrates that the List introduces significant heterogeneity in ways that are undesirable and were clearly not intended by its designers. First, its duration-reducing effects are not uniform—the List has no effect on long or short duration motions; its influence is *only* felt on motions that would have taken roughly five to eight months to process in its absence.

Second, and perhaps even worse, the List generates heterogeneity by filing date. A motion filed on August 29 will on average be disposed of roughly one month earlier than one filed on September 1, just because the former faces a 214-day deadline while the latter is not List eligible for 396 days. A motion’s filing date is an arbitrary occurrence, and offers no legitimate basis for disparate treatment by the courts.

Finally, the data on motion closures at least hints at that possibility that judges may actually *reduce* their efforts to close September-filed motions between April and September of the following year. Motion closures fall-off substantially during this period, and do not rise again (secondary bunching) until the September deadline begins to loom. In line with “Student Syndrome” or Parkinson’s Law,<sup>171</sup> it is actually possible that the List speeds up August-filed motions but *delays* the resolution of some of the September motions relative to a world of no List at all.<sup>172</sup> This kind of induced disparity seems impossible to justify.

#### b. *Excessive Duration Risk Borne by Judges*

A second failure of the existing rules is their sensitivity to noise, which imposes an unnecessary risk on judges. Cases are randomly allocated to judges, so purely by luck, a judge could end up with a set of cases that are unusually difficult or

---

<sup>171</sup> See *supra* note 77.

<sup>172</sup> The September motions are serving as the control group for the August motions, so there is no counterfactual baseline available for what would have happened to the September motions absent the List. We freely acknowledge that we are speculating beyond the evidence here.

complex. That means that in any six-month interval, a judge might end up on the List just because she happened to draw a set of cases or motions that proved unusually hard to resolve quickly. If judges are risk-averse—as most people are—forcing them to bear this risk is costly. Welfare would be enhanced if we could provide the same incentives for speed without imposing as much risk on judges.<sup>173</sup>

## 2. *Better Incentives (If Any)*

Until Congress acts, the best option is to supplement the List with an evaluation of a judge's aggregate performance over all the cases she handles, combined with a risk adjustment measure.

We suggest that the interval for such measurement should be more frequent than every six months, so as to dilute the incentive to bunch decisions into two weeks of the year. Reasonable minds may differ as to whether the average or median duration would be a better measure of each judge's typical performance in this context. Either of these, or a more nuanced option,<sup>174</sup> would be superior to the outcome measure employed by the List.

Focusing on mean or median duration also enhances judicial discretion. Judges can trade extra time spent on cases that require additional effort against those that can be disposed of more quickly. Current rules focus only on whether a case is cleared before the relevant deadline and do not recognize tradeoffs of this sort.

Because we suspect that judges are risk averse, enhancing speed of processing while limiting risk could be achieved with some version of risk adjustment,<sup>175</sup> as well as by focusing on

---

<sup>173</sup> The tension between providing incentives and insuring against risk has been widely recognized in economics. See KENNETH ARROW, *THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS* 613 (1962); *supra* notes 2 and 44. Here, however, we have a rare example of a free lunch. If we want to, we can make incentives 'stronger' without forcing judges to bear more risk.

<sup>174</sup> For example, we could also publicize the standard deviation of duration measured across a judge's portfolio of cases. Or we could focus on the Xth percentile—the length of time such that X percent of all cases were resolved faster than this, where X might be some large number like 90%.

<sup>175</sup> For a recent theoretical analysis on this topic, see Henry Y. Mak, *Provider Performance Reports and Consumer Welfare*, 48 *RAND J. ECON.* 250, 253, 267 (2017). There are various techniques for adjusting performance measures to compensate for the difficulty of achieving the desired results. In medicine, such techniques are called "risk-adjustment." In teaching, the same principles underlie the notion of "value-added," in which a teacher is rewarded not for how well his pupils actually perform, but for how much their results exceed *ex ante* expected performance.

typical outcomes instead of extreme cases. Risk adjustment techniques are widely used to evaluate performance in other contexts, and are designed to distinguish between results that are due to luck and those that reflect the actual contribution of a particular agent. For example, a heart surgeon who operates on older, sicker and harder-to-treat patients may have a lower apparent “success” rate than one whose patients are younger and healthier. But the surgeons’ measured success rates reflect both their skill and effort *and* the mix of patients they treat. To avoid this kind of confounding, risk adjustment methods try to control for the difficulty of the assigned task when evaluating performance.

A judge who randomly receives a portfolio of hard-to-close cases should not be viewed harshly for her failure to resolve them within the required time limit. If it were possible to assess the “difficulty” of resolving a case or its expected duration based on objective factors at the time the case is filed,<sup>176</sup> each judge’s actual performance could be adjusted to reflect the difficulty of the mix of cases she or he was assigned. We also note that there are some selection effects operating in the federal system. Although most judges are assigned their cases randomly, some judges are selected to oversee complex, multidistrict litigation<sup>177</sup> and some courts have worked to attract particular types of cases.<sup>178</sup> A risk adjustment system could account for these variations.

At least in the ordinary situation faced by judges randomly assigned their caseload, a focus on average or median outcomes could also reduce risk. The Six-Month List punishes a judge for the duration of her longest cases or motions, regardless of how well she does in handling the typical case. But the number of such outliers is more variable than the average or median duration of a *group* of cases. Using outliers to measure

---

<sup>176</sup> Risk adjustment does require a good model for predicting the relevant outcome (such as duration). Without it, the comparison between actual and predicted outcomes is worthless. This may be a problem for judges, given that the empirical estimates we have come up with do a poor job of explaining case durations, and most of the variance in case durations is not explained by the independent variables we can actually measure (although admittedly we have only a few variables at our disposal). This is not promising, but we know that there are more data inside the judicial system that are not made available to the general public, and perhaps these could be used to fit better predictive models.

<sup>177</sup> See 28 U.S.C. § 1407(b) (2012) (multidistrict litigation statute). The Judicial Panel on Multidistrict Litigation determines to which judge to assign these lawsuits.

<sup>178</sup> Daniel Klerman & Greg Reilly, *Forum Selling*, 89 S. CAL. L. REV. 241, 280 (2016) (describing the successful attempt by at least one federal district court to attract certain types of cases).

performance thus imposes more risk on the judge for a given level of incentive.<sup>179</sup> Not only is the mean or median duration of a judge's cases more likely to approximate what society actually cares about, but it is also less noisy than the measure used to compute the Six-Month List.

#### CONCLUSION

Our study of the Six-Month Lists yields four important lessons for the design of incentives.

First, conformity to social norms appears to be a very substantial motivator, at least for this group of elite professionals, even when unaccompanied by tangible rewards or punishments. Second, incentives that seem on their face simple (closing motions within six months of filing) end up with unintended consequences that, with a little thought, should have been predicted. In this case, bunching and heterogeneity in duration are both unintended but entirely foreseeable consequences of imposing a calendar deadline. Third, there are likely to be tradeoffs between speed and accuracy, even if these are hard to measure; these require more careful normative evaluation before imposing incentives for speed. Finally, even if it makes sense to provide an incentive for judges to work faster, existing incentives are poorly designed for that task. But given the difficulty of evaluating the key output of judges—high quality decisions—our preferred approach is to eliminate incentives altogether. In this context, “accountability” cannot survive a cost-benefit analysis.

---

<sup>179</sup> The coefficient of variation is defined as the standard deviation divided by the mean. Given plausible parameters, the coefficient of variation for the duration of the average case (in a random sample of 150 cases) is about one-half as large as the coefficient of variation for the number of cases exceeding three years in length.

## APPENDIX 1

FIGURE A: SAMPLE PAGE FROM THE MARCH 31, 2016,  
SIX MONTH LIST

CJRA Table 8W—Report Of Motions  
Pending Over Six Months For  
Period Ending March 31, 2016  
DC Circuit, District Judge HOGAN, THOMAS F.

Nature of Suit	Case Title	Motion Text	Status	Status Description
Drug Related Seizure of Property	UNITED STATES OF AMERICA v. &#036;189,880.00 IN U.S. CU	MOTION to Dismiss for Lack of Prosecution	V	Change of Counsel
Drug Related Seizure of Proper	UNITED STATES OF AMERICA v. ONE 1998 CHEVROLET K2500 PI	MOTION to Strike [1] Complaint,	V	Change of Counsel
		MOTION to Dismiss for Lack of Prosecution	V	Change of Counsel
		MOTION to Strike [1] Complaint	V	Change of Counsel
Other Contract Actions	LANNAN FOUNDATION v. GINGOLD et al	MOTION to Dismiss	B	Opinion/Decision in Draft
		MOTION for Leave to File Supplemental Complaint	B	Opinion/Decision in Draft
Other Statutory Actions	CHEROKEE NATION vs. NASH, et al	MOTION for Summary Judgment and Supporting Memo	B	Opinion/Decision in Draft
		Cross MOTION for Summary Judgment and Opposition	B	Opinion/Decision in Draft
		Cross MOTION for Summary Judgment and Opposition	B	Opinion/Decision in Draft
		MOTION to Strike Cross MOTION for Summary Judgment	B	Opinion/Decision

Total Motions for District Judge : HOGAN, THOMAS F. 10

TABLE A: SUMMARY STATISTICS, AO CASE-LEVEL DATA  
(by Termination Week)

	Pre-CJRA: 1980-1990		Post-CJRA: 1991-2017		
	Non-List Week	List Week	Non-List Week	List Week	Total
N	2,385,226	121,486	6,257,044	340,271	9,104,027
Mean Duration, in Days	339	351	365	425	361
Median Duration, in Days	219	229	231	300	230
SD	363	372	448	451	427
Plaintiff Win Rate	59.7%	57.7%	38.0%	31.3%	45.5%

TABLE B: BASELINE COVARIATE BALANCE FOR SUMMARY JUDGMENT MOTIONS

<b>Variable</b>		<b>6 days</b>	<b>10 days</b>	<b>14 days</b>	<b>28 days</b>	<b>Full Sample</b>
Nature of Suit Category	Pearson's Chi-Sq	7.474	3.704	5.197	6.664	7.935
	<i>p</i>	0.381	0.813	0.636	0.573	0.440
	<i>N</i>	129	205	275	524	757
Days Between Case and Motion Filing Dates	K-S Test	0.200	0.147	0.086	0.164***	0.178***
	<i>p</i>	0.149	0.196	0.637	0.001	0.000
	<i>N</i>	130	206	276	525	758
Party Making Motion	Diff in Means	0.055	-0.027	-0.019	-0.045	-0.067*
	<i>p</i>	0.530	0.700	0.755	0.295	0.062
	<i>N</i>	127	205	275	524	757
Cross-Motion	Diff in Means	0.073	0.077	0.085	0.064	0.074
	<i>p</i>	0.380	0.248	0.136	0.116	0.225
	<i>N</i>	129	204	274	523	754
Circuit	Pearson's Chi-Sq	7.255	9.727	9.579	7.790	6.689
	<i>p</i>	0.778	0.555	0.569	0.732	0.824
	<i>N</i>	129	205	275	524	757
District	K-S Test	0.076	0.091	0.056	0.059	0.042
	<i>p</i>	0.980	0.751	0.969	0.721	0.878
	<i>N</i>	129	205	275	524	757

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Difference-in-means is used for binary variables. Pearson's chi-square test is used for small samples, and the Kalmogorov-Smirnov (K-S) Test is used for continuous variables.



FIGURE B(1): SUMMARY JUDGMENT MOTION OFFSET RESOLUTION DATE, BY CALENDAR WEEK TREATMENT GROUP

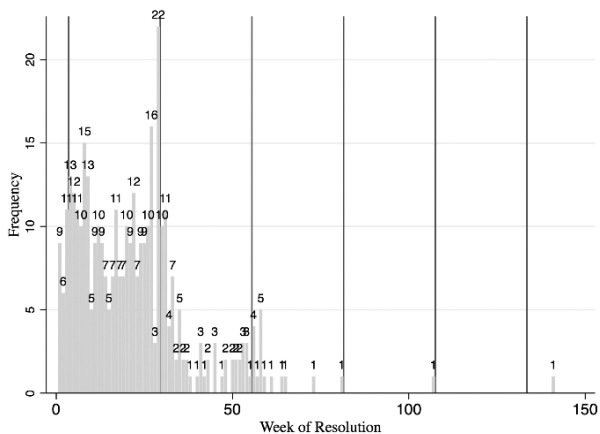


FIGURE B(2): SUMMARY JUDGMENT MOTION OFFSET RESOLUTION DATE, BY CALENDAR WEEK CONTROL GROUP

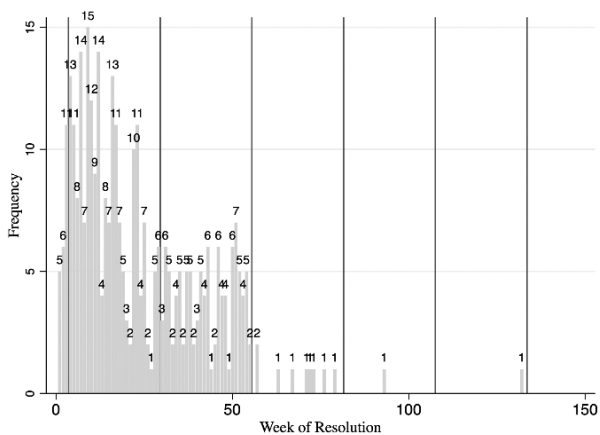


FIGURE B(3): DIFFERENCE IN THE NUMBER OF OFFSET CLOSURES BETWEEN THE TREATMENT AND CONTROL GROUPS (WEEK 29 = MARCH 31, 2012 & WEEK 55 = SEPTEMBER 30, 2012), BY CALENDAR WEEK

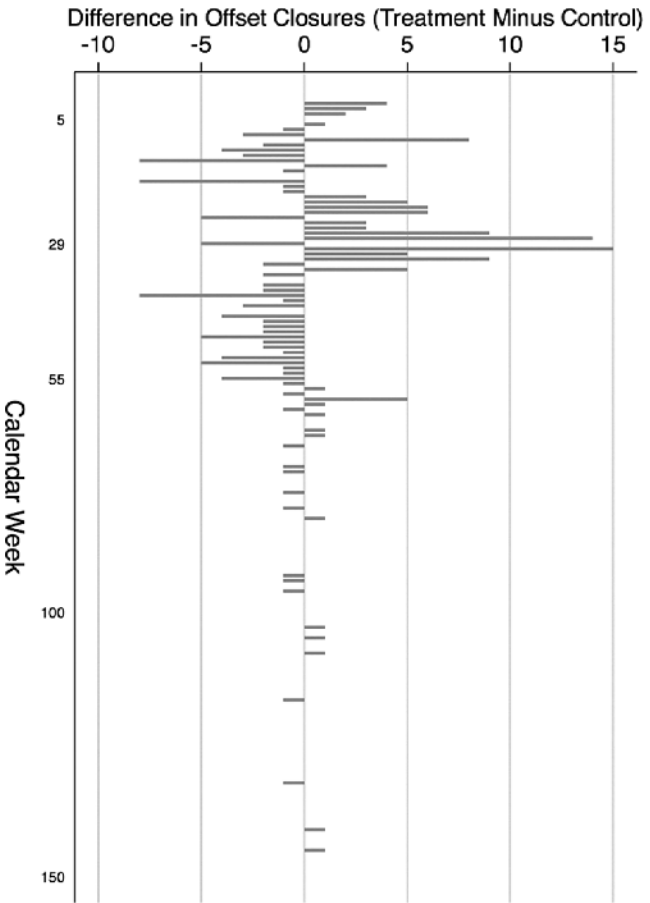


TABLE C: OLS REGRESSIONS OF NUMBER OF MOTIONS CLOSING IN 2012, BY OFFSET CALENDAR WEEK (MOTIONS SURVIVING ON OR AFTER AUG. 30, 2012)

VARIABLES	(1) Treatment	(2) Control	(3) Difference
Time	-0.122*** (0.0208)	-0.110*** (0.0140)	-0.00621 (0.0122)
Time <sup>2</sup>	0.000369*** (6.50e-05)	0.000329*** (4.46e-05)	2.40e-05 (3.72e-05)
Week 27	6.145*** (0.716)	-2.402*** (0.415)	11.83*** (0.479)
Week 28	-7.336*** (1.182)	1.929*** (0.501)	-5.811*** (0.672)
Week 29	11.62*** (1.524)	0.828** (0.414)	13.69*** (1.587)
Week 53	1.087*** (0.260)	1.522*** (0.259)	-0.788*** (0.287)
Week 54	-1.731*** (0.420)	1.864*** (0.407)	-4.000*** (0.341)
Week 55	-1.438*** (0.528)	-0.564 (0.871)	-1.213*** (0.416)
Constant	8.930*** (1.481)	8.259*** (0.968)	0.266 (0.908)
Observations	255	255	255
R-squared	0.641	0.558	0.355

Standard errors in parentheses

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

Figures B(1), B(2), and B(3) present these “age-standardized” data. This allows us to compare what happens to these two groups in the 207th to 214th days of their lives (regardless of when they were born), i.e., *holding “age” constant*.<sup>180</sup> These figures demonstrate that motions in the treatment group are being decided around the 214 day mark while motions in the control group are not. In other words, we see

<sup>180</sup> Because 2012 was a leap year, we had to make some adjustments to week lengths so that we could start on August 30, 2011 and have a 7-day interval that ended on March 31, 2012. Details of these adjustments are available from the authors.

bunching in the treatment group decisions but not the control group. This analysis does not show secondary bunching, however, because if we control for motion age *within* the treated group, we lose the connection to the real-time calendar that drives judges to process motions.

Evidence from the offset regressions shows the effect of the List on the number of orders per week when we consider the age of motion as a function of days rather than place in the calendar year. Consistent with our findings, Table 2 demonstrates that there is an increase in closures of 13.7 cases in the treatment group relative to the control group in week 29 (the List week). Notably, we found a coefficient of -5.8 for the treatment group for week 28 (the week before the List week) meaning that judges are deciding fewer motions right before the deadline in the treatment group than the control group. No definitive conclusion can be drawn from this behavioral anomaly.

TABLE D(1): OLS ESTIMATES FOR THE IMPACT OF THE LIST ON SUMMARY JUDGMENT MOTION DURATION (ALL MOTIONS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	10 days	10 days	10 days	14 days	14 days	14 days	28 days	28 days	28 days	Full Sample	Full Sample	Full Sample
Estimate	3.40	10.91	37.71	-14.03	-6.65	.83	-27.07*	-21.01	-20.22	-15.92	-12.14	-8.37
SE	26.27	25.67	49.78	21.46	21.36	22.11	15.99	15.58	20.03	13.08	12.79	14.43
<i>p</i>	0.90	0.67	0.47	0.51	0.76	0.94	0.09	0.18	0.33	0.22	0.34	0.57
Observations	206	204	204	276	274	274	525	523	523	758	754	754
Cross-Motion Control	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
District FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

TABLE D(2): OLS ESTIMATES FOR THE IMPACT OF THE LIST ON SUMMARY JUDGMENT MOTION DURATION (MOTIONS <2 YEARS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	10 days	10 days	10 days	14 days	14 days	14 days	28 days	28 days	28 days	Full Sample	Full Sample	Full Sample
Bandwidth	10 days	10 days	10 days	14 days	14 days	14 days	28 days	28 days	28 days	Full Sample	Full Sample	Full Sample
Estimate	-21.59	-13.78	-6.77	-27.51*	-19.21	-28.12	-25.49**	-18.89*	-24.09*	-12.96	-8.95	-10.79
SE	18.13	17.53	18.53	15.27	14.94	17.27	10.92	10.38	13.18	9.16	8.74	8.12
<i>p</i>	0.24	0.43	0.72	0.07	0.20	0.13	0.02	0.07	0.10	0.16	0.31	0.21
Observations	204	202	202	272	270	270	517	515	515	746	742	742
Cross-Motion Control	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
District FE	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

APPENDIX 2:  
AN ALGEBRAIC DEMONSTRATION OF HOW TIME SHIFTING  
COULD LOWER LIST WEEK WIN RATES

Here is a demonstration that shifting “easy” cases to List weeks will lead to a decline in the plaintiff win rate for those weeks.<sup>181</sup> Suppose that cases (or motions) can be divided into 3 groups:

P = number of easy plaintiff wins

D = number of easy defendant wins

H = number hard cases.

The plaintiff win rate for easy cases is just

$$\varepsilon = \frac{P}{P + D}$$

The win rate for “hard” cases is 50 percent, which is in a sense what it means to be “difficult.” Suppose that “easy” cases (P + D) are  $\alpha\%$  of the total.<sup>182</sup> Then the overall win rate is

$$w = \alpha\varepsilon + (1 - \alpha)(1/2),$$

which is just the weighted average of the win rates for the two kinds of cases. We know that the overall win rate is roughly 40 percent. Then the win rate in easy cases is

$$\varepsilon = \frac{w - (1 - \alpha)/2\alpha}{\alpha}$$

which is necessarily less than the overall win rate,  $w$ .

It follows that as the share of easy cases decided over some period of time increases, the overall win rate must fall. That is,

$$\partial w / \partial \alpha = \varepsilon - 1/2$$

which is less than zero, since  $\varepsilon < 40$  percent.

If judges disproportionately shift “easy” cases to List weeks, the composition of List week decisions will be more heavily weighted towards defendant-favored cases and the plaintiff win rate will drop. Given that plaintiffs win less than 40 percent of the easy cases, shifting *difficult* cases (with a 50 percent win rate) to List weeks will *raise* the plaintiff win rate.

<sup>181</sup> It will also lead to a *rise* in the plaintiff win rate for non-List weeks.

<sup>182</sup> That is  $\alpha = (P + D)/(P + D + H)$ .