

Deadlines, Judicial Decisions, and Firm Outcomes: Evidence from Productivity Quotas in Croatia

Daniel L. Chen and Viknesh Nagarathinam

Abstract

Judicial backlogs slow growth by delaying enforcement, and governments worldwide increasingly hold judges to numerical caseload targets. We ask whether such targets buy speed at the cost of quality, using Croatia's quota regime—annual point targets that court presidents monitor monthly, with shortfalls carrying disciplinary consequences. With administrative micro-data on 4.1 million civil and commercial suits decided between 2010 and 2016, one million matched to defendant-firm balance sheets, we document a desk-clearing pattern. Case closures surge in the final days of each month, in every year, court, and case type. Decisions in the surge are issued faster after the last hearing and are roughly nine percent more likely to be appealed; the appeal premium is stable as month, weekday, week, and year fixed effects are added. Defendant firms receiving a month-end ruling have substantially lower subsequent profit—on the order of a third—an effect equally stable across fixed-effect specifications and larger when the ruling is appealed. The most natural form of selection, in which judges save easy cases for the deadline, would bias against these findings; and the hardest cases are not the ones held for month-end, since late-month decisions are faster, not slower. The results suggest that rolling evaluation windows, complexity-adjusted weights, and re-review of deadline-period rulings could preserve speed while limiting collateral harm.

I. Introduction

Courts underpin market exchange. By enforcing contracts, protecting property, and resolving disputes, they reduce the uncertainty that deters investment and entrepreneurship (North 1990; Djankov et al. 2003). When adjudication is slow, contracts are riskier to write and costlier to enforce, with measurable consequences for credit and firm growth (Chemin 2012; Di Vita 2010). Yet courts in much of the world remain burdened by backlog and delay (World Bank 2015). In response, governments increasingly evaluate judges on quantitative output, requiring each to close a set number of cases, or accumulate a set number of weighted points, within a

fixed period. The premise is intuitive: judges measured on output will work faster, shrinking backlogs and widening access to justice.

Numerical targets for judges are now a global phenomenon, and a contested one. Spain tied a portion of judicial pay to workload benchmarks (“módulos”) until its Supreme Court struck the scheme down as a threat to judicial independence. The Netherlands funds its courts by weighted caseload, giving them de facto disposition targets. Brazil’s National Council of Justice publishes annual per-judge productivity in its *Justiça em Números* reports. China’s Case Quality Evaluation System, in force since 2011, ties bonuses, promotion, and demotion to disposal and reversal targets. Malaysia cut its backlog by roughly half in two years under court-level disposal quotas. And in the United States, the Department of Justice has set immigration judges an annual target of 700 completed cases. These regimes share a structure—reward speed, penalize backlog—and a recurring worry, that pressure to produce will degrade the quality of what judges produce. Croatia, which has held judges to formal case quotas since the early 2000s and which digitized its entire civil docket in 2010, offers an unusually clear window onto whether that worry is justified.

Whether output targets degrade quality is, in principle, an empirical question, and a large literature outside the judiciary suggests the answer is often yes. Numerical targets and deadlines induce bunching, and output produced near a deadline tends to be worse. Equipment dealers paid on fiscal-year targets book a surge of discounted sales each December (Oyer 1998); enterprise sales forces game quarter-end quotas (Larkin 2014); Navy recruiters relax standards at quarter-end (Asch 1990); federal contracting officers facing “use-it-or-lose-it” budgets sign lower-quality contracts in the final week of the fiscal year (Liebman and Mahoney 2017); patent examiners given less review time grant more low-quality patents (Frakes and Wasserman 2017);

and drug reviewers clear more applications before salient deadlines, with worse post-market safety (Carpenter, Zucker, and Avorn 2008; Cohen, Gurun, and Li 2021). The common mechanism is desk-clearing: when output is rewarded and the clock is fixed, agents dispose of pending work in a rush, producing both more output and more error at predictable points in the cycle.

Direct evidence for the judiciary is thinner, because few systems record the precise timing of decisions. The closest precedent is de Figueiredo, Lahav, and Siegelman (2020), who show that the U.S. federal “Six-Month List”—a semiannual public report of each judge’s backlog—leads district judges to dispose of more cases just before the reporting date, with suggestive evidence of more errors. That setting is a soft disclosure requirement rather than a binding quota, rests on a hand-coded sample of summary-judgment motions, and is not connected to the economic fortunes of litigants. Croatia differs on each count: its deadline is a formal quota whose shortfall carries disciplinary consequences, not merely an informal benchmark of the kind documented for drug regulators (Cohen, Gurun, and Li 2021); the ICMS records the near-universe of civil and commercial cases; and a match to firm balance sheets lets us follow rushed adjudication into the real economy.

Croatia’s regime supplies the institutional structure the question requires. The Ministry of Justice publishes annual point quotas by case category, but court presidents review each judge’s progress monthly and must refer persistent shortfalls for discipline, so judges face a recurring, soft monthly deadline; judges below 80 percent of target risk sanction, while those above it gain standing for promotion. Since 2010 the Integrated Case Management System (ICMS) has logged every civil and commercial case—filing, hearings, decision, closure, litigant identities, claim

value, decision type, and appeals—at a daily granularity that makes the timing of adjudication observable.

We document three facts. *First*, closures bunch sharply at month-end: counts climb through the first three weeks of each month and surge in the final days, with the spike present in every year of the sample, in each major court tier, and across case types. *Second*, decisions in the surge bear the marks of haste. Conditional on claim value, court, and month-year, the interval between the last hearing and the written decision is shorter for late-month closures, and those closures are roughly nine percent more likely to be appealed—an estimate that barely moves as month, weekday, week, and year fixed effects are added. *Third*, the consequences reach beyond the courtroom. Among defendant firms we observe before and after closure, those receiving a month-end ruling have substantially lower subsequent profit—the estimated effect is on the order of a third, is stable across fixed-effect specifications, and is larger when the ruling is appealed.

These patterns are difficult to attribute to selection, even though judges choose what to close when. The cases a judge would clear to pad a count are the easy ones, which are appealed less—so selection biases against our findings—and the decisions are issued faster, not slower, near the deadline, indicating haste rather than a queue of hard cases held for month-end; the estimates are, moreover, stable across rich controls. The paper connects the economics of courts to the literature on deadlines and bureaucratic incentives, and is, to our knowledge, the first to trace the within-month timing of adjudication through to the economic outcomes of the firms being judged.

The findings bear directly on a policy choice many governments are making. Numerical targets can buy speed, but the Croatian evidence indicates they also invite corner-cutting whose costs fall on litigants. We close by discussing designs that may preserve the gains while blunting

the incentive to sprint: rolling evaluation windows, complexity-adjusted weights, reversal-rate penalties, and re-review of decisions issued in the deadline window. The remainder of the paper proceeds as follows. Section II describes Croatia's quota regime; Section III the data; Section IV the empirical strategy and identification; Section V the results; and Section VI concludes.

II. Institutional Setting

Croatia has held judges to numerical workload standards since well before the digital reforms that make its case data observable; productivity norms were already in force in the early 2000s. The legal basis is the Courts Act, which delegates to the Ministry of Justice the authority to set framework criteria for judicial work. Each year the ministry, on the proposal of the Supreme Court, publishes a schedule assigning a point value to every category of lawsuit and stating how many points a judge in each division must earn over the year. Case types are weighted by complexity, so that a routine enforcement action counts for a fraction of a contested trial, and a judge's annual obligation is expressed as an equivalent number of standard cases—on the order of 165 for a municipal civil judge, with lower counts of more heavily weighted matters in the commercial and county courts. Productivity targets of this kind are common in civil-law judiciaries (Buscaglia and Dakolias 1999; Palumbo et al. 2013).

Although the yardstick is annual, monitoring is continuous. The Courts Act obliges court presidents to verify throughout the year that judges are fulfilling their duties, and in practice presidents run a progress report each month. A judge who falls behind receives a written warning; persistent shortfall without good reason obliges the president to refer the matter to the State Judicial Council, and repeated failure to reach 80 percent of target can lead to removal. In the other direction, judges who exceed their target accrue recognition that raises their standing for promotion. The result is a soft monthly deadline: presidents rarely sanction a judge who is

briefly behind in a given month, but consistent deficits draw scrutiny, so judges manage their dockets to remain visibly on pace. Because backlogs are large, there is always a queue of cases awaiting a final revision or opinion that can be completed quickly when a judge needs to close the gap before the president's review.

Evaluation pairs this quantity standard with quality and timeliness measures. The State Judicial Council's methodology scores judges on whether they met the required number of decisions, whether they issued decisions within prescribed time limits, and the rate at which their decisions are upheld on appeal—reversal rates computed over a multi-year window, with low reversal rewarded. The system is thus designed, on paper, to discourage trading quality for volume. Whether it does so in the final days of the month is the empirical question.

The regime has drawn sustained criticism. In 2020 the ministry raised targets in several divisions—by as much as half in some, as in a proposed increase for investigating judges from roughly 600 to 900 matters—arguing that falling backlogs justified recalibration. Judicial associations objected that the new norms were unrealistic: at the High Commercial Court, traditionally among the better-performing courts, roughly half of judges were unable to meet the 2020 target, and the association warned that quotas exceeding a realistic year's work (about 220 working days) would pressure judges toward rubber-stamping. A 2021–22 review supported by the European Union and Norway examined the methodology with an eye to making it fairer and more sensitive to case complexity. Adjustments followed, but the core structure—numerical quotas under monthly surveillance—remains in place. That combination of high-powered targets and continuous monitoring is what the rest of the paper examines.

III. Data

Our analysis draws on the ICMS, the case-management system the Ministry of Justice introduced in 2010 to record and track court proceedings electronically. The system covers Municipal, County, Commercial, and Supreme Court cases; administrative and misdemeanor courts maintain separate systems and lie outside our data. For each case the ICMS records a unique identifier; the court and division; identities of plaintiff and defendant (national ID for individuals, tax ID for firms); the filing date; the date of each hearing; the date the dispositive decision is entered; the date of formal closure after service; the decision type (merits judgment, settlement, dismissal, enforcement order, and so on); the claim value; and whether and when an appeal was lodged within the statutory window.

We restrict the sample to suits with a recorded dispositive decision between January 2010 and December 2016, yielding 4,071,618 cases. We drop the 2.7 percent of records with missing or inconsistent dates (for example, closure before filing) and, where a case has multiple dispositive steps, retain the earliest decision date, which generally coincides with the judge's signature on the written opinion and is the moment that counts toward quota points. For each case we record the day of the month on which the decision falls and define *EndMonth* as an indicator for a decision in the last ten days of its month. Claim values are reported but subject to measurement error and enter our specifications in logs.

To trace decisions into the real economy, we match plaintiff and defendant tax identifiers to Bureau van Dijk's Orbis, which compiles Croatian corporate registries and annual financial statements. We match 1,063,774 cases to at least one corporate defendant filing balance-sheet data between 2008 and 2018, and for these firms extract annual net income, total assets, liabilities, employment, and NACE industry codes. Requiring at least one observation in the

three years before resolution and one in the two years after yields a panel of 248,326 defendant firms.

IV. Empirical Strategy

Our design exploits an exogenous, calendar-driven deadline. The monthly checkpoint at which court presidents assess progress is fixed by the calendar, not by the cases before any judge; judges respond by concentrating closures just before it; and we ask whether the decisions produced in that surge are of lower quality and what they cost the parties. The analysis proceeds in three steps.

A. Bunching at the deadline

We first establish that closures concentrate at month-end, plotting the daily count of dispositive decisions against the day of the month and summarizing the spike by the excess share of closures in the final days relative to the within-month daily average, estimated separately by year, court tier, and case type. The bunching is the behavioral response to the monthly checkpoint and is itself the first finding. We note that it precludes a regression-discontinuity design in “days to month-end”: bunching at the threshold is manipulation of the running variable, the very violation that invalidates RDD, so the discontinuity in closure density is evidence of the incentive rather than a basis for identifying its consequences.

B. Decision speed and appeals

We then compare decisions closed late in the month with those closed earlier in the same month. For the interval between the last hearing and the decision, and for the probability of appeal, we estimate

$$y_{ic} = \beta \cdot \text{EndMonth}_{ic} + X_{ic} \gamma + \lambda_t + \alpha_c + \varepsilon_{ic} ,$$

X includes log claim value, case category, and party roles; λ_t are month-by-year effects; and α_c are court (and, in the within-judge specification, judge) effects. The coefficient β compares otherwise-similar cases by their position in the monthly cycle.

C. Firm outcomes

Finally, we ask what month-end rulings cost defendant firms. Because firm performance is persistent and a single case is one of many shocks a firm faces, we collapse the panel into pre- and post-closure averages following Bertrand, Duflo, and Mullainathan (2004) to avoid serially correlated errors, and estimate the change in log profit as a function of receiving a month-end ruling, interacting an indicator for the post-closure period with the *EndMonth* indicator and absorbing court, category, and party-role effects. A placebo that compares the year before closure with two years before probes pre-existing trends.

D. Identification

Cases decided at month-end are not a random sample, and the question is whether their timing reflects the rushed decision-making we posit or pre-existing differences in the cases. The pattern is hard to square with selection. The natural way to game a count is to clear one's easiest cases at the deadline—but easy cases are appealed less and cost litigants less, so that form of selection would push the estimates toward zero, whereas we find month-end decisions appealed more and followed by larger firm losses. Nor are difficult cases being saved for the deadline: were that so, late-month cases would take longer to resolve, but the interval from final hearing to decision is shorter at month-end (Figure 2), the signature of compressed review rather than an accumulated queue of hard cases.

The estimates are also insensitive to controls. The appeal premium barely moves as month, weekday, week, and year fixed effects are added (Table 1), and the firm-outcome

coefficient is stable across category, court, and party-role fixed effects (Table 2); such stability bounds the scope for unobserved selection to overturn the result (Altonji, Elder, and Taber 2005; Oster 2019). The pattern further holds within judge and within cases where both sides are represented by counsel. To corroborate the mechanism in the decisions themselves, we are assembling text-level measures from published judgments—opinion length, citations to statute and precedent, and disposition type—to test whether late-month opinions are substantively thinner, not merely more often appealed.

V. Results

A. Bunching

Closures rise gradually over the first three weeks of each month and surge in the final days. On the last working days of the typical month, courts close substantially more cases than the within-month daily average—on the order of 60 percent more on the penultimate day and roughly double on the final day. The pattern recurs in every year from 2010 to 2016, in each of the Municipal, Commercial, and County tiers, and across case types (Figure 1).

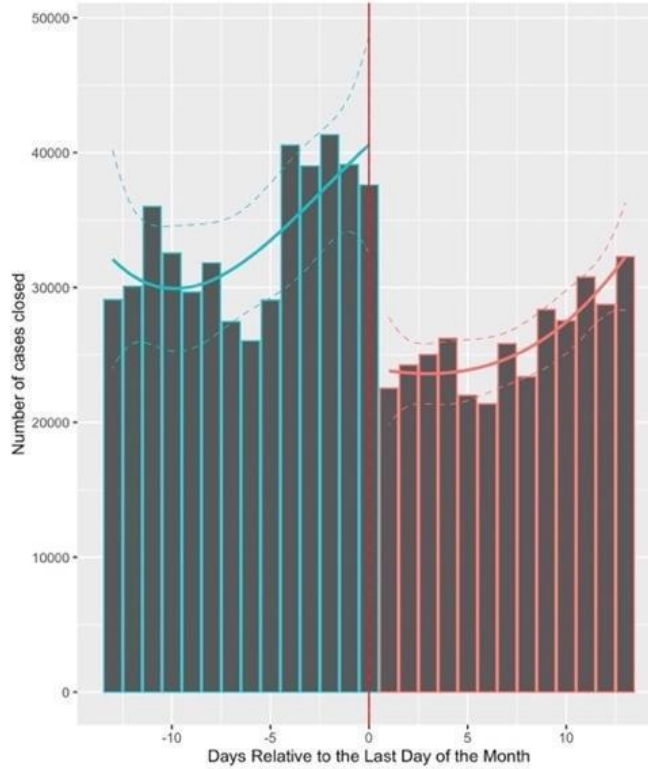


Figure 1. Cases closed by day of the month.

Notes: Counts of dispositive decisions by day of the month, expressed relative to the last day of the month (0 = last day). Pooled across 2010–2016, all court tiers, and case types. The rise into the final days reflects desk-clearing at the monthly checkpoint.

B. Decision speed and appeals

Decisions closed in the final days of the month are finalized faster after the last hearing than decisions closed earlier in the same month, conditional on claim value, court, and month-year (Figure 2). Faster need not mean worse, but late-month decisions are also more likely to be appealed. In a logit for whether a case is appealed, the coefficient on closing at month-end is about 0.09—an increase of roughly 0.4 percentage points on a base appeal rate near 4.3 percent, or about nine percent in relative terms—and it is stable across specifications that add month, weekday, week, and year fixed effects (0.093, 0.093, 0.094, 0.076, 0.093; Table 1, Figure 3). The premium also holds in the within-judge and both-represented subsamples.

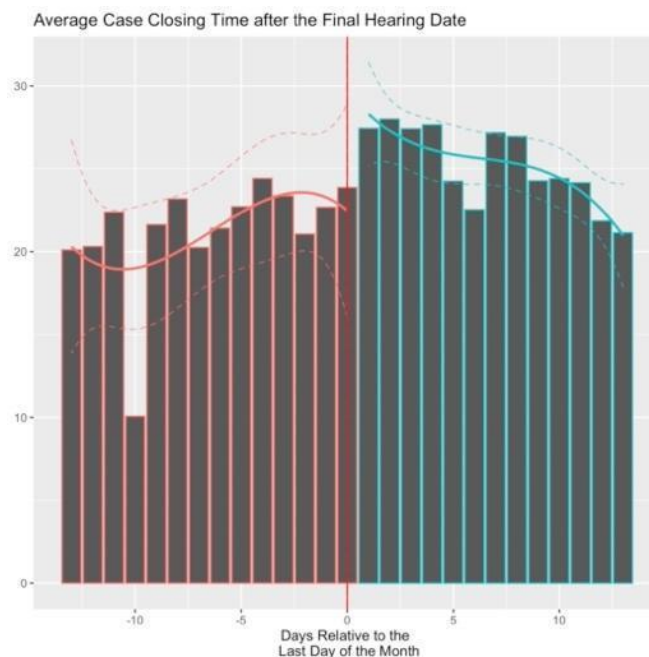


Figure 2. Days from final hearing to closure, by day of the month.

Notes: Mean number of days between the final hearing and case closure, by day of the month relative to month-end (0 = last day). The shorter interval near month-end indicates compressed review under the deadline rather than the accumulation of harder cases.

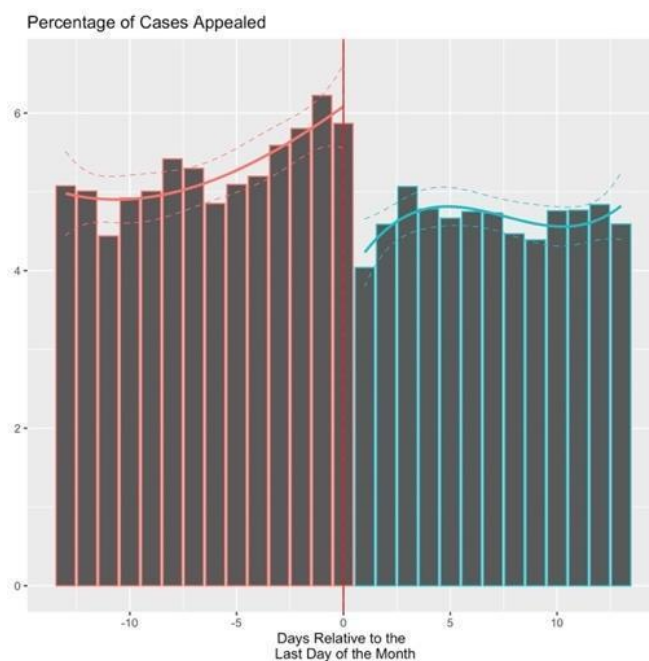


Figure 3. Appeal rate by day of the month.

Notes: Share of dispositive decisions appealed within the statutory window, by day of the month relative to month-end (0 = last day). Appeal rates are higher for cases closed in the final days of the month.

C. Firm outcomes

Among the 248,326 defendant firms we track across closure, those receiving a month-end ruling have lower subsequent profit than comparable firms whose cases closed earlier in the month. The interaction of the post-closure indicator with the month-end indicator is about -0.39 in log profit and is stable across category, court, and party-role fixed effects (-0.392 , -0.368 , -0.397 ; Table 2)—an effect corresponding to roughly a one-third reduction. The placebo on the year before closure relative to two years before (Figure 5) shows no comparable pre-existing differential, supporting the parallel-trends assumption. The estimated effect is larger when the ruling is appealed, consistent with part of the cost operating through prolonged appellate litigation, and survives firm fixed effects and industry-year controls (Figure 4).

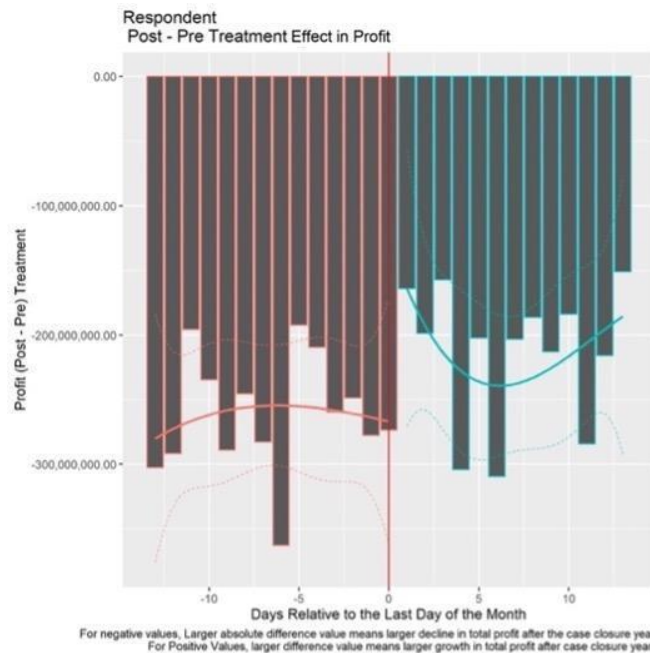


Figure 4. Defendant-firm profit, post-minus-pre change, by day of closure.

Notes: Change in defendant-firm profit from before to after case closure (post minus pre), by day of closure relative to month-end (0 = last day). Larger declines appear for firms whose cases were decided in the final days of the month.

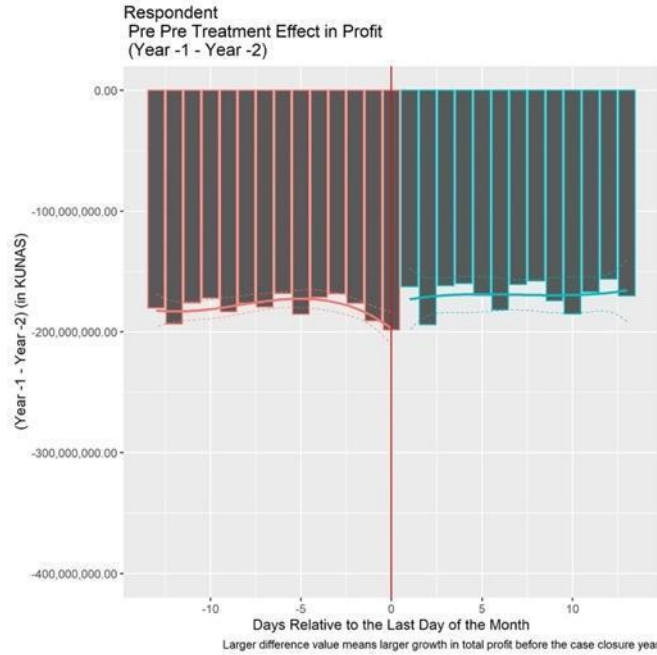


Figure 5. Placebo: pre-closure profit trend, by day of closure.

Notes: Change in defendant-firm profit from two years to one year before closure, by day of closure relative to month-end (0 = last day). The absence of a comparable pre-closure differential supports the parallel-trends assumption underlying the firm-outcome estimates.

VI. Conclusion

Croatia’s experience illustrates a predictable consequence of hard numerical quotas in the judiciary. Faced with annual targets monitored on a monthly cycle, judges bunch case closures in the final days of each month. Those late-month decisions are produced more quickly after the last hearing, are appealed more often, and are followed by lower profitability for the defendant firms they affect—a desk-clearing pattern of the kind Cohen, Gurun, and Li (2021) document for drug regulators, here under a formal quota and traced through to the parties. Because the most natural selection runs against the result, because the hardest cases are decided faster rather than slower, and because the estimates are stable across rich controls, the pattern is hard to explain as anything other than deadline-driven haste.

The policy tension is general. Output benchmarks can shrink backlogs, but when tied to rigid periods they invite haste that erodes quality and imposes costs beyond the courtroom—the

same logic that governs procurement, patent examination, and drug approval. Several designs would weaken the desk-clearing incentive without abandoning accountability for speed. Rolling or quarterly evaluation windows spread the deadline. Complexity-adjusted weights and explicit penalties for reversals realign priorities toward accuracy. And re-review of decisions issued in the deadline window can catch errors before they harden. The broader lesson for institutional design is that quantitative goals work best paired with qualitative review: give agents flexibility in timing, but hold them accountable for errors. Striking that balance is what would let judicial quotas deliver faster service without sacrificing the quality on which the value of courts depends.

References

- Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–184.
- Asch, Beth J. 1990. "Do Incentives Matter? The Case of Navy Recruiters." *Industrial and Labor Relations Review* 43 (3): 89S–106S.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (1): 249–275.
- Buscaglia, Edgardo, and Maria Dakolias. 1999. *Judicial Reform in Latin American Courts*. World Bank Technical Paper 280. Washington, DC: World Bank.
- Carpenter, Daniel, Evan J. Zucker, and Jerry Avorn. 2008. "Drug-Review Deadlines and Safety Problems." *New England Journal of Medicine* 358 (13): 1354–1361.
- Chemin, Matthieu. 2012. "Does Court Speed Shape Economic Activity? Evidence from a Court Reform in India." *Journal of Law, Economics, and Organization* 28 (3): 460–485.
- Cohen, Lauren, Umit G. Gurun, and Danielle Li. 2021. "Internal Deadlines, Drug Approvals, and Safety Problems." *American Economic Review: Insights* 3 (1): 67–82.
- de Figueiredo, Miguel F. P., Alexandra D. Lahav, and Peter Siegelman. 2020. "The Six-Month List and the Unintended Consequences of Judicial Accountability." *Cornell Law Review* 105 (2): 363.
- Di Vita, Giuseppe. 2010. "Judicial Efficiency and Delay in the Italian Courts." *International Review of Law and Economics* 30 (4): 350–358.
- Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2003. "Courts." *Quarterly Journal of Economics* 118 (2): 453–517.
- Frakes, Michael D., and Melissa F. Wasserman. 2017. "Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from Microlevel Application Data." *Review of Economics and Statistics* 99 (3): 550–563.

- Larkin, Ian. 2014. “The Cost of High-Powered Incentives: Employee Gaming in Enterprise Software Sales.” *Journal of Labor Economics* 32 (2): 199–227.
- Liebman, Jeffrey B., and Neale Mahoney. 2017. “Do Expiring Budgets Lead to Wasteful Year-End Spending? Evidence from Federal Procurement.” *American Economic Review* 107 (11): 3510–3549.
- North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics* 37 (2): 187–204.
- Oyer, Paul. 1998. “Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality.” *Quarterly Journal of Economics* 113 (1): 149–185.
- Palumbo, Giuliana, Giulia Giupponi, Luca Nunziata, and Juan Mora-Sanguinetti. 2013. “The Economics of Civil Justice: New Cross-Country Data and Empirics.” OECD Economics Department Working Paper 1060.
- World Bank. 2015. *World Development Report 2015: Mind, Society, and Behavior*. Washington, DC: World Bank.

Tables

Table 1. Probability of appeal and month-end closure

	Dependent variable: case appealed (= 1)				
	(1)	(2)	(3)	(4)	(5)
Case closed end of month	0.093*** (0.006)	0.093*** (0.006)	0.094*** (0.006)	0.076*** (0.006)	0.093*** (0.006)
Month fixed effects	No	Yes	Yes	Yes	Yes
Weekday fixed effects	No	No	Yes	Yes	Yes
Week fixed effects	No	No	No	Yes	Yes
Year fixed effects	No	No	No	No	Yes
Observations	4,181,076	4,181,076	4,181,076	4,181,076	4,181,076
Log likelihood	−784,703	−783,657	−783,462	−782,574	−729,445

Notes: Logit estimates. The dependent variable equals one if the case was appealed within the statutory window. “Case closed end of month” equals one if the dispositive decision falls in the last ten days of the month. The estimation sample is all civil and commercial suits with a dispositive decision between 2010 and 2016; the base appeal rate is approximately 4.3 percent, so the coefficient of about 0.09 corresponds to roughly a nine percent relative increase. The coefficient is stable as month, weekday, week, and year fixed effects are added. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2. Defendant-firm profit and month-end closure

	Dependent variable: $\log(\text{profit} + 1)$		
	(1)	(2)	(3)
After case closure	0.397*** (0.016)	0.377*** (0.016)	0.403*** (0.016)
After case closure $\times$ end of month	−0.392*** (0.022)	−0.368*** (0.021)	−0.397*** (0.021)
Case-category fixed effects	Yes	No	Yes
Court fixed effects	No	Yes	Yes
Party-role fixed effects	Yes	Yes	Yes

Observations	264,417	287,505	264,417
R-squared	0.237	0.240	0.274

Notes: OLS estimates. The dependent variable is the log of one plus the defendant firm's profit. "After case closure" equals one for firm-years following the case's closure; its interaction with end-of-month closure is the difference-in-differences coefficient of interest, on the order of -0.39 log points (roughly a one-third reduction) and stable across fixed-effect specifications. The estimation sample is defendant firms matched to Orbis with balance-sheet data in the three years before and two years after closure. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.