Information Frictions and Court Performance: Experimental Evidence from Chile[1]

Paloma Carrillo*, Daniel Chen†, Manuel Ramos-Maqueda⋆,
and Bernardo Silveira∘

September 2026

Abstract

This paper examines how reducing information frictions can improve public-sector performance. In Chilean family courts, managers have access to administrative data but rarely use them effectively. In a randomized trial, we tested three low-cost interventions: email nudges, simplified feedback, and peer-comparison feedback. Courts receiving the simplified feedback resolved cases faster: case clearance rates rose and case duration fell, with no detectable decline in case quality. The peer-comparison feedback produced similar gains on several margins, while the email nudges showed suggestive evidence of shifting attention toward longstanding cases. These results show that even modest information tools can substantially enhance institutional efficiency.

JEL codes: D73; D83

Keywords: Information frictions, court efficiency, and court managers.

∗ International Labour Organization; email: carrillo@ilo.org

† Toulouse School of Economics; email: daniel.chen@iast.fr

⋆ World Bank; email: mramosmaqueda@worldbank.org

∘ University of Zurich; email: bernardo.silveira@econ.uzh.ch

[1]We thank the Department of Institutional Development in the Chilean Judiciary's Administrative Corporation (CAPJ) for the partnership in this project. We would like to thank in particular Esteban Paiva, Mario Lara, Pablo Cabezas and Sebastián Montero for their excellent support and coordination throughout the project. Nithya Nagarathinam and Viknesh Nagarathinam provided excellent research support. We would also like to thank Cristián Eyzaguirre and audience participants at CELS and at the Rule of Law and Democracy: Innovations and Challenges Symposium for helpful comments and suggestions. This project received IRB Approval from Toulouse School of Economics on January 18, 2020. The pre-analysis plan is registered in the AEA RCT Registry at https://www.socialscienceregistry.org/trials/5512 and a copy of it can be found in Appendix A.1 with further explanations. All errors are our own.

1 Introduction

Despite the growing abundance of administrative data in public institutions, effectively harnessing these resources for evidence-based decision-making remains a significant challenge. Recent literature highlights a notable gap in data-management and analytical skills within the public sector, which inhibits the full utilization of available data (OECD, 2014; Savoldelli et al., 2014). This gap underscores the urgent need for strategies that can alleviate such constraints and enable institutions to capitalize more effectively on their data resources.

In this context, the implementation of performance measures in public administration offers a promising avenue. These measures, designed to assess and enhance the efficiency, effectiveness, and overall productivity of public services, have been increasingly recognized for their potential to inform and refine management and policy decisions (Ostrom and Hanson, 2010; Hanson and Ostrom, 2014; Durham and Becker, 2016). However, their adoption remains limited, often constrained by managerial attention (Gabaix, 2014) and analytical capacity (Kleiman, 2009). Our study addresses this gap by examining how simplifying the presentation of performance data in the judiciary can reduce informational frictions, thereby enhancing the use and impact of such data in judicial decision-making.

This paper investigates how information frictions affect the management of Chilean family courts. We first show that, even when performance data are available, many court managers have inaccurate beliefs about key metrics, such as case clearance rates. Using a randomized controlled trial, we then find that low-cost informational interventions can meaningfully improve performance. By presenting performance data in a clear and digestible format, these interventions prompt managers to adjust their daily decisions, leading to faster case resolution, better motion processing, and higher clearance rates—without any detectable decline in the quality of judicial outcomes.

Chilean family courts provide an ideal setting for our analysis. They handle cases related to child custody and alimony, gender-based violence, and divorce. Delays in resolving these matters can have serious consequences for vulnerable individuals, such as children and at-risk women. For example, Cooper et al. (2023) finds that, in Chile, prolonged foster care proceedings increase youth involvement in crime and reduce school attendance. To improve court efficiency, the Chilean Judiciary created a new administrative position in family courts in the early 2000s: the court manager. Court managers plan, organize, and monitor day-to-day court operations and supervise between 40 and 60 staff members each.

According to a survey we conducted in 2020, 70% of managers agree or strongly

agree that tracking court indicators is among the most important tasks of their job (see Section 3.4 for methodology and results). Yet our evidence suggests that managers are often unfamiliar with their own court's performance. When asked to recall their court's case clearance rate—the primary indicator of court speed—over the preceding two months, 31% of family court managers were off by more than 25% from the actual value. This lack of familiarity persists despite the availability of accurate data. Since 2018, managers have had access to these measurements through an online court management platform called Quantum. While 71% of managers agree or strongly agree that Quantum is useful for their work, platform usage is limited: between July 2018 and August 2019, the median number of logins per manager was six, and over 35% of managers never logged in.

To address the challenges of underused administrative data, we partnered with the Department of Institutional Development of the Chilean Judiciary to conduct a randomized controlled trial in family courts. The experiment tested three types of low-cost informational interventions. In the first one, the *email promotion* intervention, we employed nudges encouraging managers to engage with the existing online court management platform. In a second intervention, which we refer to as the *simplified feedback*, we redesigned the platform to make performance information more accessible and easier to interpret for court managers. Finally, a third intervention, named *peer comparison feedback*, combined the redesigned platform with comparisons to peer courts, leveraging benchmarking and social norms. Our primary goal was to assess how these interventions influenced court productivity, while also exploring the behavioral channels through which managers respond to information.

Courts receiving the simplified feedback dashboard resolved cases faster, with case clearance rates rising by about 0.36 standard deviations. Case-level data show the same pattern: the hazard rate of resolution for incoming cases rises by roughly six percent. Other metrics more directly under managerial control, such as timely motion resolution and hearing programming, also improved. The peer comparison treatment showed directionally positive but statistically insignificant effects on case clearance, alongside comparable gains in case duration.

The email promotion treatment modestly increased platform logins and drew managers' attention to longstanding cases. Although smaller in magnitude, these effects complemented the feedback interventions, demonstrating that even modest nudges can encourage engagement with performance data and reinforce improvements in court management.

The evidence suggests that these interventions worked primarily by making key information more accessible and actionable. The feedback interventions helped managers

focus on the most relevant aspects of daily operations, while email nudges encouraged attention to older pending cases. Together, the results highlight that small, low-cost informational improvements can meaningfully enhance managerial decision-making and operational efficiency without undermining the quality of judicial outcomes.

This paper contributes to the growing literature on mitigating information frictions in public sector management, a field that has gained momentum following substantial investments in high-frequency data collection. A central focus of this work has been the use of informational dashboards, examined across social protection (Dodge et al., 2021), healthcare (Callen et al., 2020; Whidden et al., 2018), civil service (Mattsson, 2021), education (Dizon-Ross, 2019), procurement (Celhay et al., 2024), and taxation (Pomeranz, 2015). Our study is unique in examining dashboards within the judicial system—a sector traditionally resistant to reform and empirical evaluation (Engel, 2013; Greiner and Matthews, 2016).

In that sense, our work connects to a growing body of research on interventions designed to improve court efficiency. Whereas our focus is on internal information tools, other recent efforts have sought to enhance judicial performance by engaging the broader public. For example, Sadka et al. (2024) and Chemin et al. (2023) show that public-oriented strategies can strengthen court functioning in Mexico and Kenya, respectively. Along similar lines, Kondylis and Stein (2023) examine procedural reforms in Senegal and find that incentives for faster resolution of civil and commercial cases can meaningfully stimulate economic activity. Our study complements these contributions by demonstrating the potential of internally focused information tools to improve the day-to-day workings of the judicial system.

More broadly, our findings speak to how internal government processes shape the quality of public service delivery. We focus on low-cost, technologically modest interventions in court administration, paralleling recent work on governance improvements through employee monitoring (Banerjee et al., 2008; Dhaliwal and Hanna, 2017), phone surveillance (Muralidharan et al., 2021), and bespoke mobile applications (Callen et al., 2020; Dodge et al., 2021). Unlike these approaches, which often rely on the introduction of new technologies, our intervention leverages existing infrastructure. This strategy minimizes fixed and ongoing costs, scales readily across settings, and allows rapid adaptation to evolving challenges, while avoiding the steep learning curve typically associated with new technology.

2 Court Managers and Quantum Platform

Court managers (*administradores de tribunal*) primarily facilitate the administrative functioning of the court. The position was introduced as part of Chile's procedural reforms beginning in the early 2000s, which created hearing-based procedures and reorganized courts so that judges focus on adjudication while a professional manager oversees day-to-day operations.[2] Each manager oversees one court and roughly 40–60 employees, planning, organizing, directing, and monitoring court operations and optimizing staff performance (Chilean Judiciary, 2020). Managers can adjust agendas and redistribute tasks in response to case inflows (see Table A2). The role is stable: many managers in our sample have held it since its inception (see Section 3).

Court managers' responsibilities sit within a broader internal accountability structure. Managers undergo annual performance evaluations conducted by the President of the relevant Court of Appeals, based on (i) a report from the Judges' Committee of the manager's court and (ii) a separate report from the Administrative Corporation of the Judiciary (CAPJ) (Código Orgánico de Tribunales, Articles 273–278). In addition, the judiciary sets annual court-level management targets, covering indicators such as case clearance rates and hearing scheduling, and meeting these targets qualifies court staff—including the manager—for an institutional performance bonus. Importantly, neither the annual rating nor the bonus is mechanically tied to any single indicator: evaluations rely on supervisory judgment and multiple information sources rather than automated KPIs. Thus, while court managers are evaluated holistically, they have real incentives to care about operational performance.

In July 2018, the Department of Institutional Development launched the online Quantum platform, providing court managers and staff with comprehensive performance metrics, including cases filed, case clearance rates, average durations, and hearing schedule adherence. The platform covers all courts and aggregates data monthly and yearly, allowing cross-court comparisons.

Despite Quantum's capabilities, take-up has been limited. Between July 2018 and August 2019, the median number of logins per manager was only six, and over 35 percent of managers never logged in during this period. Although the platform is technologically sophisticated and information-rich, it is unclear whether it has affected

[2]The separation of adjudication from court administration is not unique to Chile. Many judicial systems rely on professional court administrators to run scheduling, staffing, and caseflow operations (e.g., the United States, the UK and Australia, as well as some Latin American countries such as Ecuador, Panama or Costa Rica). Where a dedicated manager role does not exist, these managerial functions are often performed by judges or senior clerks/registrars. Chile institutionalized this separation explicitly within reformed first-instance courts (including family courts), where high volumes and dense hearing calendars make operational management particularly salient.

court management. A baseline survey conducted in early 2020 (see Section 3.4) found that over 90 percent of managers were aware of Quantum. Yet one-third of respondents reported that Quantum lacked relevant information or features, even though it provides a vast amount of data. This pattern suggests that the platform's design may hinder engagement.

One key challenge is that managers without a quantitative background may find Quantum overwhelming or unclear. Users encounter a dense homepage (Figure A2, Appendix A.5), making it difficult to locate disaggregated data or specific metrics. In other words, it is easy to get lost on the platform. The next section describes our intervention to make Quantum's information more accessible and digestible.

3 Intervention and Data

3.1 Experimental Design

Our experiment focuses on 49 family courts in Chile, each overseen by a court manager.[3] As of 2019, these courts employed 1,697 staff, including 254 technical advisers, 1,106 administrative officers, 85 heads of unit, and 252 judges. In 2021, they handled 270,952 new cases. The most common case types were contested proceedings (26.6%), domestic violence (20.8%), protection measures (18.5%), enforcement procedures (16.0%), and mediation (12.6%).[4]

The first intervention we consider—the *email promotion*—encouraged managers to use Quantum via targeted emails, sent to treated individuals with a monthly frequency for the duration of the experiment. A second type of intervention—the *feedback intervention*—redesigned Quantum's homepage to display a simplified dashboard summarizing key performance statistics. The dashboard streamlined the original table of monthly and cumulative statistics—which included information on incoming and terminated cases, hearings, protection measures, legal writings, resolutions, proceedings,

[3]As outlined in the pre-analysis plan, the intervention was originally intended for all 346 Chilean courts with court managers. Due to institutional constraints, implementation was restricted to family courts. Of the 60 family courts, 10 were excluded because they helped design the simplified Quantum homepage in 2020, and one was excluded due to missing performance data.

[4]These categories correspond to procedural tracks in the judiciary's case-management system (SITFA), denoted by the letter that precedes the court's internal docket number (RIT). In family courts, contested proceedings (RIT C) comprise ordinary contentious litigation (e.g., child support, custody, visitation, divorce-related matters); domestic violence cases (RIT F) refer to cases of intra-family violence; protection measures (RIT P) are child-protection proceedings triggered by serious threats or violations of children/adolescents' rights; enforcement procedures (RIT Z) are enforcement/compliance matters (often enforcement of child-support obligations or other prior orders); and mediation (RIT M) refers to the pre-filing mediation track, mandatory for child support, custody, and visitation disputes.

notifications, and customer service contacts—and added statistics in percentage or rate terms (see Figure A2 and Figure A3 in Appendix A.5 for a comparison). We refer to the baseline version of the feedback intervention as *simplified feedback*. Additionally, some managers were exposed to an alternative version of the feedback intervention—the *peer comparison feedback*—which also included a pop-up showing each court's best and two worst indicators relative to a peer court, emphasizing relative strengths and weaknesses.[5] This pop-up was updated on a monthly basis. At each monthly update, the pop-up recomputed the court's one best and two worst indicators relative to the assigned peer based on the largest gaps, so the highlighted strength and weaknesses could change from month to month as relative performance evolved.

To administer the interventions, we randomly assigned the 49 courts into six treatment arms: (C) control; (T1) email; (T2) simplified feedback; (T3) email + simplified feedback; (T4) peer comparison feedback; and (T5) email + peer comparison feedback. Randomization was stratified by 2019 incoming caseload, a proxy for court size, yielding eight courts in the control group, eight in T1, nine in T2, eight in T3, nine in T4, and seven in T5. The factorial design allows separate estimation of feedback and email effects: 16 courts received no feedback, 33 received simplified or peer comparison feedback; 23 received email promotion, 26 did not.[6] Given the sample size and study period, we can detect effects starting at 0.4–0.6 standard deviations—see Appendix A.4 for details on the power calculation. Table 1 summarizes the treatment categories described in this section, which are used throughout the paper.

The experiment ran from January 26 to June 2021. On January 26, all 49 managers received an introductory email explaining how to access Quantum and retrieve forgotten passwords. Managers in Treatments 2–5 (feedback intervention—either the simplified or peer comparison variety) were additionally informed about homepage changes. Managers in Treatments 1, 3, and 5 (email promotion) received three monthly emails (Feb 22, Mar 22, May 25[7]) encouraging use of Quantum. Emails were personalized based on prior login activity, congratulating top users or reminding inactive managers of the benefits of tracking court performance. All emails highlighted positive user feedback, the link between Quantum use and higher case clearance rates, and included a prominent link and button to increase platform accessibility. Control group managers experienced no homepage changes and received no further Quantum

[5]Peer courts are designated by the Chilean Judiciary based on competence (family court) and size, proxied by annual incoming cases. Pairing is one-to-one and reciprocal. The assigned peer court remained fixed throughout the intervention period.

[6]The design was originally conceived for over 300 courts, but institutional constraints limited the sample to 49 (see Footnote 3). The main analysis (Section 4) separately estimates feedback and email effects, while Appendix A.6.2 reports full treatment interactions.

[7]See Appendix A.3 for full versions.

promotion, continuing to use the existing table-based interface. All estimated effects should therefore be interpreted relative to this format, which is common in many countries' court dashboards.[8] This makes the results informative for settings where performance information exists but is delivered in dense, reporting-centric formats rather than streamlined dashboards with trends and peer comparisons.

Table 1: Treatment Category Labels

	C	T1	T2	T3	T4	T5
Email		✓		✓		✓
Simplified feedback			✓	✓		
Peer comparison feedback					✓	✓
Feedback			✓	✓	✓	✓
Email × Feedback				✓		✓

Notes: (C) control; (T1) email; (T2) simplified feedback; (T3) email and simplified feedback; (T4) peer comparison feedback; and (T5) email and peer comparison feedback.

3.2 Conceptual Framework

Court managers continually allocate limited time and attention across competing operational tasks, such as scheduling hearings, monitoring pending cases, reassigning staff, troubleshooting bottlenecks, and coordinating with judges, clerks and other court actors. These decisions are costly: they require effort, follow-up, and often generate internal friction. At the same time, managers have clear reasons to care about operational performance, considering their annual performance evaluations and those of their own courts (including performance bonuses). This setting implies a central tension: while managers have incentives to improve performance, translating those incentives into outcomes is constrained by information frictions (e.g., knowledge about absolute and relative standing), and by the attention and coordination costs of acting on that information.

Our interventions alter the manager's decision environment by changing: (i) exposure to performance information, (ii) the ease of diagnosing operational problems, and (iii) the salience of specific benchmarks. We outline these mechanisms below, while emphasizing they are non-exclusive.

From interventions to information. The most immediate effect of the interventions is on the information managers acquire about their courts, which can occur through

[8] European Commission for the Efficiency of Justice , CEPEJ, describes the "simple grid/data table" (p.11) as the most common way to present many indicators in one snapshot, and notes its familiarity and ease of development. Similar designs are used in the Dominican Republic, Brazil, or India's National Judicial Data Grid, among others.

two pathways. First, the interventions may increase the frequency with which managers access Quantum. This is the explicit goal of the email promotion, but it could also arise from the feedback treatments: being informed of the redesigned homepage may prompt logins, and exposure to a more usable interface may encourage repeat visits. Because our data include individual-level login activity, we can directly assess these effects. Second, conditional on accessing Quantum, the feedback interventions may reduce the cost of obtaining and interpreting key performance information. By reorganizing content around a small number of operationally relevant indicators, the redesigned interface can make it easier for time-constrained managers to diagnose where performance is lagging and to track whether corrective actions are working. Thus, even if logins were not to increase, reducing the costs of processing information and diagnosing key metrics can still shift decisions by making available information more usable.

From information to action. How does increased access to performance information translate into changes in managerial behavior? Our preferred interpretation is that the interventions relax information and attention constraints that affect whether managers' incentives translate into performance.

(i) Simplified feedback primarily reduces information-processing costs and uncertainty about current performance. In complex organizations, managers often have sufficient incentives but insufficient clarity about which operational margins matter most or where the binding bottlenecks lie. Presenting a small set of key metrics more prominently can redirect attention toward high-return actions (e.g., enforcing scheduling routines or reallocating administrative effort toward the stages where delays accumulate.) This mechanism is consistent with evidence that simplifying information and making key metrics salient can improve decisions in complex environments (Thaler and Sunstein, 2008; Chetty et al., 2009).

(ii) Email promotion primarily addresses inattention and inertia. Even when data are accessible, day-to-day operational pressures can crowd out monitoring activity. Reminder emails can increase the likelihood that managers check the platform. This treatment should therefore be most directly linked to intermediate engagement outcomes (logins) and may translate into performance effects when the information accessed is already sufficiently interpretable.

(iii) Peer comparison feedback adds an additional reference-point channel beyond simplification. Relative performance information can influence managers' choices even holding formal incentives constant, by activating professional norms and reputational concerns (Ashraf et al., 2014; Allcott and Rogers, 2014), and by shifting beliefs about what is feasible with existing resources. To keep the message actionable under limited

attention, the pop-up summarizes comparisons using one relative strength and two relative weaknesses. This design emphasizes the largest shortfalls relative to a peer court while still including a positive indicator to avoid a purely negative frame that could discourage engagement. Because the pop-up selectively highlights only a few indicators, the peer comparison arm also creates scope for heterogeneous effects across outcomes, a mechanism we discuss below under *outcome heterogeneity.*

Organizational interactions. A further mechanism arises from the relationship between managers and judges. As we document in Section 3.4, managers frequently cite judicial interference in administrative affairs as the main obstacle to effective court management. Clear, court-specific performance metrics—especially when prominently displayed—may strengthen managers' bargaining position when negotiating for autonomy over administrative decisions, by providing a shared factual basis for identifying problems and prioritizing operational changes. At the same time, enhanced performance visibility could also be perceived as enabling monitoring by superiors, which may affect managerial autonomy and effort in either direction—motivating effort through accountability or discouraging initiative if perceived as punitive (Humphrey et al., 2007; Rasul and Rogger, 2018).

Outcome heterogeneity through attention channeling. Finally, the interventions may generate heterogeneous effects across outcomes via attention channeling. In the peer comparison arm, the pop-up highlights each court's single best-performing and two worst-performing indicators relative to a peer court, mechanically directing attention to a subset of metrics. If managers respond by prioritizing the highlighted margins, this can lead to differential improvements across indicators. Evidence on task selection suggests that individuals allocate effort toward salient and feasible tasks (Amar et al., 2011), while studies of dashboard design show that presentation choices shape which information users attend to (Van der Heijden, 2013). This mechanism can explain patterns where some performance measures improve while others remain unchanged, consistent with recent evidence on performance feedback in public service delivery (Mattsson, 2025).

Summary. While our data does not allow us to fully disentangle the relative contribution of each mechanism, the framework clarifies what each treatment is designed to capture: email reminders primarily target inattention and inertia; simplified feedback primarily reduces processing and diagnostic costs; and peer comparison adds peer benchmarks that can shift norms and beliefs about feasible performance. Interpreting treatment effects through this lens helps explain how low-cost informational interventions can strengthen the link between managers' existing incentives and court performance.

3.3 Data

3.3.1 Administrative and Case-Level Data

Unless otherwise specified, we obtained data directly from the Chilean Judiciary's Administrative Corporation (CJAC). Our main court performance measures span five monthly indicators, covering four months before the intervention through its fifth month. Computed by the CJAC and fed into Quantum, these indicators are: (1) *case clearance rate* – resolved cases divided by incoming cases; (2) *timely motion resolution rate* – motions signed within three working days divided by total motions signed[9]; (3) *hearing schedule adherence rate* – hearings held on their originally scheduled date and time, divided by total scheduled hearings; (4) *timely case resolution rate* – cases with entry dates under two years divided by total inventory; and (5) *rate of timely hearing programming* – hearings scheduled within 70 days divided by total scheduled hearings.[10] These measures capture performance across different case components and horizons, with higher values indicating faster case processing and, all else equal, greater judicial efficiency.

We complement the CJAC's court-month data with publicly available case-level microdata from *Judicial Power in Numbers*,[11] which links each case to a unique court (and its manager). Only the case clearance rate can be reconstructed at the case level, and it closely matches the official indicator (correlation 0.996). The microdata also allow detailed analysis of incoming and resolved cases, average duration, and hearing scheduling. Table 2 reports descriptive statistics for the 129,887 cases opened during the treatment. By December 31, 2022, 23% remained ongoing; the remaining 77% concluded in ten ways, most commonly judgment, followed by cases where the demand was not accepted, and mediation. Conditional on conclusion, the average duration was 72.4 days. Across these cases, 73,640 hearings were held, with a mean scheduling time of 47.08 days.

We also include in our analysis pre-treatment court characteristics and performance indicators from 2019: number of incoming cases, number of hearings, case clearance, and inventory older than one and two years. Manager-level information includes age, sex, and tenure. Tables 3 and 4 provide descriptive statistics for the treatment and

[9] A motion is a written request to the court for an order, ruling, or direction, typically filed according to the type of case.

[10] Four of these indicators correspond to pre-registered primary outcomes; hearing schedule adherence was added because it is computed by the CJAC, displayed prominently in Quantum, and directly relevant to understanding managerial responses to information about hearing scheduling (see Appendix A.1 for details on deviations from the pre-analysis plan).

[11] In Spanish, *Poder Judicial en Números*. Available at `https://numeros.pjud.cl/Competencias/Familia`.

Table 2: Case-level data for cases filed during the treatment period

Hearings during treatment period		
Days to schedule (mean, SD)	47.08	30.9
Resolved cases		
Duration in days (mean, SD)	72.4	102.2
Resolution type (including ongoing cases)		
Judgment	39,554	30.5%
Ongoing	29,776	22.9%
Motion to dismiss in limine	20,833	16.0%
Mediation	14,682	11.3%
Incompetent court	8,194	6.3%
Conciliation	7,317	5.6%
Involuntary dismissal	2,552	2.0%
Settlement	2,020	1.6%
Voluntary dismissal (no resub.)	2,018	1.6%
Settlement (II)	1,762	1.4%
Voluntary dismissal (resub. ok)	1,179	0.9%

Notes: Summary statistics based on 129,887 cases filed during the treatment period, of which 100,111 (77%) were resolved by December 31, 2022.

control groups—see Section 3.3.2 for a discussion of the balance of our treatment assignment. About two-thirds of managers are men, with an average age of 52 and 12 years in their position.

Additionally, we have survey data on managers' perceptions of their roles and Quantum, collected in January 2020 (see Section 3.4). We also track logins to Quantum one month before and during the five-month intervention, including exact login date and time, but not session duration or pages visited. In the pre-treatment month, 33% of managers logged in at least once. A summary of all data sources is provided in Table 5.

3.3.2 Balance Check

Tables 3 and 4 report baseline values for the feedback and email promotion treatment arms. We consider five manager characteristics, nine annual court indicators from 2019, five monthly pre-treatment indicators from October 2020 to January 2021, and three monthly aggregated case-level statistics.

Before the treatment, the groups were similar on most variables. In the feedback arms, only three of the 22 variables show statistically significant differences: the monthly resolution rate is higher, while the hearing schedule adherence rate and the number of hearings held are lower in the treatment groups (all significant at the 1% level, except for case resolution rate in the simplified feedback group, which is significant at 5%). In contrast, for the email arm, none of the differences are significant at conventional levels, suggesting that random assignment was largely successful.

Table 3: Court Characteristics for No-Feedback and Feedback at Baseline

	No Feedback	Simplified Feedback	Peer comp. Feedback	Difference (1-2)		Difference (1-3)	
	(1)	(2)	(3)	Diff	SE	Diff	SE
Panel A. Manager's Characteristics							
Age	52.06	52.29	51.69	-0.23	(2.24)	0.38	(2.09)
Men	0.69	0.76	0.56	-0.08	(0.16)	0.13	(0.18)
Years in position	12.31	11.82	12.81	0.49	(1.81)	-0.50	(1.91)
January Quantum logger	0.31	0.35	0.31	-0.04	(0.17)	0.00	(0.17)
Number of logins in January	0.94	0.76	0.94	0.17	(0.45)	0.00	(0.55)
Panel B. Court Indicators (2019)							
2019 incoming cases	8799.19	9636.00	8197.63	-836.81	(1953.32)	601.56	(1746.75)
2019 hearings	5320.13	5788.24	4730.38	-468.11	(1173.53)	589.75	(1033.33)
2019 case clearance	75.54	73.59	75.04	1.95	(3.25)	0.50	(2.75)
2019 inventory older than 1 year	0.32	0.70	0.69	-0.39	(0.28)	-0.37	(0.30)
2019 inventory older than 2 years	0.06	0.23	0.11	-0.17	(0.14)	-0.06	(0.07)
2019 optimal number staff gap	-0.50	0.24	-0.44	-0.74	(0.50)	-0.06	(0.56)
2019 optimal number judge gap	0.31	0.76	0.55	-0.45	(0.38)	-0.24	(0.35)
Observations	16	17	16	33	33	32	32
Panel C. Court Indicators (Oct–Jan)							
Timely motion resolution rate	96.13	97.44	97.63	-1.32	(0.52)	-1.50	(0.54)
Timely hearing programming	56.05	63.72	62.52	-7.67	(5.03)	-6.47	(4.89)
Hearing schedule adherence rate	69.16	61.75	58.16	7.41	(2.72)	11.00	(2.61)
Case clearance rate	70.39	69.35	71.69	1.04	(2.07)	-1.30	(2.11)
Timely case resolution rate	92.16	91.53	91.72	0.63	(2.16)	0.44	(2.15)
Panel D. Micro Data (Oct–Jan)							
Number of hearings	289.00	262.03	207.30	26.97	(27.60)	81.70	(26.39)
Number of incoming cases	557.06	589.15	491.48	-32.08	(65.00)	65.58	(57.52)
Number of concluded cases	376.05	395.13	352.05	-19.09	(40.93)	24.00	(38.90)
Observations	64	68	64	132	132	128	128

Notes: The no-feedback group includes those in the control group (C) and those receiving the email promotion (T1). The January Quantum logger is a dummy variable equal to one if the court manager logged into Quantum at least once in January 2021. The optimal number of staff and judge gaps refer to indicators built by the courts in order to understand the deficit or surplus in the number of judges and staff given the workload of the court. The number of years in the position of court manager has an upper bound of 16 years when the position was first created. 26 of the 49 court managers have been in the position since its inception.

Table 4: Court Characteristics for No-Email and Email at Baseline

	No Email (1)	Email (2)	Difference (1-2) (3)	SE (4)
Panel A. Manager's Characteristics				
Age	51.35	52.78	-1.44	(1.77)
Men	0.62	0.74	-0.12	(0.14)
Years in position	11.85	12.83	-0.98	(1.51)
January Quantum logger	0.35	0.30	0.04	(0.14)
Number of Quantum logins in January	0.85	0.91	-0.07	(0.39)
Panel B. Court Indicators (2019)				
2019 incoming cases	9074.15	8688.39	385.76	(1566.84)
2019 hearings	5347.85	5224.52	123.32	(902.47)
2019 case clearance	73.39	76.18	-2.79	(2.37)
2019 inventory older than 1 year	0.67	0.46	0.21	(0.26)
2019 inventory older than 2 years	0.18	0.08	0.10	(0.10)
2019 optimal no. staff gap	-0.27	-0.17	-0.10	(0.43)
2019 optimal no. judge gap	0.46	0.63	-0.16	(0.30)
Observations	26	23	49	49
Panel C. Court Indicators (Oct–Jan)				
Timely motion resolution rate	97.28	96.84	0.44	(0.39)
Timely hearing programming	58.72	63.20	-4.47	(4.12)
Hearing schedule adherence rate	62.63	63.40	-0.77	(2.32)
Case clearance rate	70.27	70.66	-0.39	(1.65)
Timely case resolution rate	93.00	90.43	2.57	(1.76)
Panel D. Micro Data (Oct–Jan)				
Number of hearings	267.33	236.73	30.60	(21.59)
Number of incoming cases	552.95	539.80	13.15	(50.96)
Number of concluded cases	380.24	368.72	11.52	(33.37)
Observations	104	92	196	196

The no-email group includes those in the control group (C) and those receiving the feedback treatment (T2 and T4). The January Quantum logger is a dummy variable equal to one if the court manager logged into Quantum at least once in January 2021. The optimal number of staff and judge gaps refer to indicators built by the courts to capture the deficit or surplus in the number of judges and staff given the court's workload.
†The number of years in the position of court manager has an upper bound of 16 years, when the position was first created. 26 of the 49 court managers have been in the position since its inception.

Table 5: Data Sources

Source	Timeline	Usage
CJAC monthly indicators	Oct 2020–Jun 2021	Five monthly court indicators, used for pre-treatment controls and main outcome variables.
CJAC case-level data	Jan 2019–Dec 2022	Case clearance, case duration, hearings, and case resolution type.
CJAC yearly indicators	2019	Pre-treatment balance check and number of cases (used as strata variable).
CJAC personnel data	Oct 2021	Court manager characteristics (age, sex, and years of experience) used as covariates.
Court manager survey	Jan 2020	Managers' perceptions of their positions and the Quantum platform.
CJAC Quantum login data	Jul 2018–Aug 2019; Jan 2021–Jun 2021	Court manager login activity.

Nevertheless, to account for the few variables that differ significantly in the feedback arm—and for any other imbalances that may be substantively important but imprecisely estimated given the sample size—we include as covariates the court manager's characteristics, all pre-treatment monthly indicators, and the pre-treatment number of hearings in our main regression specifications (Section 5). We also present results without these controls as a benchmark.

3.4 Survey Data: Manager's Perception of their Roles and Use of Quantum

We administered a baseline survey to all 346 court managers—including those in family courts targeted by our intervention—to document their perceptions of their roles and the Quantum platform.[12] The survey was sent in January 2020, about a year before the intervention, and included three modules on managerial responsibilities, court performance, and knowledge and use of Quantum, with the third module administered to a random subset. After standard data cleaning (see Appendix A.2), the final sample comprised 116 responses for the first two modules and 97 for the third, representing 34% of all courts and 53% of intervention courts.

Surveyed managers reported that knowledge of court performance indicators is important: 70% agreed that tracking indicators is one of their most important tasks (Table A1). Indicators are used mainly to guide the distribution of workload and improve hearing schedules (Table A2).

However, many managers lacked accurate information about their courts' performance. When asked to recall their own court's case clearance rate from two months prior, 56% were off by more than 25%, with the error rate among family court managers at 31% (Figure A1). It should be noted that accurate information on case clearance—as well as other relevant court performance statistics—is readily available to all managers through Quantum. However, the survey indicates that the platform is underutilized: among respondents who reported accessing Quantum, 69% denied relying on it for decisions, often citing a lack of relevant statistics.

Finally, the survey revealed tensions between managers and judges over administrative responsibilities, with power dynamics being the most commonly cited obstacle for improved court functioning (26%). Managers noted that judges often intervene in administrative matters and resist direction. In open-ended questions about challenges for improving their performance at work, respondents stated that "court administration

[12]There are more courts than managers because some civil courts had not been reformed by the time of the survey.

continues to face barriers and limitations in carrying out its functions effectively, as it remains subordinate to judges in various administrative areas," and complained about "the involvement of judges in decisions that fall within the administrative domain."

4 Empirical Strategy

4.1 Main Specification

To assess the impact of the email promotion and feedback treatments on court efficiency, we estimate the following specification:

$$\begin{aligned} Y_{nct} = \alpha_{n0} + \alpha_{n1} Email_c + \alpha_{n2} Simplified\ Feedback_c + \alpha_{n3} Peer\ Comp.\ Feedback_c \\ + \alpha_{n4}\mathbf{X}_c + \alpha_{n5}\mathbf{Y}_{nct_0} + \alpha_{n6}\mathbf{H}_{nct_0} + \alpha_{n7}\gamma_t + \xi_{nct}, \end{aligned} \tag{1}$$

where Y_{nct} denotes indicator n, which can be one of the five CJAC productivity measures; measures computed from the case-level microdata, such as case duration or days to schedule hearings; or metrics of Quantum login activity, measured for court c in month $t \in$ {February, March, April, May, June}. The dummy variables $Email_c$, $Simplified\ Feedback_c$, and $Peer\ Comp.\ Feedback_c$ indicate whether the manager of court c received the email promotion, simplified feedback, or peer comparison feedback treatments, respectively. $\mathbf{X}_c$ is a vector of court characteristics, including the manager's sex, age, age squared, a dummy for high experience (12 or more years in the position), and a dummy for large courts (used as the strata variable in the treatment randomization).[13] The vectors $\mathbf{Y}_{nct_0}$ and $\mathbf{H}_{nct_0}$ contain the pre-intervention values of indicator n and the number of hearings, respectively, for October 2020 through January 2021.[14] Finally, γ_t is a month fixed effect.

The coefficients α_{n1}, α_{n2}, and α_{n3} measure the intention-to-treat (ITT) effects of the respective treatments.[15] ITT interpretation is appropriate because not all managers assigned to feedback treatments received them (requiring a Quantum login), and managers may not have read the email messages. For regressions of the official monthly indicators, Y_{nct} is standardized using control group observations over October–June, giving each indicator a standard deviation of one.

[13]The dummy for large courts equals one if the number of incoming cases in 2019 is above the median.

[14]Number of hearings in pre-intervention months is added as a control because Table 3 shows statistically significant differences between courts in the feedback and no-feedback treatments.

[15]A concise model is used due to limited sample size (49 courts), see Footnote 3 and Appendix A.1 for details. Full treatment interactions are reported in A.6.2.

4.2 Additional Specifications

To complement the duration analysis, we estimate a Cox proportional hazards model, where the hazard function for case i resolved in court c at time t is

$$\lambda(t|X_{ic}) = \lambda_0(t)\exp(\beta_{n1}Email_c + \beta_{n2}Simplified\ Feedback_c + \beta_{n3}Peer\ Comp.\ Feedback_c + \beta_{n4}\mathbf{X}_c + \beta_{n5}\mathbf{Y}_{nct_0} + \beta_{n6}\mathbf{H}_{nit_0}),$$

using the same explanatory variables as specification (1) but without month fixed effects. The Cox model accounts for right-censoring, incorporating information from unresolved cases.

Finally, we examine how treatments affect case resolution types using a multinomial logistic regression. Let k denote one of ten resolution types (see Table 2), with ongoing cases as the reference category. The probability that case i in court c resolves as k is

$$\Pr(E_{ic} = k) = \frac{e^{\boldsymbol{\delta}_k \cdot \mathbf{X}_c}}{1 + \sum_{j=1}^{K-1} e^{\boldsymbol{\delta}_j \cdot \mathbf{X}_c}}, \tag{2}$$

with

$$\boldsymbol{\delta}_k \cdot \mathbf{X}_c = \delta_0 + \delta_1 Email_c + \delta_2 Simplified\ Feedback_c + \delta_3 Peer\ Comp.\ Feedback_c + \delta_4\mathbf{X}_{ct} + \delta_5\mathbf{E}_{t_0} + \delta_6\mathbf{I}_{ct_0} + \delta_7\mathbf{H}_{ct_0} + \delta_8\gamma_t,$$

where $\mathbf{E}_{t_0}$ and $\mathbf{I}_{ct_0}$ denote the pre-intervention average share of each conclusion type and the inventory older than two years (October 2020–January 2021), with $\mathbf{E}_{t_0}$ effectively controlling for lagged dependent variables.

5 Results

5.1 Monthly Court Productivity Measures

Table 6 reports the intention-to-treat (ITT) effects of the three treatments on the five monthly CJAC productivity measures.[16] Outcomes are presented in sequence: first, indicators of immediate managerial activity—timely motion resolution, timely hearing programming, and hearing schedule adherence—and then downstream case resolution

[16] Appendix A.6.1 shows results for additional dependent variables, such as measures of Quantum login activity and auxiliary productivity indicators constructed from case-level microdata. Appendix A.6.2 reports results from parsimonious specifications that progressively add covariates. Appendix A.6.3 reports p-values adjusted for multiple hypothesis testing. For specifications that exploit the factorial design by accounting for interactions between interventions, see Appendix A.6.2.

outcomes—the case clearance rate and timely case resolutions—which are less directly under managerial control but represent broader measures of court productivity.

Table 6: Intention-to-Treat Effects of Email Promotion, Simplified and Peer Comparison Feedback on CJAC Productivity Measures

	Timely Motion Resolution (1)	Timely Hearing Programming (2)	Hearing Schedule Adherence (3)	Case Clearance† (4)	Timely Case Resolution (5)
Email promotion	0.0517	-0.0375	-0.371***	0.199	0.0751**
	(0.0333)	(0.0561)	(0.0905)	(0.144)	(0.0331)
Simplified feedback	0.101**	0.249***	-0.259**	0.361**	0.0302
	(0.0455)	(0.0680)	(0.126)	(0.157)	(0.0407)
Peer comparison	0.0976***	0.00937	-0.380***	0.128	0.00427
	(0.0379)	(0.0793)	(0.137)	(0.163)	(0.0491)
p-value: Simp. = Peer	0.915	0.001	0.365	0.161	0.490
N	245	245	245	245	245

Notes: OLS estimates. All regressions control for a dummy for large courts (the strata variable); manager characteristics (sex, age, and age squared) and a dummy for tenure $\geq$ 12 years; a dummy for whether the manager logged into Quantum in January 2021; pre-intervention values of the dependent variable and the number of hearings (Oct 2020–Jan 2021); and month fixed effects. Standard errors clustered at the court level and bootstrapped. † Similar magnitudes are estimated for case clearance if the dependent variable is constructed from case-level data.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns (1) and (2) show effects on immediate managerial activities. Both feedback treatments significantly improve timely motion resolution (about 0.1 standard deviations). For timely hearing programming, simplified feedback has a particularly strong and significant effect (0.25 standard deviations, significant at the 1% level). By contrast, email promotion has no measurable impact on these outcomes.

Column (3) shows that all three treatments reduce hearing schedule adherence, defined as the share of hearings held on their originally scheduled date. This decline does not reflect fewer hearings overall. On the contrary, Appendix Table A4 shows that the interventions increased the total number of hearings per case, with significant effects for email promotion and peer comparison feedback. This pattern suggests that treated managers scheduled more hearings (a quantity not directly observed), which in turn required more rescheduling. Consistent with this interpretation, both feedback interventions also reduce the average length of realized hearings, with the effect statistically significant at the 10% level for simplified feedback (Appendix Table A4, column (3)). Taken together, these results suggest that treated managers increased the scheduling of hearings, forcing the court to reschedule some and to shorten the duration of those that took place.

Turning to the downstream indicators, simplified feedback has a statistically significant positive impact on the case clearance rate, increasing the indicator by 0.36 standard deviations on average (column (4)). The impacts of email promotion and peer comparison feedback on case clearance are positive but not statistically signif-

icant at conventional levels. The simplified feedback result survives correction for multiple hypothesis testing: after Šidák-Holm p-value adjustment across all 15 hypotheses, it remains statistically significant at the 5% level (adjusted p = 0.045); see Appendix Table A7. Section 5.2 shows that cases also resolve faster under simplified feedback in case-level data, with the hazard rate of resolution for incoming cases rising by roughly six percent. The estimate is, however, more sensitive to the covariate set than our other results. As Appendix Table A5 shows, the simplified feedback coefficient is positive but imprecise without manager characteristics (0.237) and larger once they are included (0.373 in column 3, 0.361 in column 4); the change in significance reflects the larger point estimate rather than a smaller standard error.

For the timely case resolution rate (the share of cases with a duration shorter than two years), the feedback treatments show no statistically significant effects (column (5)). By contrast, email promotion raises timely case resolution by 0.07 standard deviations, significant at the 5% level. Given the longer horizon of this outcome and the short span of the panel, muted feedback effects are unsurprising, while email promotion appears to shift attention toward older cases. However, this effect shows a significant positive interaction with simplified feedback (Appendix Table A6), complicating interpretation of the underlying mechanism. Overall, the evidence suggests that feedback encourages managers to act on relatively newer cases, whereas email promotion prompts attention to older ones. We next turn to case-level data to examine these mechanisms more directly.

5.2 Case Duration

Table 7 reports Cox regression results for three subsamples: all cases (column (1)), only incoming cases (column (2)), and ongoing cases older than two years (column (3)). Consistent with the improved case clearance reported above, both simplified and peer comparison feedback increase the hazard rate of case resolution by about six percent in the full sample and in the incoming-case subsample. For incoming cases, the effects are statistically significant at the 5% (simplified feedback) and 10% (peer comparison feedback) levels, while the full-sample estimates are similar in magnitude but statistically insignificant. For older cases, the estimated effects are considerably larger, and email promotion has a significant impact, suggesting that the intervention encouraged managers to prioritize longstanding cases. This pattern aligns with the earlier finding that email promotion improved timely case resolution.

To further probe case-level mechanisms, columns (4) and (5) present OLS estimates from specification (1) for incoming cases, using case duration (in days) and the

number of days to schedule hearings (conditional on hearings being held) as outcomes. Because OLS does not account for right-censoring, these results exclude cases that remained open at the end of the sample and should therefore be interpreted cautiously. Nonetheless, the estimates mirror the hazard rate results: simplified and peer comparison feedback reduce average case duration by roughly 4 and 7 days, respectively, and also shorten the time required to schedule hearings, though the estimates are somewhat noisy.

Table 7: Treatment Impacts on Micro Determinants

	Hazard Rate (All cases) (1)	Hazard Rate (Incoming) (2)	Hazard Rate (Ongoing >2y) (3)	Case Duration (Days) (4)	Days to Schedule (Hearings) (5)
	(1)	(2)	(3)	(4)	(5)
Email promotion	0.035	-0.031	0.328*	-1.463	3.210
	(0.027)	(0.028)	(0.195)	(2.834)	(2.312)
Simplified feedback	0.059	0.064**	0.254	-4.254	-1.715
	(0.037)	(0.031)	(0.199)	(5.565)	(3.349)
Social comparison	0.059	0.062*	0.142	-7.296	-3.321
	(0.040)	(0.036)	(0.272)	(4.949)	(3.463)
Observations	293430	129887	8705	100111	73640

Notes: Columns (1)–(3) report Cox regression estimates, and columns (4) and (5) report OLS estimates. All regressions control for manager characteristics (sex, age, and age squared) and a dummy for tenure $\geq$ 12 years; a dummy for whether the manager logged into Quantum in January 2021; pre-intervention values of the dependent variable and the number of hearings (Oct 2020–Jan 2021); Columns (2), (4), and (5) include month fixed effects. The Cox models stratify on court size (large vs. small courts); the OLS models include it as a covariate. Standard errors are clustered at the court level. Case horizons: (a) all cases not yet closed before treatment start date plus those started during treatment months; (b) all cases started during treatment months; (c) all unclosed cases older than 2 years at treatment start date.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.3 Effects on Quality of Proceedings

We next examine whether faster case processing altered the nature of case conclusions. Using a multinomial logistic regression on 129,887 cases filed during the treatment period, with ongoing cases as the reference category (Table 8), we estimate changes in the probabilities of different case endings. For instance, simplified feedback reduces the relative log odds of involuntary dismissal by 0.413, corresponding to a 0.9% decrease in its probability, holding other covariates at their means.[17] Overall, the interventions do not meaningfully affect the most common outcomes (judgments, motions to dismiss in limine, mediations, or ongoing cases). They slightly reduce involuntary dismissals

[17]Coefficients reflect log-odds changes relative to the reference category and may not always be statistically significant, while marginal effects capture probability changes and may show significance even when individual coefficients do not.

(2% of cases) and assignments to incompetent courts (6.3% of cases), while slightly increasing settlements (11.3% of cases) and voluntary dismissals (0.9% of cases). Although the effect on ongoing cases is not statistically significant, it aligns with earlier evidence that feedback reduces case duration and increases clearance rates. Taken together, these results suggest that managers exposed to the interventions process cases more quickly without compromising—and perhaps even improving—the quality of judicial outcomes.

5.4 Discussion

In sum, the interventions influenced both the pace and focus of case management. Simplified and peer comparison feedback improved immediate managerial tasks, reduced case duration, and raised clearance rates, primarily by accelerating the resolution of newer cases. The evidence for the email promotion is more tentative, but it points to a shift in attention toward longstanding cases and modest gains in timely case resolution. [18] Importantly, these efficiency improvements were not offset by any detectable decline in the quality of judicial outcomes.

The results highlight two distinct channels through which information interventions can influence managerial performance. The first operates through increased use of the Quantum platform. As shown in Appendix A.6.1, the email promotion raised the probability that managers logged into Quantum by 0.08 standard deviations (significant at the 10% level). The treatment is also associated with a larger coefficient for the number of logins (about 0.3 standard deviations), although this effect is not statistically significant at conventional levels. By contrast, the feedback interventions did not meaningfully affect login activity. These patterns suggest that the emails modestly increased exposure to performance information by encouraging some managers to reengage with the Quantum platform.

The second mechanism operates through improved access to information once on the platform. Both feedback treatments made key indicators more salient and easier to interpret. The performance gains associated with these treatments likely stem from this transparency channel rather than from higher platform use. Taken together, the results suggest that more frequent logins, as encouraged by the email promotion, may have helped managers identify and address older pending cases, while the more transparent information provided by the feedback treatments altered how managers handled newer cases, streamlining daily operations and accelerating case resolution.

[18] The interaction specification (Table A6, column 5) reveals a positive email × simplified feedback interaction (0.179, $p < 0.10$), indicating the effect is larger when combined with the redesigned dashboard, which suggests a more complicated mechanism.

Table 8: Multinomial Logistic Regression by Type of Case Endings

Panel A. Case Endings (Part I)

	Coef.	Marg. Eff.
Ongoing case (ref. category)		
Simplified Feedback	0.000	-0.007
	(.)	(0.011)
Peer Comparison Feedback	0.000	0.001
	(.)	(0.012)
Email Promotion	0.000	-0.011
	(.)	(0.009)
Involuntary dismissal		
Simplified Feedback	-0.413**	-0.009**
	(0.183)	(0.004)
Peer Comparison Feedback	-0.392*	-0.008*
	(0.226)	(0.005)
Email Promotion	-0.225	-0.005*
	(0.153)	(0.003)
Settlement (II)		
Simplified Feedback	-0.046	-0.001
	(0.198)	(0.003)
Peer Comparison Feedback	0.117	0.002
	(0.189)	(0.003)
Email Promotion	0.076	0.000
	(0.161)	(0.002)
Conciliation		
Simplified Feedback	-0.007	-0.002
	(0.097)	(0.004)
Peer Comparison Feedback	-0.044	-0.002
	(0.129)	(0.006)
Email Promotion	-0.102	-0.008*
	(0.095)	(0.005)
Voluntary dismissal (no resubmission)		
Simplified Feedback	0.126	0.001
	(0.113)	(0.001)
Peer Comparison Feedback	0.194*	0.003**
	(0.112)	(0.002)
Email Promotion	-0.031	-0.001
	(0.085)	(0.001)

Panel B. Case Endings (Part II)

	Coef.	Marg. Eff.
Incompetent court		
Simplified Feedback	-0.116	-0.009***
	(0.082)	(0.003)
Peer Comparison Feedback	-0.213**	-0.013***
	(0.092)	(0.003)
Email Promotion	0.126	0.005
	(0.084)	(0.003)
Mediation		
Simplified Feedback	0.110	0.009
	(0.082)	(0.009)
Peer Comparison Feedback	0.017	0.002
	(0.082)	(0.009)
Email Promotion	0.126*	0.009
	(0.070)	(0.007)
Motion to dismiss in limine		
Simplified Feedback	0.068	0.006
	(0.099)	(0.010)
Peer Comparison Feedback	0.046	0.008
	(0.113)	(0.012)
Email Promotion	-0.006	-0.008
	(0.080)	(0.008)
Voluntary dismissal (resubmission allowed)		
Simplified Feedback	0.257**	0.002**
	(0.126)	(0.001)
Peer Comparison Feedback	0.121	0.001
	(0.165)	(0.001)
Email Promotion	0.070	0.000
	(0.143)	(0.001)
Judgement		
Simplified Feedback	0.056	0.008
	(0.063)	(0.007)
Peer Comparison Feedback	0.011	0.004
	(0.067)	(0.008)
Email Promotion	0.101*	0.016**
	(0.054)	(0.007)
Settlement		
Simplified Feedback	0.132	0.002
	(0.119)	(0.001)
Peer Comparison Feedback	0.172	0.003**
	(0.122)	(0.001)
Email Promotion	0.315***	0.004***
	(0.102)	(0.001)

Notes: All regressions control for a dummy for large courts (the strata variable); manager characteristics (sex, age, and age squared) and a dummy for tenure $\geq$ 12 years; a dummy for whether the manager logged into Quantum in January 2021; pre-intervention values of the dependent variable and the number of hearings (Oct 2020–Jan 2021); and month fixed effects. Standard errors clustered at the court level. The ongoing case category is the reference group for coefficient results.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The survey evidence provides additional context for interpreting the transparency channel. The substantial dispersion in managers' recalled performance — documented in Figure A1 — suggests that a meaningful share of managers were operating with considerable uncertainty about their court's actual performance. This uncertainty does not appear to reflect systematic over- or under-confidence, but rather imprecision in managers' working knowledge of their own metrics. By making key metrics accessible and digestible, the feedback dashboards may have closed the gap between available information and managers' working knowledge of their court's performance, enabling more informed operational decisions.

Between the two feedback designs, the estimated effects are generally similar, though the simplified feedback treatment produces significantly larger improvements for timely hearing programming ($p = 0.001$, Table 6) and marginally larger effects for case clearance. Other outcomes show no statistically significant differences between the two feedback designs. This pattern likely reflects the design of the peer comparison intervention. As discussed in Section 3.2, the popup window highlighted each court's single best-performing and two worst-performing indicators relative to peers. This selective highlighting appears to have channeled managerial attention toward specific metrics, generating outcome-specific effects, consistent with evidence in literature (Van der Heijden (2013); Mattsson, 2025). Nonetheless, both feedback versions produced clear improvements in core performance indicators, underscoring the value of simple, well-designed information dashboards in bureaucratic settings.

Following the intervention period, the Chilean Judiciary redesigned the Quantum platform for operational reasons. While this required building an entirely new system, the redesign incorporated many features from the simplified feedback dashboard developed for this study, though it did not include the peer comparison pop-up.

6 Conclusion

In this study, we examined the underutilization of performance data in Chilean family courts and its potential to improve court management. Partnering with the Judiciary's Department of Institutional Development, we implemented a randomized controlled trial to test low-cost informational interventions: promoting Quantum platform usage via email, simplifying the homepage to highlight key statistics, and introducing peer comparisons to benchmark performance.

Our results reveal substantial information gaps among court managers and demonstrate that targeted interventions can meaningfully improve managerial performance.

Simplified feedback increased timely motion resolution, hearing scheduling, and case clearance rates, while peer comparison feedback had more nuanced effects. Importantly, these efficiency gains were achieved without compromising the quality of judicial outcomes—if anything, cases were slightly less likely to end in involuntary dismissal or assignment to an incompetent court. Email promotion showed suggestive evidence of increasing platform engagement and shifting attention toward longstanding cases, though the effects were less precise relative to the feedback interventions.

Overall, the findings highlight that information constraints often limit performance in complex public-sector organizations. Low-cost interventions that increase data transparency and reduce information frictions can improve efficiency while maintaining quality. Future research could examine the persistence of these effects, explore the psychological and organizational mechanisms at play, and consider how emerging technologies—such as AI-driven dashboards—might further enhance public-sector performance by optimizing access to actionable information.

References

Allcott, H. and Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, 104(10):3003–3037.

Amar, M., Ariely, D., Ayal, S., Cryder, C. E., and Rick, S. I. (2011). Winning the battle but losing the war: The psychology of debt management. *Journal of Marketing Research*, 48(SPL):S38–S50.

Ashraf, N., Bandiera, O., and Jack, B. K. (2014). No margin, no mission? a field experiment on incentives for public service delivery. *Journal of Public Economics*, 120:1–17.

Banerjee, A. V., Duflo, E., and Glennerster, R. (2008). Putting a band-aid on a corpse: incentives for nurses in the indian public health care system. *Journal of the European Economic Association*, 6(2-3):487–500.

Callen, M., Gulzar, S., Hasanain, S. A., Khan, M. Y., and Rezaee, A. B. (2020). Data and policy decisions: Experimental evidence from pakistan. *NBER Working Paper*, 27678.

Celhay, P. A., Gertler, P., Olivares, M., and Undurraga, R. (2024). How managers can use purchaser performance information to improve procurement efficiency. Technical report, National Bureau of Economic Research.

Chemin, M., Chen, D. L., Di Maro, V., Kimalu, P., Mokaya, M., and Ramos-Maqueda, M. (2023). Data science for justice: Evidence from a randomized judicial reform in kenya.

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–1177.

Chilean Judiciary (2020). Design and description of positions in the country's tribunals: Family courts.

Cohen, J. (1988). Statistical power analysis for behavioral science. hillsdale, new jersey: Lawrance erlbaum associates. *Inc., Publishers*.

Cooper, R., Doyle Jr, J. J., and Hojman, A. P. (2023). Effects of enhanced legal aid in child welfare: Evidence from a randomized trial of mi abogado. Technical report, National Bureau of Economic Research.

Dhaliwal, I. and Hanna, R. (2017). The devil is in the details: The successes and limitations of bureaucratic reform in india. *Journal of Development Economics*, 124:1–21.

Dizon-Ross, R. (2019). Parents' beliefs about their children's academic ability: Implications for educational investments. *American Economic Review*, 109(8):2728–65.

Dodge, E., Neggers, Y., Pande, R., and Moore, C. (2021). Updating the state: Information acquisition costs and public benefit delivery. Technical report, Working paper, July.

Durham, C. and Becker, D. (2016). A case for court governance principles.

Engel, C. (2013). Legal experiments-mission impossible? *Available at SSRN 2276566*.

European Commission for the Efficiency of Justice (CEPEJ), C. o. E. (2021). Handbook on court dashboards. Technical Report CEPEJ(2021)8REV1, European Commission for the Efficiency of Justice (CEPEJ), Council of Europe.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129(4):1661–1710.

Greiner, D. J. and Matthews, A. (2016). Randomized control trials in the united states legal profession. *Annual Review of Law and Social Science*, 12:295–312.

Hanson, R. A. and Ostrom, B. J. (2014). Achieving better court management through better data. *J. App. Prac. & Process*, 15:19.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70.

Humphrey, S. E., Nahrgang, J. D., and Morgeson, F. P. (2007). Integrating motivational, social, and contextual work design features: a meta-analytic summary and theoretical extension of the work design literature. *Journal of applied psychology*, 92(5):1332.

Kleiman, M. (2009). Building a better understanding of trial courts. *Justice System Journal*, 30(3):321–326.

Kondylis, F. and Stein, M. (2023). The speed of justice. *Review of Economics and Statistics*, 105(3):596–613.

Mattsson, M. (2021). Information systems, service delivery, and corruption: Evidence from the bangladesh civil service. *Service Delivery, and Corruption: Evidence from the Bangladesh Civil Service (December 16, 2021)*.

Mattsson, M. (2025). Information systems, service delivery, and corruption: Evidence from the Bangladesh civil service. *American Economic Journal: Applied Economics*, 17(1):102–133.

Muralidharan, K., Niehaus, P., Sukhtankar, S., and Weaver, J. (2021). Improving last-mile service delivery using phone-based monitoring. *American Economic Journal: Applied Economics*, 13(2):52–82.

OECD (2014). Data-driven innovation for growth and well-being.

Ostrom, B. J. and Hanson, R. A. (2010). *Achieving high performance: A framework for courts*. National Center for State Courts.

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, 105:2539–69.

Rasul, I. and Rogger, D. (2018). Management of bureaucrats and public service delivery: Evidence from the nigerian civil service. *The Economic Journal*, 128(608):413–446.

Sadka, J., Seira, E., and Woodruff, C. (2024). Information and bargaining through agents: Experimental evidence from mexico's labour courts. *Review of Economic Studies*, 91(6):3677–3711.

Savoldelli, A., Codagnone, C., and Misuraca, G. (2014). Understanding the e-government paradox: Learning from literature and practice on barriers to adoption. *Government Information Quarterly*, 31:S63–S71. ICEGOV 2012 Supplement.

Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, 62(318):626–633.

Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press, New Haven, CT.

Van der Heijden, H. (2013). Evaluating dual performance measures on information dashboards: Effects of anchoring and presentation format. *Managerial Auditing Journal*, 28(7):564–579.

Whidden, C., Kayentao, K., Liu, J. X., Lee, S., Keita, Y., Diakité, D., Keita, A., Diarra, S., Edwards, J., Yembrick, A., et al. (2018). Improving community health worker performance by using a personalised feedback dashboard for supervision: a randomised controlled trial. *Journal of global health*, 8(2).

A Appendix

A.1 Pre-Analysis Plan

The pre-analysis plan can be seen below. The initial intervention was intended to encompass all three hundred and forty-six Chilean courts supervised by court managers; however, due to institutional constraints, the implementation of the intervention was confined solely to family courts.

Moreover, it is important to note that the primary outcomes listed in the pre-analysis plan explain the definition of court indicators, while in the main paper, we use the name of the indicator for simplicity. The indicators with identical naming include the *case clearance rate* and the *number of days to program (or schedule) hearings*. The average length for ending cases is addressed as *case duration* in the paper; the percentage of writing resolved within 3-5 days is labeled as *timely motion resolution*; and the percentage of cases pending for more than 1-2 years is labeled as *timely case resolution rate*.

Regarding the outcome of the average time the court needs to provide a written submission during the consultation period, we were unable to get access to this statistic, but it is captured to a certain extent by timely motion resolution. The percentage of hearings that started with a delay of 15 minutes was not applicable during the pandemic period, as many of the hearings were done virtually, and this indicator captured only in-person hearings. We were unable to secure data related to appeals.

INFORMATION PROVISION AND COURT PERFORMANCE: EXPERIMENTAL EVIDENCE FROM CHILE

LAST REGISTERED ON JUNE 17, 2021

VIEW TRIAL HISTORY

Restricted Access Edit Trial

Pre-Trial

Trial Information

GENERAL INFORMATION

Title
Information Provision and Court Performance: Experimental Evidence from Chile

RCT ID
AEARCTR-0005512

Initial registration date
December 06, 2020

Initial registration date is when the trial was registered.

It corresponds to when the registration was submitted to the Registry to be reviewed for publication.

First published
December 07, 2020, 10:54 AM EST

First published corresponds to when the trial was first made public on the Registry after being reviewed.

Last updated
June 17, 2021, 4:42 AM EDT

Last updated is the most recent time when changes to the trial's registration were published.

LOCATIONS

Country
Chile

Region
National

PRIMARY INVESTIGATOR

Name
Paloma Carrillo

Affiliation
Toulouse School of Economics

Email
paloma.carrill@gmail.com

OTHER PRIMARY INVESTIGATOR(S)

PI Name
Daniel Chen

PI Affiliation
IAST/ Toulouse School of Economics

PI Email

PI Name
Bernardo Silveira

PI Affiliation
University of California, Los Angeles

PI Email

PI Name
Manuel Ramos

PI Affiliation
World Bank

PI Email
mramosmaqueda@worldbank.org

ADDITIONAL TRIAL INFORMATION

Status
In development

Start date
2020-01-01

End date
2021-09-30

Keywords
Behavior, Governance

Additional Keywords
Judicial, tribunal, judicial system, behavioral, information provision, Chile

JEL code(s)

D04

Secondary IDs

H43, D63

Abstract

Previous studies have shown that behavioral nudges can be a cost-effective tool to influence changes in people's actions. In this study, we aim to test whether nudging court managers through informing them on how their court performs in absolute and relative terms can improve court productivity. Moreover, we test if there is any difference if the information about the court performance is given in contrast and relation to self past performance or if the information is relative to other courts' performance.

External Link(s)

REGISTRATION CITATION

Citation

Carrillo, Paloma et al. 2021. "Information Provision and Court Performance: Experimental Evidence from Chile." AEA RCT Registry. June 17. https://doi.org/10.1257/rct.5512-1.1

Sponsors & Partners

PARTNER

Name

Department of Institutional Development (DDI) at the Administrative Corporation of the Chilean Judiciary (CAPJ)

Type

government

URL

Experimental Details

INTERVENTIONS

Intervention(s)

In this study, we aim to test whether nudging court managers through informing them on how their court performs in absolute and relative terms can improve court productivity.

Intervention (Hidden)

The Department of Institutional Development (DDI) of the Chilean judicial system developed an electronic platform called Quantum in 2018. Quantum displays comprehensive information on court indicators, such as the number of cases filed, the case clearance rate, the average duration of cases finished in a month, and the percentage of realized hearings. It also allows users to compare their courts to other courts in the

same jurisdiction. Quantum was launched in 2018, but take-up has been limited: 20% of court managers have never logged in, and overall there is an average of 20 logins per court manager in 1 year and 2 months. The platform is technologically well developed and rich in information, yet it is unclear whether it has any impact on the management of the court. Our project consists of evaluating an intervention through a randomized controlled trial (RCT) with two main branches. First, we will (randomly) promote the Quantum platform in multiple ways, such as sending court managers a survey that implicitly markets the platform, making phone calls, and sending them emails. Second, we will (also randomly) provide court managers a new dashboard that summarizes the main statistics displayed in Quantum and compares them to themselves to a reference group of courts. There would be a total of six distinct intervention or treatments:

Treatment 0: Control. No change to their Quantum dashboards nor provided with any Quantum promotion.

Treatment 1: Quantum Promotion
The tribunals randomized into Treatment 1 will have their court staff receive both emails with a Quantum link to increase accessibility and salience of Quantum, and a small baseline and post-intervention survey that includes Quantum promotion. The baseline survey given to the court managers at the beginning of the RCT will ask them about their beliefs about some productivity metrics, how much these metrics affect their decision at work, inform them that these metrics can be seen in Quantum, and describe the effect of Quantum usage on people's productivity through the results of an event study using historical data. A sample survey is provided in the appendix.

Treatment 2: No Quantum Promotion + New Dashboard
The tribunals randomized into treatment 2 will not receive any promotion but will have their home page in Quantum, what we call the dashboard, present various statistics at the tribunal level.

Treatment 3: Quantum Promotion + New Dashboard
The tribunals randomized into treatment 3 will receive the same promotion as that in treatment 1 and the new dashboard as in Treatment 2.

Treatment 4: No Quantum Promotion + New Dashboard + Comparative to others
The tribunals in treatment 4 will receive the new dashboard plus another tab or pop-up window that focuses on the tribunal's best performing and worst-performing dimension from the previous month in comparison to the performance of peer tribunals (same competence) in the same month. This comparison leans into social comparison motivation.

Treatment 5: Quantum Promotion + New Dashboard + Comparative to others
The tribunals in treatment 4 will receive the new dashboard plus another tab or pop-up window that focuses on the tribunal's best performing and worst-performing dimension from the previous month in comparison to the performance of peer tribunals (same competence) in the same month. This comparison leans into social comparison motivation.

The court managers' job satisfaction level will be measured with pre and post-surveys to court managers that measure perceptions of their tribunals and their satisfaction with their positions. By informing the court managers about their court's standing in the new dashboard and comparative to others, this could change how empowered or satisfied a court manager is with his or her position and power role.
Overall, the objectives of the survey for court staff are threefold. First, it will measure their knowledge of Quantum statistics. How close or far is their perception of their court's performance from the truth. We can later use this information (prior beliefs) to understand if greater access to Quantum updates the beliefs closer to the truth when we measure their posteriors (survey at the end of intervention). Notice that the measurement of beliefs and opinions is something unique to the survey that the rest of the interventions cannot. Second, the survey will allow us to understand if the court staff find the statistics important and in what order of importance. This is useful for Quantum to know which variables are important for users and make them more salient in the dashboard or in the rest of the Quantum pages. That is, the results from the survey can help tailor the intervention to make it more effective. Third, the survey will promote Quantum as a source of accurate and useful information through the event study results. This may help influence those that are skeptical of Quantum to give it a chance.

Intervention Start Date

2020-12-14

Intervention End Date

2021-09-30

PRIMARY OUTCOMES

Primary Outcomes (end points)

Case clearance rate, average length for filing cases (days), average length for ending cases, the average time the court needs to provide a written submission during the consultation period, percentage of writing resolved with 3-5 days, average number of days to program a hearing, percentage of hearing that started with a delay of 15 minutes, percentage of cases pending for more than 1-2 years, appeal rate, and number of cases appealed.

On the promotion intervention, the main primary outcome is number of logins per court manager to the Quantum platform

Primary Outcomes (explanation)

SECONDARY OUTCOMES

Secondary Outcomes (end points)

Court managers job satisfaction

Secondary Outcomes (explanation)

The court managers' job satisfaction level will be taken from the pre and post-surveys to court managers that measure perceptions of their tribunals and their satisfaction with their positions. By informing the court managers about their court's standing it could change how empowered or satisfied a court manager is with his or her position and power role.

EXPERIMENTAL DESIGN

Experimental Design

The program will have six distinct treatments. The treatments will combine promoting the usage of an electronic platform that contains information on their court performance and providing distinct homepages in this platform that will summarize the courts performance stressing the weaknesses and strengths of the court in comparison to a reference group.

Experimental Design Details

First, we will (randomly) promote the Quantum platform in multiple ways, such as sending court managers a survey that implicitly markets the platform, making phone calls, and sending them emails. Second, we will (also randomly) provide court managers a new dashboard that summarizes the main statistics displayed in Quantum and compares them to themselves in the past or to a reference group of courts. There would be a total of six treatments: (0) no quantum promotion no new dashboard (control) (1) quantum promotion, (2) no quantum promotion and new Quantum dashboard, (3) quantum promotion and a new dashboard, (4) no quantum promotion, new dashboard, and comparative that emphasizes the strongest and weakest indicators for that month in comparison to a similar group of courts in that same month, and (5) quantum promotion, new dashboard, and comparative that emphasizes the strongest and weakest indicators for that month in comparison to a similar group of courts in that same month.

Given that the information in the Quantum platform is updated daily and our dashboards are updated with monthly data, we will have multiple pre-treatment observation and many post-treatment observations.

Randomization Method

Randomization done in office by a computer.

Randomization Unit

The unit of randomization is the court. The randomization was stratified by size (small and big) and court type (7 distinct ones).

Was the treatment clustered?

No

EXPERIMENT CHARACTERISTICS

Sample size: planned number of clusters

346 courts

Sample size: planned number of observations

346 courts

Sample size (or number of clusters) by treatment arms

57 courts stay as control, 57 courts receive T1, 58 courts receive T2, 58 courts receive T3, 58 courts receive T4, and 58 courts receive T5.

Minimum detectable effect size for main outcomes (accounting for sample design and clustering)

Supporting Documents and Materials

IRB

INSTITUTIONAL REVIEW BOARDS (IRBS)

IRB Name

Toulouse School of Economics

IRB Approval Date

2020-01-18

IRB Approval Number

N/A

Analysis Plan

Post-Trial

Post Trial Information

Data Publication

Reports, Papers & Other Materials

© Copyright 2012-2023, MIT.

A.2 Survey to Court Managers

We administered a baseline survey to all 346 court managers to capture their views on their professional roles and on the Quantum platform. The Department of Institutional Development of the Chilean Judiciary distributed the survey electronically at the end of January 2020, and responses were collected through March 2020. The survey comprised three modules: (i) the nature of managerial responsibilities, the main obstacles to effective court management, and the role of court performance indicators; (ii) perceptions of actual court performance, focusing on four indicators (incoming cases, resolved cases, pending cases, and the hearing schedule adherence); and (iii) knowledge and use of the Quantum platform. The third module was administered only to a randomly selected subset of respondents.

A total of 132 responses were received, of which 128 contained substantive answers. We excluded responses based on three criteria: (a) identical numerical values reported for the first three performance indicators; (b) implausible disparities among indicators—such as one value being less than 1% of either of the others; or (c) a hearing schedule adherence rate of zero. In addition, five administrators submitted duplicate responses, of which only one was retained. After applying these exclusions, the final analytic sample comprised 116 valid responses for the first two modules and 97 for the module on Quantum. The valid responses span all types of courts and all seventeen judicial jurisdictions in the Chilean court system.

The first module asked court managers about their perceptions of their work and the obstacles they face. Responses indicate that managers view their role as essential to court functioning but also report significant challenges (Table A1, Panel A). The most frequently cited obstacle (26%) was tension between administrators and judges, particularly regarding the division of responsibilities. In open-ended responses, managers noted that judges often intervene in administrative matters and resist being directed, resulting in frequently rescheduled hearings. Other common obstacles included lack of trained staff, absenteeism, and resistance to change (14%); problems with scheduling and the high rate of rescheduled hearings (7%); and limited technical resources and support (5%).

Court managers also expressed strong support for the use of performance indicators (Table A1, Panel B). Specifically, 97% agreed that indicators are important for their job; 70% agreed that tracking indicators is one of their most important tasks; and 93% agreed that decisions should be based on indicators. Managers reported using indicators mainly to redistribute workloads and to improve hearing schedules (Table A2).

Nevertheless, the survey suggests that many court managers are not well-informed about their court's actual performance indicators. In the second module, respondents were asked to recall the case clearance rate from two months prior. Approximately 56% of managers were off by more than 25% from the true value (based on 116 courts—see Figure A1 for the full distribution of recall errors).[19] Among family courts, the share of large errors dropped to 31% (based on 26 courts).

The third module, which focused on Quantum, indicates that the platform is underutilized. About 10% of managers had not heard of Quantum, and 17% reported not using it. Older administrators were less likely to use the platform. Among those who did use Quantum, 69% reported not relying on it to guide their decisions. The most common suggestion for improvement was to expand the range of available statistics and information.

Table A1: Managerial Responsibilities and Obstacles to Court Management

Statement	Mean	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
Panel A. Court managers' work						
Managers' work relevant for court	4.69	0	0	0	36	80
Managers' actions can improve performance	4.40	0	1	5	57	53
Satisfied with own work's impact	4.38	0	1	3	63	49
Existence of obstacles to improve performance	3.78	3	15	19	47	32
Panel B. Performance indicators						
Importance of knowledge of indicators	4.67	0	1	2	31	82
Decisions should be taken based on indicators	4.44	0	0	8	49	59
Indicators should influence decisions	4.18	0	1	13	66	36
Importance of tracking indicators	3.91	1	5	28	52	30

Responses from the court managers baseline survey, administered from January to March 2020. Statistics based on 116 valid responses. Values in columns 3–7 report the number of responses.

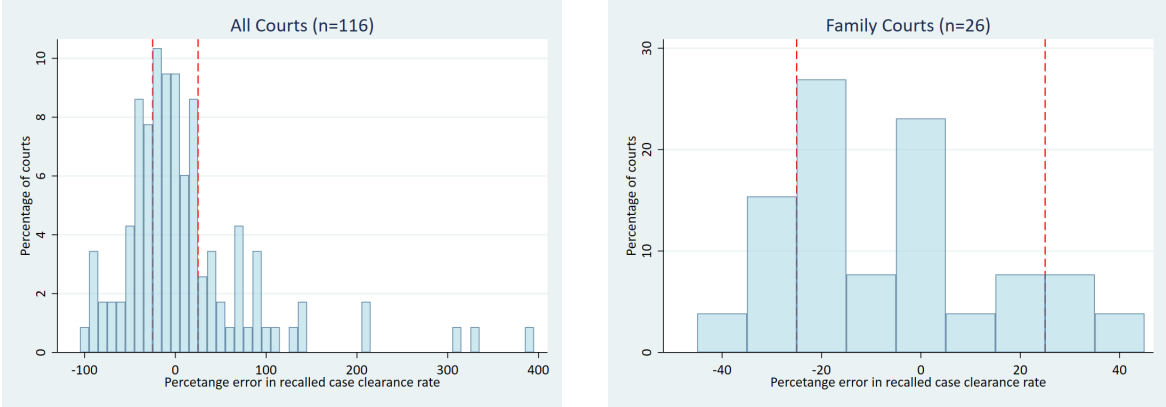
Table A2: Actions Reported by Court Managers to Improve Court Performance

Number of incoming cases	**Number of pending cases**	**Number of resolved cases**	**Hearing schedule adherence rate**
Courtroom distribution / agenda (24%)	Review causes of cases on hold (16%)	Avoid cancelling / rescheduling hearings and monitor agenda (21%)	Effectiveness in scheduling hearings and preventing cancellation (43%)
Balance of workload / redistribution of tasks (15%)	Ensure hearings occur (13%)	Strengthen reviewing of cases (14%)	Improve usage of courtrooms (13%)
Increase number of judges, employees, and/or hearings (9%)	Increase number of hearings / schedule simultaneous hearings / improve scheduling (10%) Balance workload (8%)	Improve effectiveness of courtrooms (7%)	Distribute hearings by topic (8%)

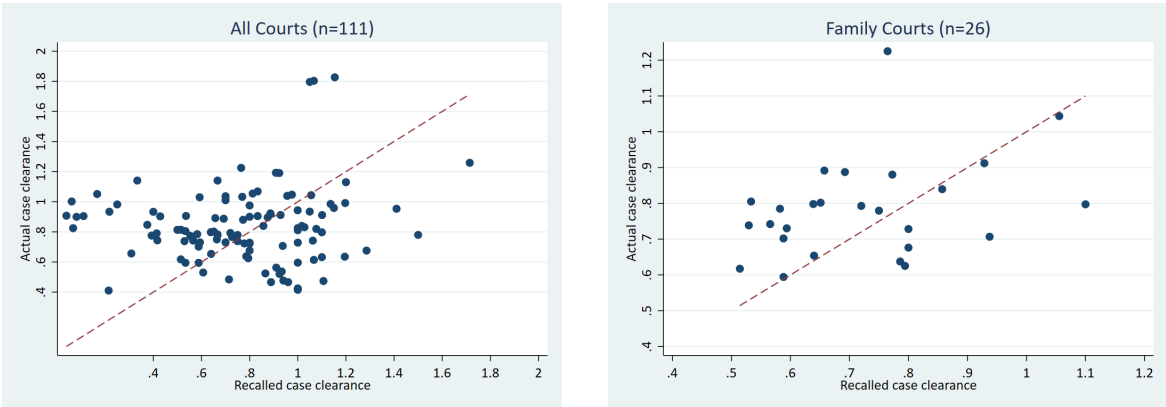
Percentages based on text analysis of baseline survey responses from court managers. Values reflect the share of respondents mentioning each action.

[19] Respondents and non-respondents were similar in age and gender composition. However, respondents had higher clearance rates, suggesting that the knowledge gap we document is likely a lower bound.

Figure A1: Court managers' knowledge of case clearance rates



(a) Distribution of percentage errors in recalled case clearance rates



(b) Actual versus recalled clearance rates

Note: Panel (a) shows vertical dashed lines at ±25% error thresholds. Panel (b) restricts to observations with percentage errors within ±200% (n=111 for all courts, n=26 for family courts) for readability; the dashed line is the 45-degree line. 56% of all court managers and 31% of family court managers had errors exceeding ±25%.

A.3 Email Versions

A.3.1 Email on January 26

Dear Court Managers,

Hoping that you are well, we remind you that for a couple of years, the institution has advanced in the construction of the Quantum management tool, which consolidates different jurisdictional indicators by court, in order to provide information for decision-making. This tool has been valued by many courts, since it allows simplifying the consolidation of statistical data to support jurisdictional work.

Did you know ...?

Courts that use Quantum the most generally have a higher term rate. That is, for each case admitted, they tend to solve a greater number of cases in proportion.

Courts that use Quantum the most generally have a smaller inventory of old cases.

For those with feedback treatment
"During the last months, we have worked with a team of court managers on the construction of a new dashboard, which displays key indicators necessary to support the management of family courts. In this sense, we invite you to log into Quantum and explore this new functionality. This dashboard also incorporates a window that compares the performance of the courts with a court of similar characteristics to yours."

You can access this platform at https://quantum.pjud (through VPN). To access the new dashboard, they must click on the name of their court within their jurisdiction. In case of any observation, requirement and even if you do not have your login credentials, please write to the email quantum@pjud.cl.

A.3.2 Email Promotion in February, March, and May

Dear Court [Name],

Please join us in congratulating the family court managers who used the Quantum platform the most during [previous month] 2021:

Name of manager 1 - Name of court 1

Name of manager 2 - Name of court 2

Name of manager 3 - Name of court 3

For those with at least one login on the previous month:
"We trust that all of you will continue to use Quantum to track your indicators, evaluate your court's performance, and compare it to other courts."

For those with no logins on the previous month:
"We noticed that you did not log into Quantum during [previous month] 2021. You are missing the opportunity to track your metrics, evaluate your court's performance, and compare it to other courts!"

Did you know that...

Most Quantum users find Quantum easy to use, displays information clearly, contains reliable indicators, and information that is useful and relevant to their work. Family courts that use Quantum multiple times a month have higher rates of termination of cases pending for more than two years.

You can access this platform at https://quantum.pjud (through VPN). To access the new dashboard, they must click on the name of their court within their jurisdiction. In case of any observation, requirement and even if you do not have your login credentials, please write to the email quantum@pjud.cl.

A.4 Minimum Detectable Effect

Table A3: Minimum Detectable Effect in Standard Deviations

	Email Promotion vs. No Email	Feedback vs No Feedback
Timely Motion Resolution	0.52	0.55
Timely Hearing Programming	0.48	0.54
Hearing schedule adherence	0.42	0.54
Case Clearance	0.43	0.46
Timely Case Resolution	0.52	0.57

These are the minimum detectable effect in standard deviations computed using the ANCOVA methodology and setting power at 80%, alpha 0.05, 25 observations for control and treatment for the email promotion vs. no email treatment, 16 observations in the control and 33 in the treatment for the feedback treatment, four monthly pre-treatment observations, and five monthly post-treatment observations.

For the five main indicators of interest, the minimum detectable effect, considering power at 80%, alpha at 0.05, four monthly pre-treatment observations, and five monthly post-treatment observations for the 49 tribunals, is between .42 and 0.57 standard deviations. According to Cohen (1988), an effect of 0.2 standard deviation is "small", 0.5 is "medium" and 0.8 is "large," even with our small sample of tribunals, we have enough power to detect "medium" size effects.

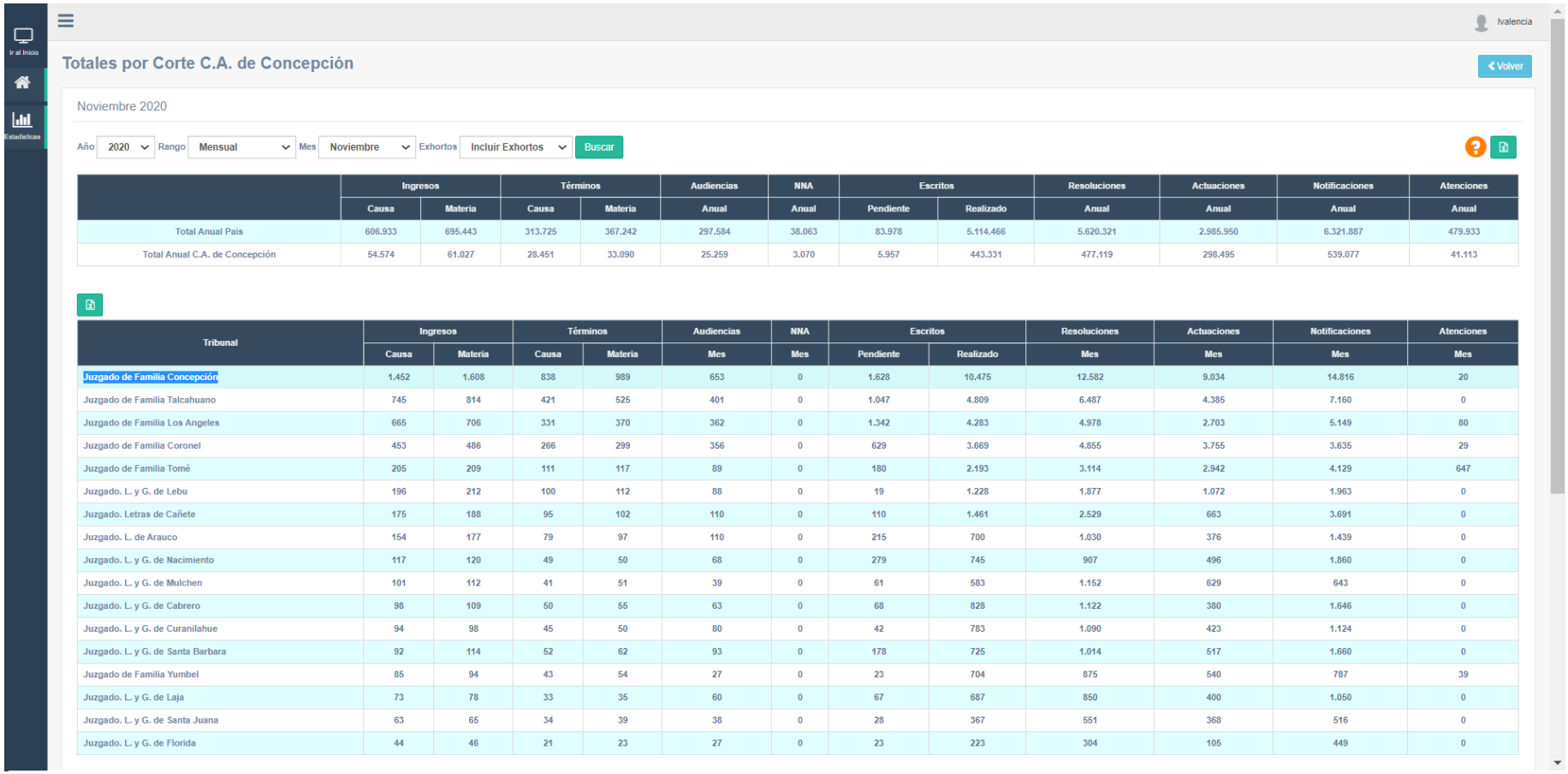
A.5 Quantum Platform Main Page Examples: Control and Treatment

Figure A2, A3, and A4 show examples of the Quantum platform for the control, dashboard, and peer comparison treatments.

A.6 Additional Empirical Results and Robustness Analyses

This Appendix reports complementary regression results that support the findings discussed in the main text.

Figure A2: Old Quantum main homepage (Control and Email-Only Treatment)



Ir al Inicio

Estadísticas

lvalencia

Totales por Corte C.A. de Concepción

Volver

Noviembre 2020

Año 2020 Rango Mensual Mes Noviembre Exhortos Incluir Exhortos Buscar

	Ingresos		Términos		Audiencias	NNA	Escritos		Resoluciones	Actuaciones	Notificaciones	Atenciones
	Causa	Materia	Causa	Materia	Anual	Anual	Pendiente	Realizado	Anual	Anual	Anual	Anual
Total Anual Pais	606.933	695.443	313.725	367.242	297.584	38.063	83.978	5.114.466	5.620.321	2.985.950	6.321.887	479.933
Total Anual C.A. de Concepción	54.574	61.027	28.451	33.090	25.259	3.070	5.957	443.331	477.119	298.495	539.077	41.113

Tribunal	Ingresos		Términos		Audiencias	NNA	Escritos		Resoluciones	Actuaciones	Notificaciones	Atenciones
	Causa	Materia	Causa	Materia	Mes	Mes	Pendiente	Realizado	Mes	Mes	Mes	Mes
Juzgado de Familia Concepción	1.452	1.608	838	989	653	0	1.628	10.475	12.582	9.034	14.816	20
Juzgado de Familia Talcahuano	745	814	421	525	401	0	1.047	4.809	6.487	4.385	7.160	0
Juzgado de Familia Los Angeles	665	706	331	370	362	0	1.342	4.283	4.978	2.703	5.149	80
Juzgado de Familia Coronel	453	486	266	299	356	0	629	3.669	4.855	3.755	3.635	29
Juzgado de Familia Tomé	205	209	111	117	89	0	180	2.193	3.114	2.942	4.129	647
Juzgado. L. y G. de Lebu	196	212	100	112	88	0	19	1.228	1.877	1.072	1.963	0
Juzgado. Letras de Cañete	175	188	95	102	110	0	110	1.461	2.529	663	3.691	0
Juzgado. L. de Arauco	154	177	79	97	110	0	215	700	1.030	376	1.439	0
Juzgado. L. y G. de Nacimiento	117	120	49	50	68	0	279	745	907	496	1.860	0
Juzgado. L. y G. de Mulchen	101	112	41	51	39	0	61	583	1.152	629	643	0
Juzgado. L. y G. de Cabrero	98	109	50	55	63	0	68	828	1.122	380	1.646	0
Juzgado. L. y G. de Curanilahue	94	98	45	50	80	0	42	783	1.090	423	1.124	0
Juzgado. L. y G. de Santa Barbara	92	114	52	62	93	0	178	725	1.014	517	1.660	0
Juzgado de Familia Yumbel	85	94	43	54	27	0	23	704	875	540	787	39
Juzgado. L. y G. de Laja	73	78	33	35	60	0	67	687	850	400	1.050	0
Juzgado. L. y G. de Santa Juana	63	65	34	39	38	0	28	367	551	368	516	0
Juzgado. L. y G. de Florida	44	46	21	23	27	0	23	223	304	105	449	0

Figure A3: New simplified homepage

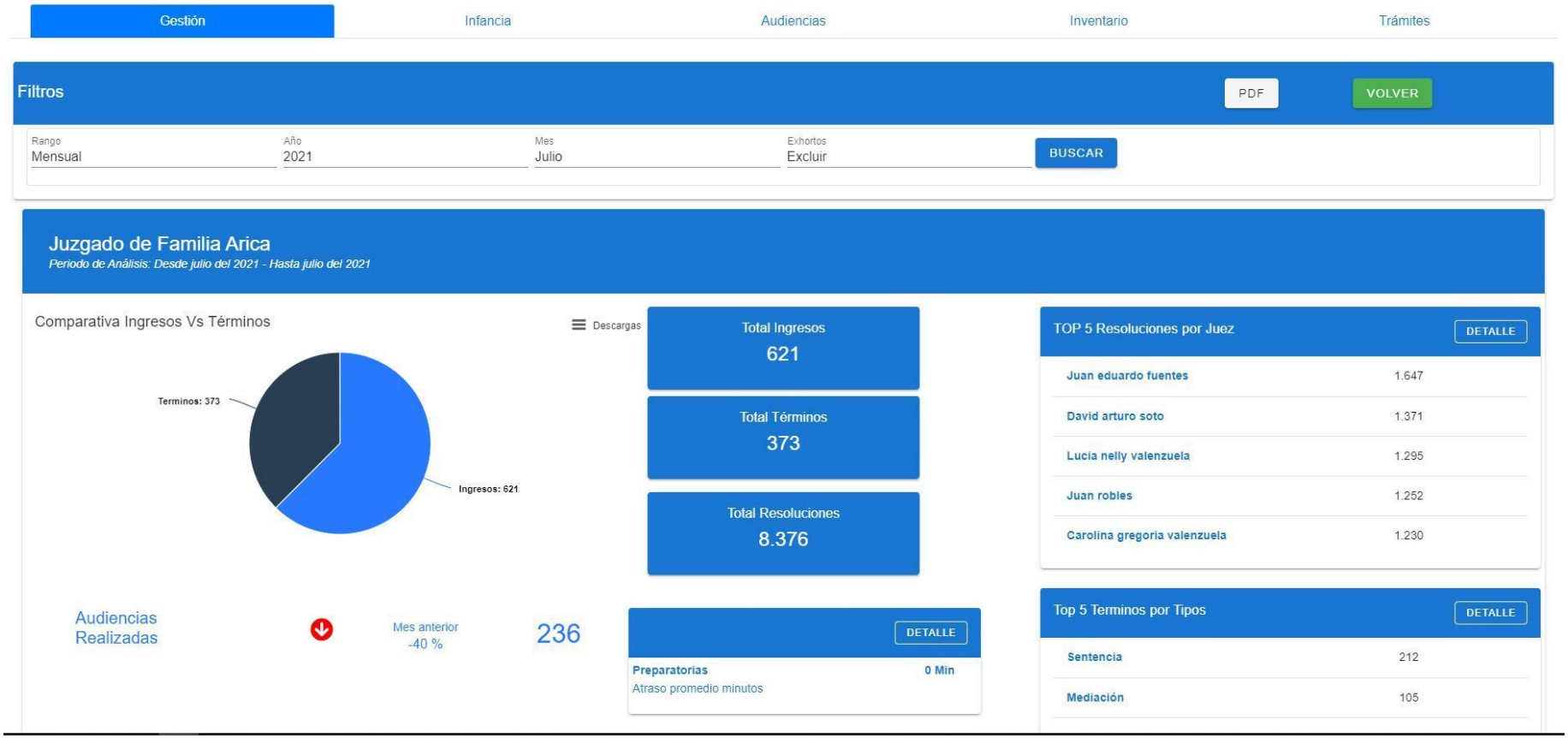
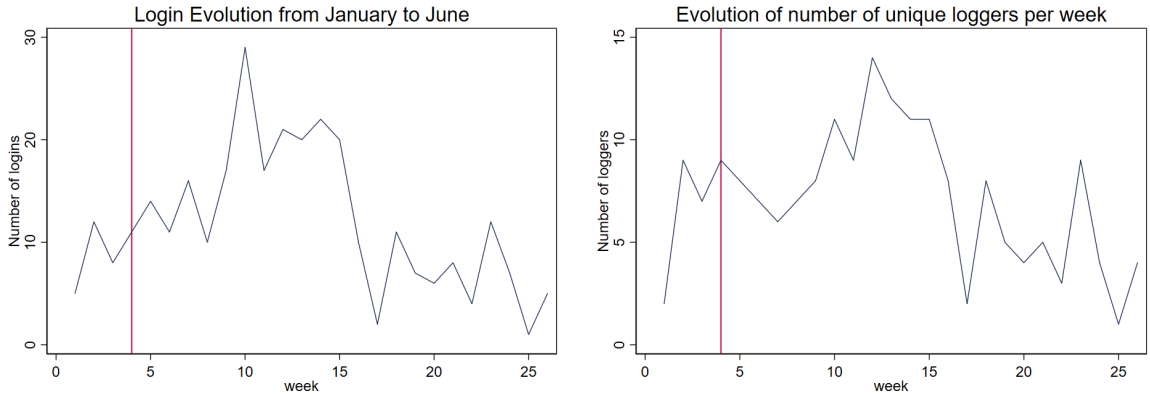


Figure A4: Pop-up window with peer comparison

A.6.1 Quantum Login and Additional Court Performance Measures

We begin by examining how the interventions affected managers' login activity on Quantum. Figure A5 plots the weekly evolution of (A) the number of logins and (B) the number of distinct managers logging in across the 49 participating courts, starting on January 1. The red vertical line marks the start of the treatment period. Although the series are noisy and the pre-treatment window is relatively short, both panels show a general upward trend in login activity, with notable peaks around weeks 10–12.

Figure A5: Evolution of Quantum Logins and Loggers



A) Number of logins B) Unique loggers

The red vertical line at week 4 marks the beginning of the intervention period.

Table A4 presents OLS estimates from specification (1), using as dependent variables a dummy for whether the manager logged in at least once (column 1) and the total number of logins (column 2). The email promotion modestly increased the likelihood that managers logged in (by about 0.08 standard deviations), consistent with the patterns in Figure A5, though its effect on the total number of logins is imprecisely estimated. By contrast, neither of the feedback treatments had a measurable effect on login behavior.

These findings suggest that the productivity gains associated with the feedback treatments, as documented in Section 5, are unlikely to be driven by increased Quantum usage. Instead, they point to changes in managerial behavior conditional on logging in, supporting the interpretation that performance improvements stemmed from the simplified information provided in the feedback dashboards. Meanwhile, the suggestive evidence that managers exposed to the email promotion logged in somewhat more frequently may indicate that renewed contact with the original platform encouraged them to focus on longstanding cases. Given the limited precision of these estimates, we interpret this evidence cautiously.

Table A4: Intention-to-Treat Effects of Email Promotion, Simplified, and Peer Comparison Feedback on Additional Outcomes

	Logger	# logins	Hearing length (minutes)	# hearings/ per case
	(1)	(2)	(3)	(4)
Email promotion	0.0831*	0.329	0.20	0.08*
	(0.0476)	(0.218)	(0.57)	(0.05)
Simplified feedback	0.0713	-0.120	-1.13*	0.06
	(0.0580)	(0.354)	(0.67)	(0.05)
Peer comparison	-0.0650	-0.0309	-1.18	0.19***
	(0.0678)	(0.357)	(0.74)	(0.05)
Control mean	0.38	1.15	25.85	2.34
N	245	245	73,640	71,624

The table reports OLS estimates. *Logger* is a dummy indicating whether the manager logged into Quantum at least once; *Logins* is the total number of logins. "Hearing duration" and "Hearings per case" are constructed from case-level microdata from *Judicial Power in Numbers*. All regressions control for a dummy for large courts (the stratification variable), manager characteristics (sex, age, and age squared), a dummy for tenure $\geq$ 12 years), pre-treatment login dummy, pre-intervention values of the dependent variable and of the number of hearings (October 2020–January 2021), and month fixed effects. Standard errors are clustered at the court level and bootstrapped.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We next extend the analysis to additional measures of court performance. Columns (3) and (4) of Table A4 report estimates for the average duration of hearings and the number of hearings per case, using case-level data. The simplified feedback reduced average hearing length by about 1.1 minutes (significant at the 10% level), with a similar but imprecise effect for the peer comparison feedback. All three interventions slightly increased the number of hearings per case, with significant effects for the email promotion (10%) and peer comparison feedback (1%).

A.6.2 Alternative Specifications

This section examines the robustness of the main results reported in Table 6 of Section 5 to alternative sets of control variables. Specifically, we re-estimate specification (1) while sequentially adding controls for pre-treatment outcomes, manager characteristics, and lagged measures of hearings.

Table A5 summarizes these robustness checks for the main performance indicators: case clearance, timely motion resolution, hearing adherence, timely case resolution, and timely hearing programming. Each panel reports the estimated coefficients for the email promotion, simplified feedback, and peer comparison feedback treatments under progressively richer specifications.

Across all outcomes, the estimated effects are generally stable with respect to the inclusion of additional controls. For timely motion resolution, the coefficients on the feedback treatments remain positive and often statistically significant, whereas the estimated effect of the email promotion is smaller and less precisely estimated. For

case clearance, the simplified feedback coefficient becomes statistically significant once manager characteristics are added (columns 3–4), compared to the more parsimonious specifications (columns 1–2). Although these characteristics are balanced across treatment arms (Table 3), they explain variation in case clearance unrelated to treatment, which in a sample of 49 courts can shift the point estimate appreciably. The change in statistical significance here reflects the larger coefficient rather than a smaller standard error. The estimate is stable across the two most saturated specifications (0.373 in column 3 versus 0.361 in column 4, which additionally includes lagged hearings), suggesting the result is not sensitive to this further control. The finding also survives Šidák-Holm correction across all 15 outcomes tested (Table A7). For hearing adherence, all three interventions have coefficients that are consistently negative and significant at conventional levels.

Turning to timely case resolution, the results for email promotion display suggestive evidence of an improvement, particularly after controlling for pre-treatment performance. Finally, for timely hearing programming, both feedback treatments show strong, positive effects in most specifications, while the email promotion has a weaker and more variable association.

Overall, the results in Table A5 confirm that the main findings are not overly sensitive to the inclusion of alternative control sets. The performance improvements observed under the feedback interventions are consistent and robust across specifications, supporting the interpretation that the simplified and peer comparison dashboards enhanced managerial efficiency through better access to performance information.

We also examined whether combining treatments produced synergistic effects by including interaction terms between email promotion and each feedback treatment Table A6. Given our limited sample size (49 courts), we have modest power to detect interactions. Nevertheless, the main effects of simplified feedback persist when accounting for interactions with email, underscoring the robustness of our core findings. For timely case resolution (column 5), the positive interaction between email and simplified feedback (0.179, $p < 0.10$)—despite insignificant isolated effects—aligns with our main finding that simplified feedback is the most impactful treatment, here working synergistically with email to address old case backlogs.

A.6.3 Multiple Hypothesis Testing

Given our hypothesis testing of multiple outcomes, we adjust to control Type I error inflation, which reduces the power to detect significant differences. Table A7 reports

Table A5: Intention-to-Treat Effects of Email, Simplified, and Peer Comparison Feedback

	Timely Motion Resolution				**Timely Hearing Programming**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Email	0.0086	0.0515*	0.0708**	0.0517	0.0617	-0.0976*	-0.0316	-0.0375
	(0.0311)	(0.0312)	(0.0311)	(0.0333)	(0.0515)	(0.0555)	(0.0497)	(0.0561)
Simplified	0.347***	0.0844**	0.155***	0.101**	0.410***	0.161**	0.236***	0.249***
	(0.0376)	(0.0397)	(0.0455)	(0.0455)	(0.0642)	(0.0708)	(0.0762)	(0.0680)
Peer Comp.	0.430***	0.133***	0.150***	0.0976***	0.361***	0.0139	0.0118	0.00937
	(0.0370)	(0.0367)	(0.0394)	(0.0379)	(0.0547)	(0.0789)	(0.0750)	(0.0793)
p-value: Simp. = Peer	0.013	0.097	0.887	0.915	0.492	0.020	0.002	0.001
	Hearing Adherence							
Email	-0.338***	-0.324***	-0.371***	-0.371***				
	(0.0858)	(0.0806)	(0.0905)	(0.0905)				
Simplified	-0.243**	-0.237**	-0.259**	-0.259**				
	(0.112)	(0.115)	(0.126)	(0.126)				
Peer Comp.	-0.262**	-0.321**	-0.380***	-0.380***				
	(0.133)	(0.136)	(0.137)	(0.137)				
p-value: Simp. = Peer	0.884	0.504	0.365	0.365				
	Case Clearance				**Timely Case Resolution**			
Email	-0.162	-0.0791	0.143	0.199	-0.524***	0.0503*	0.0611*	0.0751**
	(0.120)	(0.127)	(0.127)	(0.144)	(0.0394)	(0.0292)	(0.0320)	(0.0331)
Simplified	0.223	0.237	0.373**	0.361**	-0.0563	0.0296	0.0423	0.0302
	(0.151)	(0.163)	(0.179)	(0.157)	(0.0486)	(0.0372)	(0.0415)	(0.0407)
Peer Comp.	0.0895	0.0762	0.0571	0.128	-0.0778*	-0.00791	-0.0180	0.00427
	(0.140)	(0.163)	(0.175)	(0.163)	(0.0449)	(0.0437)	(0.0440)	(0.0491)
p-value: Simp. = Peer	0.394	0.301	0.051	0.161	0.670	0.285	0.138	0.490
Controls								
Strata and month FE	✓	✓	✓	✓	✓	✓	✓	✓
Lagged indicator		✓	✓	✓		✓	✓	✓
Manager characteristics			✓	✓			✓	✓
Lagged hearings				✓				✓
N	245	245	245	245	245	245	245	245

Regression in the last column for each indicator includes as covariates the strata variable, manager characteristics (sex, age, age squared), a dummy for tenure $\geq$ 12 years, four pre-treatment lagged values of the dependent variable, pre-treatment number of hearings, pre-treatment login dummy, and month fixed effects. Standard errors are clustered at the court level and bootstrapped.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: The Intention-to-Treat Effects with Treatment Interactions

	Timely Motion Resolution	**Timely Hearing Programming**	**Hearing Schedule Adherence**	**Case Clearance**	**Timely Case Resolution**
	(1)	(2)	(3)	(4)	(5)
Email	0.00751	0.372***	−0.807***	0.270	0.0487
	(0.0728)	(0.0786)	(0.129)	(0.253)	(0.0710)
Simplified	0.214***	0.578***	−0.375**	0.459**	−0.0683
	(0.0600)	(0.0791)	(0.186)	(0.220)	(0.0448)
Peer Comp.	−0.00960	0.277***	−0.874***	0.143	0.0484
	(0.0573)	(0.0966)	(0.176)	(0.219)	(0.0518)
Simplified × Email	−0.156*	−0.663***	0.231	−0.185	0.179*
	(0.0889)	(0.121)	(0.227)	(0.319)	(0.0945)
Peer Comp. × Email	0.227**	−0.528***	0.970***	−0.0370	−0.0859
	(0.0962)	(0.126)	(0.216)	(0.329)	(0.0867)
N	245	245	245	245	245

All regressions control for a dummy for large courts (the strata variable); manager characteristics (sex, age, and age squared) and a dummy for tenure $\geq$ 12 years; a dummy for whether the manager logged into Quantum in January 2021; pre-intervention values of the dependent variable and the number of hearings (Oct 2020–Jan 2021); and month fixed effects. Standard errors are clustered at the court level and bootstrapped.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the standard p-value adjustment method of Šidák–Holm (Šidák, 1967; Holm, 1979).[20] After the Šidák–Holm correction, the impact of the simplified feedback treatment on case clearance remains statistically significant at the 5 percent level, while the other significant coefficients in Table 6 do not survive. This is because the case clearance effect has a much lower unadjusted p-value (0.003) than the others (0.06–0.27). Under the stepdown procedure, the smallest p-value faces the full fifteen-hypothesis penalty, $1 - (1 - p)^{15}$, which maps 0.003 to 0.045; the remaining p-values face progressively smaller penalties but start too far from conventional thresholds to survive them. A Bonferroni correction yields a similar result for case clearance (adjusted $p = 0.046$), as expected at these magnitudes.

More generally, the correction should be read against the study's power. As reported in Table A.4, with 49 courts the minimum detectable effect for these indicators is 0.42–0.57 standard deviations, so a fifteen-hypothesis correction is demanding for a sample of this size. We therefore treat the other significant coefficients in Table 6 as suggestive rather than conclusive, and single out case clearance as the finding that best withstands correction for multiple testing.

[20] We compute the adjustments using the Stata command `wyoung`. We report its Šidák–Holm p-values, based on the unadjusted p-values computed analytically by `wyoung`.

Table A7: Multiple Hypothesis Testing ITT

	Timely Motion Resolution (1)	Timely Hearing Programming (2)	Hearing Schedule Adherence (3)	Case Clearance (4)	Timely Case Resolution (5)
Email promotion	0.0517	−0.0375	−0.3705	0.1987	0.0751
p-value Šidák-Holm	[0.901]	[0.991]	[0.570]	[0.345]	[0.877]
Simplified feedback	0.1010	0.2494	−0.2588	0.3614**	0.0302
p-value Šidák-Holm	[0.861]	[0.861]	[0.890]	[0.045]	[0.991]
Peer comparison feedback	0.0976	0.00937	−0.3797	0.1284	0.00427
p-value Šidák-Holm	[0.652]	[0.997]	[0.884]	[0.901]	[0.997]

Point estimates reproduce those in Table 6. All regressions control for a dummy for large courts (the strata variable); manager characteristics (sex, age, and age squared) and a dummy for tenure ≥12 years; a dummy for whether the manager logged into Quantum in January 2021; pre-intervention values of the dependent variable and the number of hearings (Oct 2020–Jan 2021); and month fixed effects. Šidák–Holm p-values are computed from the analytic p-values produced internally by the Stata command wyoung, based on a single non-bootstrapped estimation of each specification.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$