NBER WORKING PAPER SERIES

NUDGING THE NUDGER: PERFORMANCE FEEDBACK AND ORGAN DONOR REGISTRATIONS

Julian House
Nicola Lacetera
Mario Macis
Nina Mazar

Working Paper 30547
http://www.nber.org/papers/w30547

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2022, Revised November 2023

We are thankful to the Crabtree Foundation for financial support, and to ServiceOntario and Trillium Gift of Life Network for their collaboration. We thank participants to seminars at Carnegie Mellon University, the Data Colada seminar series, INSEAD, IESE, Yale SOM's Spring Decision Making Symposium, the University of Minnesota, the University of Toronto, the University of Genova and the Society for Judgment and Decision Making Conference in 2022 for their feedback. The study received approval from the Research Ethics Board of the University of Toronto (Protocol no. 32650) and the Johns Hopkins Homewood IRB (HIRB00005769), and was pre-registered at the AEA RCT Registry (No. 0001974). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Nudging the Nudger: Performance Feedback and Organ Donor Registrations
Julian House, Nicola Lacetera, Mario Macis, and Nina Mazar
NBER Working Paper No. 30547
October 2022, Revised November 2023
JEL No. C93,D90,H41,I10,J45,M50

ABSTRACT

In a pre-registered randomized controlled trial conducted over 2.5 years and involving nearly 700 customer-service representatives (CSRs) from a Canadian government service agency, we studied how providing CSRs with repeated performance feedback, with or without peer comparison, affected their subsequent organ donor registration rates. The feedback resulted in a 25% increase in daily signups compared to otherwise equivalent encouragements and reminders. Adding benchmark information about peer performance did not amplify or diminish this effect. We observed increased registration rates for both high and low performers. A post-intervention survey suggests that CSRs in all conditions found the information included in the treatments helpful and motivating, and that signing up organ donors makes their job more meaningful. Performance feedback with benchmark information was the most motivating and created the least pressure to perform, whereas feedback without benchmark increased perceived pressure.

Julian House
Behavioural Insights Unit,
Treasury Board Secretariat,
Government of Ontario
Toronto, ON M7A 1Y7
Canada
Julian.House@ontario.ca

Nicola Lacetera
University of Toronto
Institute for Management and Innovation
3359 Mississauga Road, Room KN 235
Mississauga, ON L5L 1C6
and NBER
nicola.lacetera@utoronto.ca

Mario Macis
Johns Hopkins University
Carey Business School
1717 Massachusetts Ave, Office 359
Washington, DC 20036
and IZA
and also NBER
mmacis@jhu.edu

Nina Mazar
Boston University
Questrom School of Business
595 Commonwealth Ave
Boston, MA 02215
nina@ninamazar.com

1. Introduction

A shortage of organs for transplantation exists in most countries around the world, resulting in untold human suffering and large medical costs (e.g., dialysis treatment) as people wait to receive an effective medical procedure. In Canada, for example, 2,936 organ transplants were performed in 2022. However, by the end of that same year, 3,777 people remained on the waitlist. Moreover, 701 patients dropped off the waitlist because they either died or deteriorated to a point where they were no longer eligible for a transplant.[1] The imbalance between supply and demand occurs despite the broad social support and positive attitude that the donation of organs enjoys virtually everywhere, and the expressed intention of most people to consider donating their organs upon death. For example, in Ontario, Canada's most populous province, 90% of residents support organ donation, but only 35% are registered as organ donors. The organ shortage results in one resident dying every three days while waiting for a transplant (Trillium Gift of Life 2021).

Many countries promote donations through regulatory provisions (e.g., priority rules, presumed consent or prompted-choice systems; see Kessler and Roth 2012, 2014) as well as educational and other initiatives, which include public service announcements and informational campaigns. Most of these efforts have focused directly on potential donors. In many organ donor registration contexts, however, individuals interact with intermediaries, whose influence has been largely neglected. For example, in the United States, people can join the organ donor registry when applying for a driver's license at Department of Motor Vehicle offices. Similarly, most of the organ donor registrations in Ontario (pre-Covid-19 pandemic: 85%) occurred during in-person visits to ServiceOntario centers (Trillium Gift of Life, 2027), which through their customer service representatives (CSRs) provide a wide range of services to residents ranging from driver and vehicle licensing to public health insurance registration and business licensing.

In this paper, we report on a randomized controlled trial (RCT) conducted in collaboration with ServiceOntario's in-person centers. These centers process about 25 million transactions annually; 20% of these transactions occur in the 82 publicly-owned offices (28% of all offices; the rest of the offices are privately owned), which employ several hundred CSRs. CSRs at publicly-owned

[1] The source of this information is https://www.cihi.ca/en/summary-statistics-on-organ-transplants-wait-lists-and-donors, accessed on 10/31/2023.

service centers receive a fixed salary and (unlike CSRs at privately-owned offices) do not earn a commission for the transactions that they process.

Because of their unique role, ServiceOntario CSRs are ideally positioned to promote organ donor registrations. Indeed, operational policy instructs them to implement a prompted-choice procedure with all customers, and CSRs regularly receive email reminders, about two to four times a year, encouraging them to prompt their customers to consider registering as donors. However, CSRs' performance on this task is typically neither assessed nor communicated. Our study considers the role of these intermediaries in motivating organ donations. Importantly, the CSRs at publicly-owned ServiceOntario centers do not have explicit sources of extrinsic motivation for this task, because the registration of new donors does not affect their compensation or performance reviews (Robitaille et al. 2021).

We focus on the potential role of information and, more specifically, feedback to CSRs about their performance as a motivator to increase registration rates. CSRs are plausibly unaware of their actual performance on this activity and how it compares with their colleagues. At the same time, ServiceOntario and similar agencies elsewhere can readily provide performance feedback for this socially important part of the CSRs' job and would presumably be interested in doing so, if it proved effective. Previous research has shown that receiving private feedback about one's individual performance can affect subsequent effort and outcomes (e.g., Bandiera et al. 2015). Knowledge (or perceptions) of the typical behavior of others can also affect one's behavior (Duflo and Saez 2002, Munshi and Myaux 2006). In particular, people respond to information about their performance relative to that of their peers (Alcott 2011, Croson et al. 2009), and the evidence indicates that this is not simply due to material benefits (Ball et al. 2001) or competitive preferences (Charness and Grosskopf 2001; Charness and Rabin 2002). In fact, studies have shown that people care about (and exert effort based on) their relative position ranking even when this does not produce extrinsic benefits such as financial rewards or social status (Charness et al. 2011, Tran and Zeckhouser 2012). These findings suggest that intrinsic motivation and self-image concerns may explain why information on rankings can affect people's behavior (Bénabou and Tirole 2006, Gneezy et al. 2011). However, the effect of this information is not necessarily positive. For instance, high performers might "relax" and low-performers "give up" when informed of their relative ranking, leading employees to reduce their performance upon receiving feedback (Bandiera et al. 2013; Allcott and Kessler 2019). Moreover, in a medical context, Reiff

et al. (2022) have shown that peer comparison may negatively affect outcomes that are often not measured, such as job satisfaction and burnout. Overall, the existing theories and evidence do not provide precise predictions about the effect of motivating ServiceOntario's CSRs with performance feedback in the absence of extrinsic incentives.[2]

The key intervention in our study consisted of ServiceOntario's leadership providing CSRs at all government-owned centers information via e-mail about their individual organ-donor signup performance three times over a span of one year (June 2017, January 2018, and June 2018), with or without a regional benchmark. We then measured the effect of the interventions on these CSRs' organ-donor registrations over the subsequent weeks and months, compared to a third condition that provided only a typical reminder about the importance of asking customers to join the organ-donor registry. Specifically, we randomly assigned 694 CSRs to one of three groups: (1) a "*standard-reminder*" (R) condition in which CSRs received a typical e-mail communication from ServiceOntario that reminded them of the role they play in encouraging people to join the registry and included basic up-to-date organ donor statistics, tips and facts designed to help CSRs be more effective when soliciting registrations, as well as an appeal to help further in this mission; (2) an "*individual feedback*" (IF) condition with an e-mail that, in addition to the standard reminder, included information on the CSR's individual organ donor signup performance over the previous six months (absolute and per one-hundred customer interactions); and (3) a "*regional benchmark*" (RB) condition in which in addition to individual performance, the e-mail included the performance (i.e., average and 80th percentile) of all CSRs operating in the region where the office of a given CSR was located. To minimize informational spillovers between CSRs in different conditions, we randomly assigned the conditions by office (i.e., all CRSs in any given office would receive the same condition throughout the entire experiment). Figure 1 shows examples of the intervention e-mails.

The data at our disposal contain daily observations for each CSR and span not only the experimental period but also a pre- and post-experiment period for a total time span from November 2016 to April 2019. We used these data both to compute the individual and regional-

[2] Performance evaluations relative to peers may also result in sabotage and other types of unethical behavior (Edelman and Larkin, 2014; Charness, Masclet and Villeval 2014). However, ServiceOntario's customer representatives' efforts and tasks are largely independent of each other and independent of their peers. Thus, this possibility does not apply to our context.

level statistics included in the intervention e-mails and to estimate the effects of the interventions. The final dataset for analysis includes 265,475 observations on 694 CSRs, operating in 82 offices on 745 distinct days.

Overall, providing performance feedback resulted in 0.15 additional signups per CSR per day compared to encouragement alone, corresponding to a 25% increase over the baseline. There was no difference between individual performance feedback and individual performance feedback plus the regional benchmark. The effect was particularly pronounced in the few weeks immediately following receipt of a treatment e-mail, but a substantial effect persisted for several months. Furthermore, the increase in organ donor signups was observed among both high- and low-performance CSRs.

To obtain additional insights and help interpret the results, we conducted a post-intervention survey among the CSRs. In particular, we were interested in learning whether participants paid attention to our intervention e-mails and whether the feedback interventions had any effects on CSRs' motivations and perceptions. Almost all CSRs recalled receiving the intervention e-mails, suggesting that they paid consistent attention to them regardless of their content, and that performance variations were likely due to the specific content of the messages. A significant proportion of CSRs, across all treatment conditions, felt that enrolling customers into the donor registry added value to their job, and that by consistently prompting and explaining the registry, they could obtain more sign-ups. However, there were nuanced perceptions regarding the types of performance feedback e-mails. CSRs in the individual feedback condition reported a higher perception of feeling pressured to ask. In contrast, feedback with benchmark information had a lower reported feeling of pressure and a stronger motivating effect than only individual feedback. Similar to Reiff et al. (2022), we conclude that studies of performance feedback should not only evaluate outcomes but also track measures of motivation and perceptions to ensure a comprehensive understanding of these interventions.

In the next section, we describe the context and the experimental design. In Section 3 we describe the data and present our empirical findings, and in Section 4 we discuss the results and conclude.

2. Experimental design and data

2.1 Institutional background

According to the 2017 Annual Report of the Office of the Auditor General of Ontario, ServiceOntario's in-person centers processed about 25 million transactions annually, 20% of which were handled by government-owned offices. ServiceOntario employs several hundred customer service representatives (CSRs) in their 82 offices throughout the province. CSRs at government-owned service centers receive a fixed salary and do not earn a commission for the transactions they process.[3] CSRs regularly receive reminders, about two to four times a year, via e-mail to support organ donor registrations by asking customers whether they would like to register to be organ donors. In addition to not receiving commission pay for signing customers to the registry, this specific activity does not affect the CSRs' performance reviews. There is substantial variation in signup rates across ServiceOntario CSRs (Robitaille et al., 2021). Due to a host of factors, including limited mental bandwidth, time pressure, insufficient salience of organ donor registration, a desire to avoid confrontation, as well as the absence of any material incentives for CSRs to improve their signup rates, some CSRs may fail to solicit customers consistently, or they may be less than convincing when they do. Conversely, other CSRs may be especially effective thanks to their intrinsic motivation, communication skills, or other individual traits.

2.2 Treatment conditions

The randomized controlled trial involved all government-owned ServiceOntario centers that were active at the time of implementation. We designed the following experimental conditions:

- **"Reminder" (R)**: CSRs in this condition received an e-mail including basic statistics about organ donations in Ontario, a reminder of the role that ServiceOntario plays in adding individuals to the registry, and an appeal to CSRs to help further this mission and exert effort on that activity.
- **"Individual Feedback" (IF)**: The e-mail had the same information as in Condition R plus the following additions: The number of customers the specific CSR served in the previous six

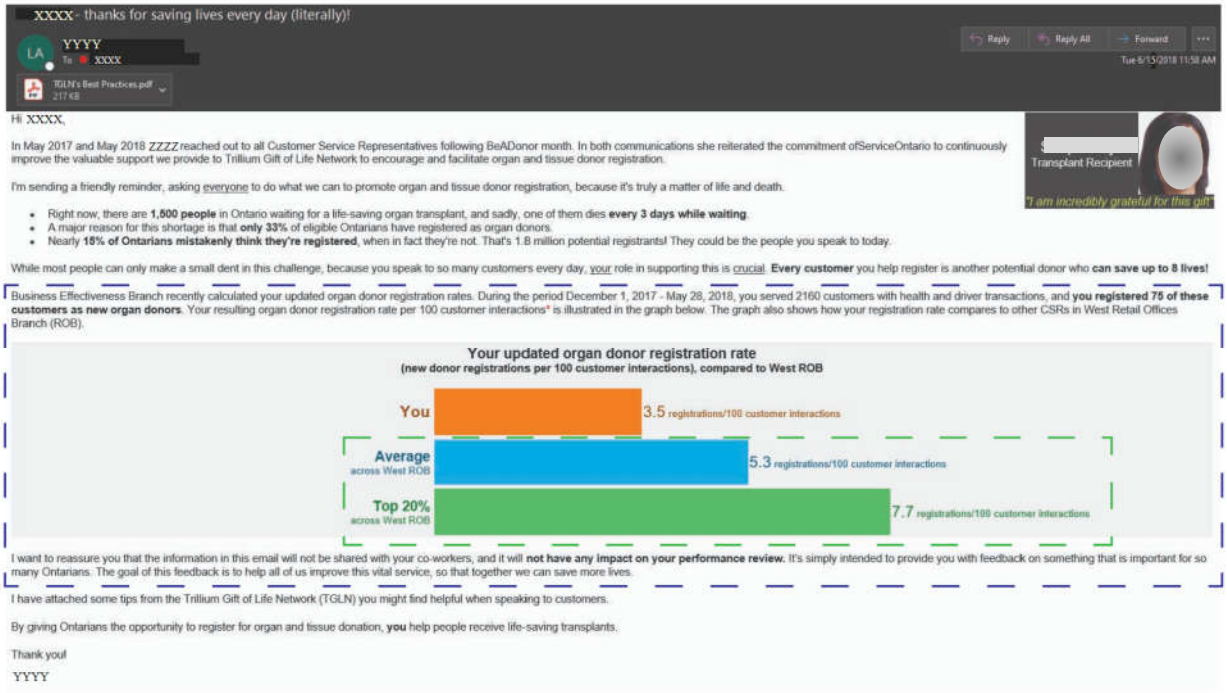
[3] Although we do not know how CSRs at privately-owned service centers are paid, privately-owned service centers themselves receive a commission for each transaction. Of the just under three hundred service centers in Ontario, approximately 72% are privately-owned.

months and how many of those customers the CSR signed up to the organ donor registry, in absolute terms and for every one-hundred customer interactions (the latter information was expressed both numerically and graphically).

- **"Benchmark" (RB)**: In addition to the same information as in Condition IF, the e-mail included the regional average and 80th percentile for the number of signups per one-hundred-customer interactions in the previous six months (graphical).

Figure 1 shows examples of the three e-mails. These e-mails were sent on June 20, 2017 (1st intervention), January 29, 2018 (2nd intervention), and June 15, 2018 (3rd intervention).

Figure 1: Sample of treatment emails sent to CSRs



Notes: The figure reports a snapshot of the emails that the CSRs received, according to their assigned treatment condition. The content of the emails was the same for all CSRs except for the individual feedback and benchmark statistics, shown in dashed blue and green boxes, respectively. The names of the email sender and receiver, as well as of the government executive who sent the information emails a few weeks before each intervention date, are redacted, and the date on the top right of the figure corresponds to the day in which CSRs received the email for the third intervention. All emails included an attached PDF "tip sheet" designed to help CSRs become more effective at prompting donations (i.e., by answering potential customer questions with accurate facts). We covered the face and name of the transplant received at the top right of the email for confidentiality reason, but both the picture and name were visible to the CSRs.

About two weeks before each intervention date, all CSRs, regardless of their treatment condition, received an e-mail from a senior provincial government executive announcing that they

would soon receive communication regarding the organ donor registry. The e-mail did not specify what information and did not mention that the communication to come would be part of an experiment; its primary purpose was to increase the CSRs' attention toward, and likelihood of opening the intervention e-mails. Strictly speaking, the reminder condition "R" is not a pure control because it includes a message that could affect performance in the activity of interest.

Following the RCT, we administered a survey to obtain individual characteristics of the CSRs (e.g., age, gender, and tenure), gauge whether they had paid attention to the intervention e-mails in the previous eighteen months and how they perceived the messages, and investigate any effect of the interventions on their motivation and perceptions (Reiff et al. 2022). All CSRs who were active at the time of the survey and who were part of the field experiment received an e-mail inviting them to complete the survey, and 283 completed it (40.8% overall completion rate; 36.7%, 36.9%, and 48.2% in conditions R, IF, and RB respectively).

2.3 Randomization

To minimize informational spillovers between CSRs in different conditions, we randomly assigned the conditions by office (i.e., all CSRs in any given office would receive the same experimental condition). We also stratified the randomization by the four regions in which ServiceOntario partitions the Province: North, East, West, and Center because these regions present socio-economic differences. Condition assignment by office also complied with requests from our partner organization to maintain equality of treatment within a specific location. One challenge was that some CSRs work in more than one office; within an intervention wave, about 30% of CSRs did. We chose to assign each of these multi-location workers to the office (and thus, condition) where they typically spent more time in the months immediately preceding the intervention. ServiceOntario staff also assisted in determining the assignment. This non-systematic deviation from full adherence to our design added "natural" variation, allowing us to control for office fixed effects in the econometric analyses (see below). When we assigned CSRs to experimental conditions, the dataset at our disposal (as of April 30, 2017) included 565 individual CSRs in 79 offices. There were 24 offices and 177 CSRs in Condition R (7.4 CSRs per office on average), 27 offices and 198 CSRs in Condition IF (7.3 CSRs per office), and 28 offices and 190 CSRs in Condition RB (6.8 CSRs per office). However, we have outcome data for 82 offices.

Thus, three offices were not used to assign conditions, although we subsequently observed CSRs working at those locations on some days.

2.4 Estimation

Equation (1) below shows the main econometric model that we estimate:

$$\begin{aligned} Y_{cdmy} = \alpha_0 &+ I_{IF}\alpha_{IF} + I_{RB}\alpha_{RB} + I_R[I_{t=1}\,\beta_{R1} + I_{t=2}\beta_{R2} + I_{t=3}\beta_{R3}] \\ &+ I_{IF}[I_{t=1}\beta_{IF1} + I_{t=2}\beta_{IF2} + I_{t=3}\beta_{IF3}] \\ &+ I_{RB}[I_{t=1}\beta_{RB1} + I_{t=2}\beta_{RB2} + I_{t=3}\beta_{RB3}] + \gamma X_{cdmy} + \mu_m + \nu_y + \eta_c \\ &+ \varepsilon_{cdmy}. \end{aligned} \tag{1}$$

Y, the outcome variable, is either the number of signups by CSR *c* on day *d*, or a binary indicator for whether a CSR *c* made at least one new signup to the organ registry on day *d*. The variables I_M, I_{MF}, I_{MFB} are binary indicators for whether a CSR was in conditions R, IF or RB, respectively (value of one if they were, and zero if they were not). $I_{t=1}$, $I_{t=2}$ and $I_{t=3}$ take a value of one if an observation is in the period after the first, second, or third intervention wave, respectively, and zero in any other period. Therefore, the estimates of the "β" parameters indicate the average differences between the number of signups by CSRs in a given condition and post-intervention period, and the signups of CSRs in the same condition in the pre-intervention period. For example, the estimate $\hat{\beta}_{IF2}$ represents the average difference in daily signups between the period after the second intervention wave (and before the third) and the period before the first intervention wave for CSRs in condition IF. Linear combinations of the parameters provide other treatment effects of interest. Within a given *condition* (e.g., condition RB), the difference between two "β" estimates represents the differential impact of a treatment in a given period as compared to the pre-intervention period – a "difference in difference" within a condition; for example, $\hat{\beta}_{RB3} - \hat{\beta}_{RB2}$ estimates the differential impact of the second and third intervention waves for condition RB with respect to the period before the first intervention wave for that same condition. Within a given post-intervention *period*, we can establish the differential treatment effect between conditions by taking the difference between parameter estimates for a given period and different conditions. For instance, $\hat{\beta}_{RB1} - \hat{\beta}_{IF1}$ estimates how condition RB changed signups in the first post-intervention period compared to the pre-intervention period, relative to the same change for condition IF – a within-period, between-condition difference-in-differences.

If we take condition R as the reference case, a natural exercise is to measure the differential impact of the feedback conditions with respect to the encouragement-email reminder. In a more fine-grained distinction, we split the pre- and post-intervention periods in intervals of about sixty to seventy days each. By looking at shorter sub-periods separately, we can assess if any effect was higher immediately after the reception of the intervention e-mails or stable throughout an intervention period.

The vector X_{cd} represents control variables. The data do not include many details about each CSR or their offices, but we have some relevant control variables. These include the total number of unique customer interactions a CSR had on a given day. The number of daily interactions may indicate the productivity of a CSR. However, for the most part, CSRs' daily volumes depend on factors beyond their control, such as the number of hours worked in a day and the haphazard assignment to particular clients or types of services. We also derive a measure of CSR experience within our sample: the number of days of activity since we begin to observe a given CSR in our data. All models include month (μ_m) and year (ν_y) fixed effects to account for time and seasonal trends. Finally, η_c indicates CSR-level fixed effects, which we include in most specifications to control for any time-invariant, unobserved individual differences (e.g., experience, personality traits, and the like). Although the assignment to experimental conditions was at the office level, some CSRs worked in more than one office during the study period while keeping the same condition assignment throughout. This variation allows us to add office-level fixed effects in some specifications.

2.5 Constructing signup performance measures

CSRs' organ donor registration performance is opportunity-dependent: namely, it is only possible for CSRs to register those customers with whom they interact. As customer interaction volumes vary across time and between ServiceOntario centers, to provide CSRs with meaningful, relative feedback it was necessary to compute organ donor registration rates, which calculated the number of organ donors registered per hundred client interactions. This computation required several steps and assumptions, given some peculiarities of the internal data collection processes at ServiceOntario and the administrative data structure.

A CSR's *interaction* with a customer may include one or more *transactions*. ServiceOntario has a prompted-choice policy for the organ donor registry; at the end of each customer interaction, a CSR is supposed to ask customers to consider joining the registry. ServiceOntario keeps track of all transactions performed by CSRs, but the records are kept in two separate systems. For health-related transactions (services related to public health insurance cards and the organ donor registry), data are recorded at the level of customer interactions (one row per customer visit, potentially recording multiple health transactions performed for that customer). Differently, data for all non-health-related transactions are recorded manually by CSRs as daily counts, displaying the number of each transaction type (e.g., driver's license renewal, license plate transfer, etc.) performed by a given CSR on a given day, and do not include a count of interactions (i.e., number of unique clients served).

To assemble complete information on each CSR's daily activity, we had to express these two data sources at the same level of aggregation. The first step of this procedure was to aggregate the customer interaction data for health-related activities at the CSR-day level. In particular, the count of entries for a given CSR on a given day provided us with the number of unique customers to whom they provided health-related services (interactions). Also, the total number of new organ-donor registry entries on each day measures a CSR's "absolute" signup performance.

Second, we merged these data, at the CSR-day level, with the data on non-health-related services, which do not directly count customer interactions. To estimate the number of daily non-health related interactions, we determined that, on average, a given customer interaction involves 1.3 health-related transactions, and assumed in consultation with ServiceOntario administrators this to be a reasonable per-interaction rate for non-health services as well. Therefore, we divided the daily non-health services by 1.3 to obtain an estimate of unique customer interactions, and added these to the daily health-related interactions for a given CSR-day to obtain total daily unique customer interactions.

For each CSR, we then added up all customer interactions and all organ donor registrations performed by each CSR during a given observation period and reported these two variables to CSRs in the IF and RB emails. Their number of registrations per 100 interactions during that period was also reported to CSRs in the e-mails as the "organ donor registration rate". To calculate the regional benchmark statistics for the RB e-mails, we computed the averages and 80^{th} percentiles of the total individual signups (per 100 overall interactions) in each of Ontario's four regions.

2.6 Data quality checks

We performed a series of checks to determine the reliability of the data and the robustness of our findings to alternative assumptions and computations of certain key variables. First, we compared the variables that we used and constructed from the experimental data with month-level transaction statistics compiled separately by the Ministries on whose behalf ServiceOntario provides services (i.e., transactions related to drivers and vehicles are recorded by the Ministry of Transportations Licensing and Control System, and health transactions are recorded by the Ministry of Health's Registered Persons Database), the latter of which is available for both publicly and privately owned ServiceOntario centers. The graphs in Appendix Figure A1 show a close overlap between these independently sourced data.

The concordance between these data sets was particularly reassuring as a means of independently validating the accuracy of the data, given both the potential for human error resulting from the manual recording of non-health transactions by CSRs, and a clerical error by back-office ServiceOntario staff who had compiled the experimental data, in which some non-health transactions types were initially inadvertently excluded from daily counts of CSRs' non-health transactions. The potential threat to our study posed by this clerical error was quantitatively assessed as a second robustness check. The clerical error occurred only in the data extract used to populate the Wave 1 e-mails and consisted in the exclusion of some transaction types for some CSRs. This resulted in Wave 1 IF and RB to overestimate organ donor registration rates by a few decimal points, which is unlikely to have been noticed by CSRs. The error also affected 58 CSRs assigned to the RB condition who as a result received slightly imprecise ranking information relative to these regional benchmarks. Reassuringly, excluding these CSRs for the period between the first and the second intervention wave (5,206 CSR-day observations) does not affect the estimates meaningfully (Section 3.2 and Table 3 below).

A third check was made possible by the fact that the month-level transaction statistics included years of data prior to our experiment. As we had confirmed that the month-level statistics aligned very closely with our experimental data, we could also use the month-level data to assess whether the time-effect trends observed during the experiment were in any way atypical, and we found no evidence of that in either the public offices (the sites of our experiment) or the private offices.

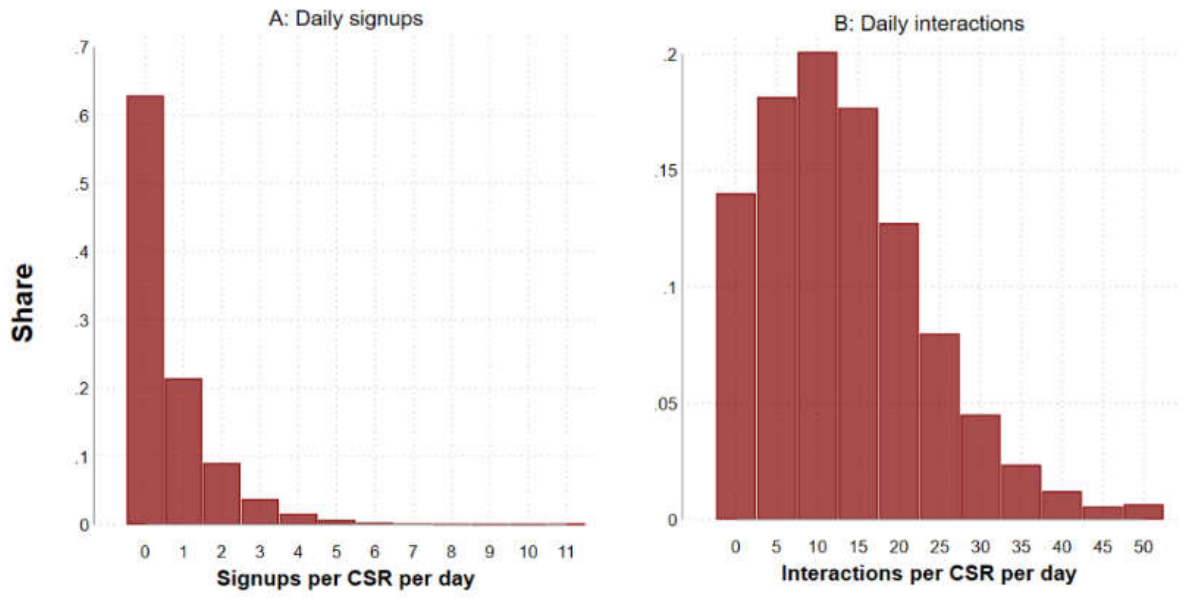
A fourth check for the reliability of the data concerned the calculation of the number of unique customers served, per day and in total, by each CSR. Again the results are robust to alternative computation choices. The details are in the Appendix (Section A1, Table A4).

2.7 Data description

The full dataset includes CSR-day-level information from November 1, 2016 through April 30, 2019, amounting to 295,884 observations. After excluding the observations that reported mail-in activities as explained above, the final dataset for analysis thus includes 265,475 observations on 694 CSRs, operating in 82 offices on 745 distinct days. We winsorized both the daily signup and daily unique customer interaction counts at the 99.9th percentile to correct for implausible large values that may be the result of reporting errors (such as reporting mail-in registrations combined with in-person signups). Neither the use of the raw counts nor the exclusions (as opposed to the winsorization) of the values in the top 0.5th percentile alters the estimates of interest meaningfully. The average (winsorized) new customer signups to the organ donor registry and total interactions per day over the entire period of observation are 0.65 (range: 0-11) and 13.4 (range: 0-50.7), respectively (Figure 2). The sample is balanced between conditions in the pre-intervention period (from November 2016 through June 15, 2017).[4]

[4] At the individual CSR level, the F statistics for the joint significance of differences in the share of women, tenure, average signups, and average total transactions per day could not reject the hypotheses of no differences between conditions.

Figure 2: Distribution of the numbers of daily signups and customer interactions per CSR over the entire period of observation



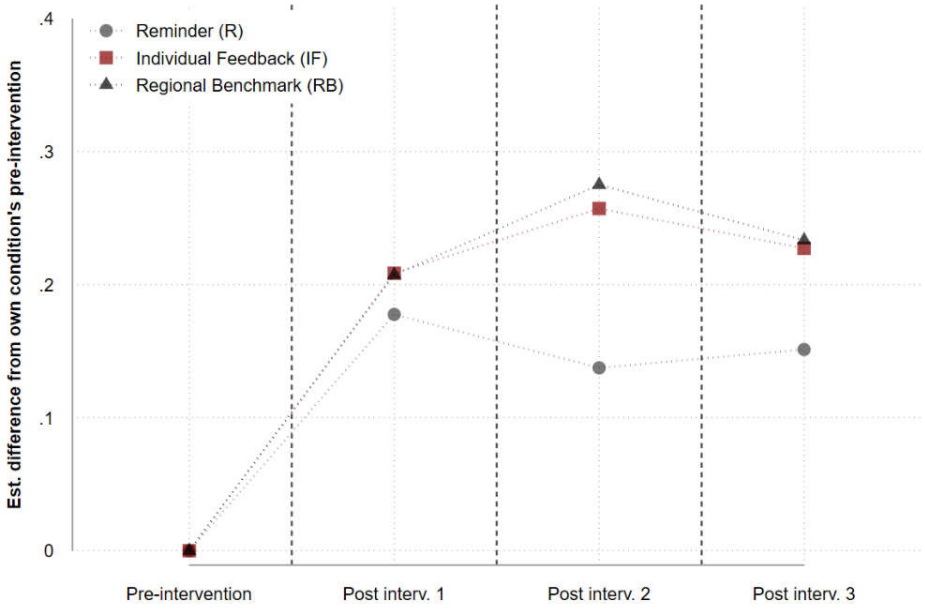
Notes: The figures report the empirical distribution of signups (panel A) and customer interactions (panel B) per day and CSR between November 1, 2016 and April 30, 2019. The figures exclude the 0.1th percentile of highest values for daily signups and interactions (values greater than 11 and 50.7, respectively). To build the graph, we winsorized the number of signups and interactions at the 99.9 percentile, and rounded the values of daily interactions to the closest integer.

3. Results

3.1 Main findings

The extended collaboration with Service Ontario made it possible to run multiple interventions and to have data over a long period (2.5 years) to study both short-term and longer-term impacts of the interventions. Figure 3 reports the nine estimated "β" coefficients from Equation (1) above. The values are from Column (2) of Table 1. Daily individual signups increased significantly in all conditions with respect to the pre-intervention period.

Figure 3: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups in each intervention period



Notes: The figure reports the estimated average changes in daily organ donor signup per CSR after each of the three interventions, compared to the average daily signups in the pre-intervention periods for each condition (normalized to zero). The estimates from which this graph is derived are in column 2 of Table S1 in the Supplementary Material. Each intervention consisted of an e-mail whose content differed according to the experimental conditions to which a CSR was randomly assigned. The e-mails were sent on June 20, 2017 (1st intervention), January 29, 2018 (2nd intervention), and June 15, 2018 (3rd intervention).

In particular, while no significant differences emerge between conditions in the period following the first email intervention, receiving information about one's performance, with or without benchmark, led to an additional, statistically significant increase over the control condition R, in the second and the third period, of about 0.13-0.15 signups per day. With a pre-intervention overall average of 0.6 daily signups per CSR as the reference, this represents an increase of roughly 25% compared to providing basic information and encouragement. The findings suggest that continuing to provide performance feedback leads CSRs to maintain, and possibly slightly increase, the gains in signup activities from the initial experience with receiving this feedback.

The addition of benchmark performance feedback does not lead to different outcomes than just providing individual performance feedback.

Table 1: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups, in each intervention period: Main regression estimates

Outcome variable:	Daily signups		
	(1)	(2)	(3)
IF	-0.051		-0.051
	(0.087)		(0.071)
RB	-0.114**		-0.109*
	(0.057)		(0.058)
R:1st int.	0.217***	0.178***	0.215***
	(0.058)	(0.040)	(0.033)
R:2nd int.	0.147**	0.137***	0.152***
	(0.057)	(0.050)	(0.046)
R:3rd int.	0.156**	0.151**	0.166***
	(0.072)	(0.065)	(0.060)
IF:1st int.	0.263*	0.209***	0.260***
	(0.137)	(0.050)	(0.055)
IF:2nd int.	0.309**	0.257***	0.294***
	(0.121)	(0.056)	(0.051)
IF: 3rd int.	0.280**	0.227***	0.269***
	(0.120)	(0.072)	(0.068)
RB:1st int.	0.239***	0.207***	0.229***
	(0.070)	(0.043)	(0.034)
RB:2nd int.	0.319***	0.275***	0.311***
	(0.077)	(0.051)	(0.044)
RB:3rd int.	0.282***	0.233***	0.265***
	(0.081)	(0.062)	(0.063)
Constant	0.396***	-0.289**	0.339***
	(0.088)	(0.114)	(0.100)
CSR fixed effects		x	
Office fixed effects			x
IF:1st int.-R:1st int.	0.047	0.031	0.044*
RB:1st int.-R:1st int.	0.023	0.030	0.014*
RB:1st int.-IF:1st int.	-0.024	-0.001	-0.030
IF:2nd int.-R:2nd int.	0.162	0.120**	0.142***
RB:2nd int.-R:2nd int.	0.172**	0.138***	0.159***
RB:2nd int.-IF:2nd int.	0.010	0.018	0.017
IF:3rd int.-R:3rd int.	0.124	0.076	0.103***
RB:3rd int.-R:3rd int.	.0.126*	0.082**	0.099***
RB:3rd int.-IF:3rd int.	0.001	0.006	-0.004
Observations	265,475	265,475	265,475
R-squared	0.145	0.297	0.172

Notes: The table reports estimates from linear regressions where the unit of observation is a CSR on a given day in which that CSR is active. The regressors are binary indicators for experimental conditions and interactions between experimental conditions and interventions. R indicates the reminder e-mail condition, IF the reminder + individual feedback condition, and RB the reminder + individual feedback + regional benchmark condition. The estimated parameter on a given interaction term (e.g., RB: 2nd int.) represents the estimated difference in daily signups between the period that the interaction term identifies, and the pre-intervention period for the same condition. The bottom part of the table reports relevant differences between estimated parameters. The regressions include variables that measure the number of days a CSR was present in the sample at any given date, its square, the total number of interactions of a CSR in a given day, and year and month fixed effects. The parameter estimates in Figure 2 in the main manuscript correspond to those reported in column 2 of this table. Standard errors are clustered both at the CSR level and at the level of office-intervention period. * p<0.1, ** p<0.05, *** p<0.01.

The fact that the difference between the reminder and the feedback conditions emerges after the second intervention, and not after the first, reinforces our interpretation that this effect was driven specifically by feedback information. If the main driver of the behavioral change was just the increased salience of the signup activity that performance feedback provided, and not its informational content, one should have expected larger differences from the Reminder condition to emerge after the first intervention wave when, arguably, the salience or novelty effect should have been stronger.

3.2 Additional analyses and robustness checks

Tables 2 and 3 report the estimates from additional analyses to investigate other potential effects of the interventions and assess the robustness of the estimates from our main specification. The parameter estimates in Table 2 are from the model described in equation (1) but with different left-hand-side variables. In Columns (1) and (2), the outcome variable is the ratio between daily signups and daily customer interactions (multiplied by 100). This is an alternative way to control for the overall activity of a CSR. The estimates in Column (1) are from a model without the number of daily interactions among the regressors, whereas those in column (2) are from a model that also includes daily interactions on the right-hand side; the estimates of interest are very similar. The estimated treatment effects show the same patterns (in size and statistical significance) as those in Table 1 above.

Column (3) of Table 2 reports results when we use a binary indicator for having signed up at least one customer in a given day as the outcome of interest. The estimates suggest that the intervention had an impact both on the extensive margin (more CSRs signing customers to the organ donor registry) and on the intensive margin (a higher number of signups per CSR).

In Column (4), the estimates are from a model where the outcome is the number of total daily transactions; as mentioned above, a CSR may provide more than one service (transaction) to the same client. One concern is that our various treatments may negatively affect the overall activity of a CSR because, for example, they might spend more time talking to customers about the organ donor registry in an attempt to sign them up. The estimates suggest that this substitution or "crowd out" effect did not occur.

Table 2: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups in each intervention period: Alternative outcomes

Outcome variable:	100*Daily signups/interactions		*I* (Daily signups>0)	Daily transactions
	(1)	(2)	(3)	(4)
R:1st int.	1.229***	1.183***	0.057***	1.130*
	(0.278)	(0.270)	(0.014)	(0.629)
R:2nd int.	0.884**	0.843**	0.038**	1.164
	(0.367)	(0.348)	(0.016)	(0.873)
R:3rd int.	0.630	0.631	0.035*	0.004
	(0.472)	(0.455)	(0.020)	(1.073)
IF:1st int.	1.444***	1.415***	0.065***	0.641
	(0.358)	(0.338)	(0.016)	(1.001)
IF:2nd int.	1.378***	1.381***	0.072***	-0.207
	(0.405)	(0.384)	(0.020)	(1.199)
IF: 3rd int.	1.037*	1.035**	0.062**	-0.118
	(0.533)	(0.506)	(0.024)	(1.325)
RB:1st int.	1.584***	1.541***	0.078***	1.142*
	(0.272)	(0.259)	(0.014)	(0.633)
RB:2nd int.	1.989***	1.948***	0.090***	1.125
	(0.361)	(0.347)	(0.017)	(0.805)
RB:3rd int.	1.631***	1.595***	0.076***	0.963
	(0.475)	(0.453)	(0.021)	(1.116)
Constant	1.595*	0.441	-0.074*	31.146***
	(0.845)	(0.844)	(0.041)	(1.579)
CSR fixed effects	x	x	x	x
IF:1st int.-R:1st int.	0.215	0.231	0.008	-0.489
RB:1st int.-R:1st int.	0.356	0.358	0.021	0.012
RB:1st int.-IF:1st int.	0.141	0.146	0.013	0.501
IF:2nd int.-R:2nd int.	0.494	0.538	0.033*	-1.372
RB:2nd int.-R:2nd int.	1.105***	1.106***	0.051***	-0.39
RB:2nd int.-IF:2nd int.	0.611*	0.567*	0.018	1.333
IF:3rd int.-R:3rd int.	0.406	0.404	0.027	-0.123
RB:3rd int.-R:3rd int.	1.000***	0.964***	0.040**	0.959
RB:3rd int.-IF:3rd int.	0.594*	0.560*	0.014	1.082
Observations	265,475	265,475	265,475	265,475
R-squared	0.118	0.120	0.252	0.357

Notes: The table reports estimates from OLS regressions where the unit of observation is a CSR on a given day in which that CSR is active. The regressors are binary indicators for experimental conditions and interactions between experimental conditions and interventions. R indicates the reminder e-mail condition, IF the reminder + individual feedback condition, and RB the reminder + individual feedback + regional benchmark condition. The estimated parameter on a given interaction term (e.g., RB: 2nd int.) represents the estimated difference in daily signups between the period that the interaction term identifies, and the pre-intervention period for that condition. The bottom part of the table reports relevant differences between estimated parameters. The regressions include variables that measure the number of days a CSR was present in the sample at any given date, its square, the total number of interactions of a CSR in a given day (except in column 1, where we compare a specification with and without this control, in columns 1 and 2, when the outcome includes this variable in the denominator), and year and month fixed effects. Standard errors are clustered both at the CSR level and at the level of office-intervention period. * p<0.1, ** p<0.05, *** p<0.01.

In Columns (1) through (5) of Table 3, the estimates are from regressions where we either excluded part of the sample or controlled for additional variables. First, we dropped the observations pertaining to CSRs whose performance, as described in Section 2.6 above, was miscalculated in a way that ended up assigning them to the wrong side of the two regional benchmarks in the first intervention wave. Second, we restricted the sample to only the CSRs who answered the post-intervention survey. Third, we added an indicator variable for CSR-day observations in which the data report zero transactions and also ran the analyses excluding these observations from the sample. A report of zero transactions may indicate a coding error in the ServiceOntario system or that a CSR was active on a given day but not in direct customer-facing tasks. Finally, we limited the sample to CSRs who never worked on mail-in registrations. This restriction is another way to isolate observations with implausibly high reported daily signups. Column (6) reports estimates from a Poisson model given the discrete-count nature of our primary outcome variable. All columns show results similar to those in Table 1.

Finally, we adopt a different approach to accounting for time and seasonal effects in signup activity and performance. Instead of adding, as regressors, indicators for each month and each year separately, we include indicators for each combination of year and month. This specification is preferable if one believes that the seasonal (month) effects are also different in each year. From the raw data on monthly signups activity (Appendix Figure A1), this does not seem to be the case, but we take this as a further robustness check. Table A3 shows estimates from the main analyses discussed above, but from a model with month-year indicators, combined, as regressors. Appendix Figure A2 replicates Figure 3, again with estimates from this alternative model specification. On the one hand, the parameter estimates imply a slightly more contained increase in signups after the first intervention, and a more continuous increase, with less tapering, in later interventions. On the other hand, the estimated differences in outcomes between treatments in each single period, which are the metrics of key interest in our study, are nearly identical to the ones from the main specification.

Table 3: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups in each intervention period: Robustness sample restrictions and econometric specifications

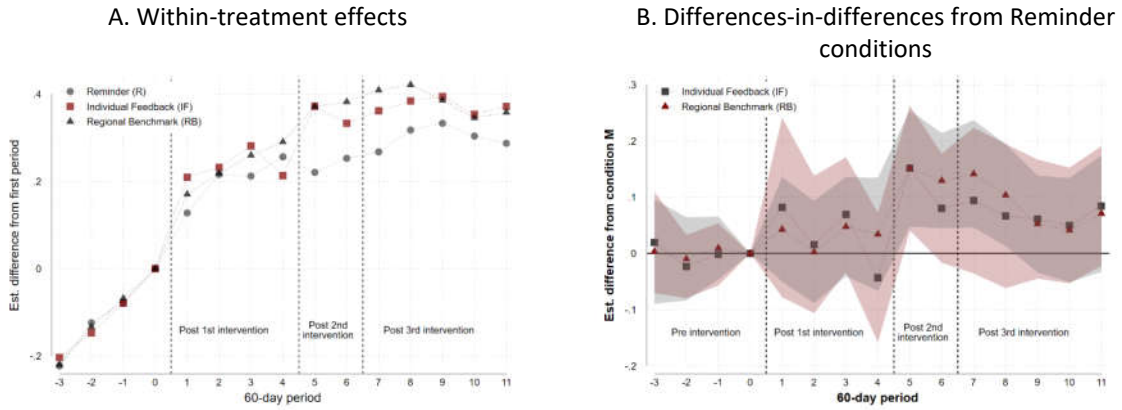
Outcome variable:	Daily signups					
Sample:	Exclude performance bechmark mismatches	CSRs who answered the survey	Full	Exclude CSR observation with no customer interactions	Exclude CSRs with any mailin	Full
	(1)	(2)	(3)	(4)	(5)	(6)
R:1st int.	0.171***	0.156***	0.180***	0.187***	0.178***	0.165***
	(0.039)	(0.054)	(0.040)	(0.042)	(0.040)	(0.033)
R:2nd int.	0.133***	0.110*	0.140***	0.140***	0.137***	0.129***
	(0.049)	(0.057)	(0.048)	(0.050)	(0.050)	(0.041)
R:3rd int.	0.148**	0.121	0.158**	0.148**	0.151**	0.144***
	(0.065)	(0.075)	(0.064)	(0.068)	(0.065)	(0.051)
IF:1st int.	0.217***	0.202***	0.210***	0.225***	0.209***	0.187***
	(0.052)	(0.054)	(0.048)	(0.047)	(0.050)	(0.040)
IF:2nd int.	0.267***	0.221***	0.265***	0.306***	0.257***	0.245***
	(0.058)	(0.064)	(0.055)	(0.062)	(0.056)	(0.049)
IF: 3rd int.	0.238***	0.247***	0.235***	0.249***	0.227***	0.226***
	(0.072)	(0.079)	(0.069)	(0.073)	(0.072)	(0.063)
RB:1st int.	0.207***	0.200***	0.206***	0.217***	0.207***	0.214***
	(0.044)	(0.054)	(0.043)	(0.045)	(0.043)	(0.036)
RB:2nd int.	0.274***	0.259***	0.274***	0.288***	0.275***	0.275***
	(0.053)	(0.061)	(0.051)	(0.055)	(0.051)	(0.043)
RB:3rd int.	0.234***	0.209***	0.235***	0.244***	0.233***	0.252***
	(0.064)	(0.070)	(0.061)	(0.064)	(0.062)	(0.055)
No interactions in a given day			-0.190***			
			(0.031)			
Constant	-0.274**	-0.404**	-0.214*	-0.219*	-0.342**	
	(0.118)	(0.163)	(0.112)	(0.122)	(0.157)	
CSR fixed effects	x	x	x	x	x	x
Model specification	Linear	Linear	Linear	Linear	Linear	Poisson
Observations	259,330	120,609	265,475	242,923	134,953	264,911
R-squared	0.299	0.288	0.298	0.287	0.339	

Notes: Columns (1) through (5) report estimates from OLS regressions where the unit of observation is a CSR on a given day in which that CSR is active. Column (6) reports the estimated marginal effects from a Poisson regression. The regressors are binary indicators for experimental conditions and interactions between experimental conditions and interventions. R indicates the reminder e-mail condition, IF the reminder + individual feedback condition, and RB the reminder + individual feedback + regional benchmark condition. The estimated parameter on a given interaction term (e.g., RB: 2nd int.) represents the estimated difference in daily signups between the period that the interaction term identifies, and the pre-intervention period for the same condition. The regressions include variables that measure the number of days a CSR was present in the sample at any given date, its square, the total number of interactions of a CSR in a given day, and year and month fixed effects. Standard errors are clustered both at the CSR level and at the level of office-intervention period. * p<0.1, ** p<0.05, *** p<0.01.

3.3 The dynamics of the treatment effects

In Figure 4, we report the estimated changes in signups over fourteen subperiods, to further gauge the dynamics of the effects of our intervention. Each subperiod is between sixty and seventy days. The response to the e-mails concentrated mainly in the few weeks immediately after the first treatment for all conditions, and diverged later on with CSRs in the performance feedback conditions IF and RB signing up more customers than their colleagues in the control condition R. To better gauge the differential impact of the two feedback conditions compared to the simple e-mail reminder, Panel B of Figure 4 displays, for each subperiod, the "differences-in-differences" for conditions IF and RB relative to condition R, again setting the two months immediately before the first intervention wave as the reference, as in an event study. The graph shows more explicitly when the feedback treatments were particularly effective compared to just sending an encouragement e-mail.

Figure 4: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups by sub-periods within each intervention

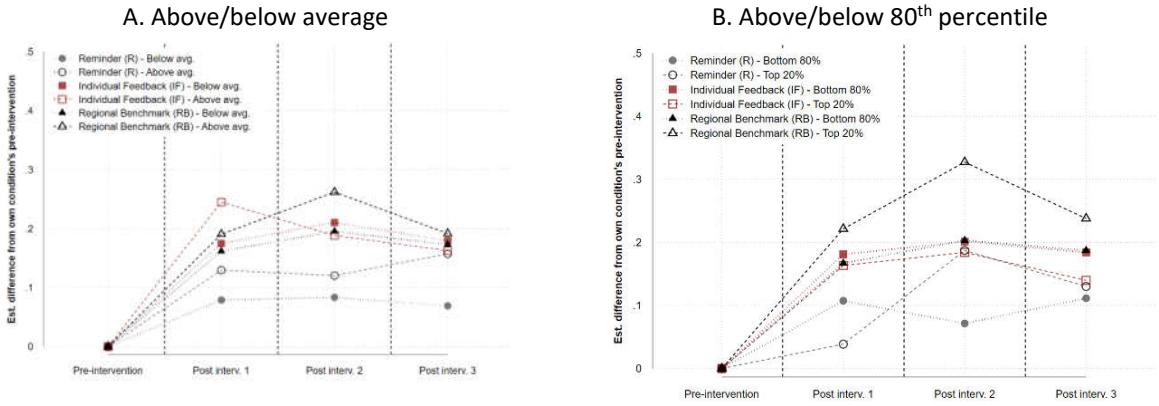


Notes: Panel A reports the estimated average changes in daily organ donor signup per CSR in sub-periods of sixty to seventy days within each post-intervention period, compared to the average daily signups in subperiod immediately preceding the first wave of treatment (subperiod 0 on the x-axis), for each single condition. The estimate in correspondence of point 1 (5, 7) on the x-axis indicates the average performance change between the approximately two months after the first (second, third) intervention date, and the two months before. The values at point 5 represent the estimate difference between the first two months after the second intervention. Panel B displays the same estimates as differences of the IF and RB conditions from the R condition. The values in Panel A are reported in column 1 of Table S2 in the Supplementary Material, and the values in Panel B are reported in column 2 of the same table. Shaded areas represent 95% confidence intervals, with standard errors clustered both at the CSR level and at the level of office-intervention period. Each intervention consisted of an e-mail whose content differed according to the experimental conditions to which a CSR was randomly assigned. The e-mails were sent on June 20, 2017 (1st intervention), January 29, 2018, (2nd intervention) and June 15, 2018 (3rd intervention).

3.4 Heterogeneous effects by prior performance

One question in the literature on the effects of performance feedback is whether this information is equally effective irrespective of CSRs' prior signup performance. One concern is that his feedback is that it might backfire. For example, high-performing individuals might "relax" and reduce their effort, whereas low-performing ones might get discouraged and further reduce their effort. As shown in Figure 5, in our context reminders (R) and the two types of feedback (IF and RB) were effective both for high-performing and low-performing CSRs. Again, the response of the CSRs in the R condition was weaker than in the two feedback conditions, and receiving information about the regional benchmark did not have a large additional effect over just receiving one's individual past signup performance.

Figure 5: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups, by signup performance



Notes: The figures reports the estimated average changes in daily organ donor signup per CSR after each of the three interventions, compared to the average daily signups in the pre-intervention periods for each single condition. In panel A, the estimates per condition are separate between CSRs whose performance in the period immediately before a given intervention was above the regional average (Table S5, column 2), and those with a performance below average (Table S5, column 1). In Panel B, the separation is between the CSRs with pre-intervention performance in the top 20% (Table S5, column 4), and those with performance in the bottom 80% in a given region (Table 5, column 3). Each intervention consisted in sending an e-mail whose content differed according to the experimental conditions to which a CSR was randomly assigned. The e-mails were sent on June 15 2017 (1st intervention), January 29 2018, (2nd intervention) and June 20 2018 (3rd intervention).

3.5 Post-intervention survey

Table 4 reports the average values of the responses to the post-intervention survey questions (or level of agreement with a statement) in the post-intervention survey, by assigned treatment condition of the participating CSRs. For each question, we report the scale or value range of the possible answers. we asked the CSRs whether they remembered receiving one or more of the intervention e-mails (from several months before); only a minority reported that they did not remember, with no differences between experimental conditions, on average. We interpret this as evidence that CSRs paid equal attention to the emails, regardless of their content, and that the performance differences that we observe are due to the specific content of the messages.

The post-intervention survey also shows that large shares of CSRs across all conditions indicated that signing up customers to the donor registry makes their job more meaningful, that customers are more likely to sign up if asked, and that by remembering to ask and by explaining the registry, CSRs have agency to sign up more people. Although small, there were some differences in how CSRs perceived the two types of performance feedback e-mails. Feedback with benchmark information (RB) had a stronger motivating effect than only individual feedback (IF) or encouragement reminder (R). At the same time, CSRs in the IF condition reported a higher perception of feeling pressured to ask. Although creating a sense of urgency may be seen as evidence of the effectiveness of an intervention meant to encourage performance, the lower motivating effect coupled with a stronger feeling of pressure in the IF condition suggest that providing both individual feedback and a benchmark might be preferable both to encourage CSRs and to increase their overall job satisfaction. Providing a frame of reference through the benchmark information may allow CSRs to better interpret the information about their individual activity. For example, the benchmark might have provided the CSRs with evidence of the fact that, in general, signup rates are low, or lower than one may predict; as such, this may have reduced the potentially negative feeling from receiving information about one's own low-looking signup rate without putting it into context.

The survey responses, however, indicate that there was no statically significant difference between respondents in the three conditions in their reported interest in receiving other similar messages in the future.

Table 4: Responses to the post-intervention survey

	Conditions			p-values from pairwise t-tests		
	R	IF	RB	IF - R	RB-R	RB-IF
Customers are more likely to signup if asked (1-5)	3.586 [0.107]	3.452 [0.107]	3.686 [0.092]	0.378	0.481	0.100
Customers are less likely to signup if asked (1-5)	2.300 [0.100]	2.342 [0.102]	2.284 [0.084]	0.767	0.905	0.659
Signing up donors makes job more meaningful (1-5)	3.800 [0.113]	3.479 [0.116]	4.020 [0.086]	0.049	0.117	0.000
Most want to be donors, bit didn't sign up (1-5)	3.029 [0.110]	3.027 [0.103]	3.127 [0.104]	0.994	0.524	0.508
By remembering to ask, I'd sign up more people (1-5)	4.086 [0.101]	3.973 [0.125]	4.402 [0.071]	0.485	0.009	0.002
By explaning registry, I'd sign up more people (1-5)	3.943 [0.106]	3.795 [0.112]	4.029 [0.089]	0.337	0.532	0.098
I remember receiving the email(s) (1-3)	1.457 [0.099]	1.431 [0.095]	1.275 [0.064]	0.846	0.107	0.158
I remember receiving the email(s) (0-1)	0.757 [0.052]	0.764 [0.050]	0.833 [0.037]	0.926	0.220	0.258
I don't remember receiving the email(s) (0-1)	0.029 [0.020]	0.042 [0.024]	0.059 [0.023]	0.675	0.358	0.617
Email(s) included new information (1-5)	3.284 [0.095]	3.101 [0.099]	3.247 [0.076]	0.187	0.763	0.237
Content of email(s) was accurate (1-5)	3.672 [0.086]	3.493 [0.084]	3.677 [0.075]	0.140	0.960	0.106
I discussed email(s) with colleaugues (1-5)	2.896 [0.109]	2.942 [0.134]	3.022 [0.110]	0.788	0.430	0.645
Email(s) motivated me to register donors (1-5)	3.433 [0.098]	3.087 [0.123]	3.527 [0.086]	0.030	0.474	0.003
Email(s) made me think it is possible to sign up more (1-5)	3.478 [0.089]	3.130 [0.115]	3.667 [0.067]	0.018	0.085	0.000
Email(s) gave me tips to sign up more (1-5)	3.448 [0.096]	3.116 [0.102]	3.247 [0.082]	0.019	0.114	0.312
Email(s) me me think managers value more signups (1-5)	3.448 [0.103]	3.188 [0.104]	3.548 [0.087]	0.078	0.456	0.008
Email(s) increased how important I see signups (1-5)	3.373 [0.108]	3.014 [0.131]	3.398 [0.099]	0.037	0.868	0.019
Email(s) made me feel in competition (1-5)	2.552 [0.111]	2.725 [0.144]	2.989 [0.110]	0.348	0.007	0.139
Email(s) made me feel pressured to ask (1-5)	2.612 [0.118]	3.000 [0.149]	2.559 [0.104]	0.044	0.740	0.013
I would like to receive more email(s) (1-3)	1.410 [0.079]	1.354 [0.074]	1.337 [0.065]	0.605	0.478	0.867
I would like to receive more email(s) (0-1)	0.656 [0.061]	0.708 [0.057]	0.733 [0.048]	0.535	0.320	0.738
I would not like to receive more email(s) (0-1)	0.279 [0.058]	0.231 [0.053]	0.198 [0.043]	0.541	0.254	0.625
I register (less, same, more) donors than avg. of my colleagues (1-3)	2.000 [0.066]	2.015 [0.076]	2.077 [0.063]	0.882	0.415	0.531
I register (less, about same, more) donors than 80% of my colleagues	1.721 [0.085]	1.712 [0.071]	1.747 [0.062]	0.933	0.800	0.710
I register (less, about same, more) donors now than before June 2017	2.339 [0.079]	2.209 [0.084]	2.304 [0.063]	0.265	0.734	0.356
My job increases welfare in the community (1-5)	3.783 [0.130]	3.848 [0.100]	3.789 [0.094]	0.689	0.972	0.669
I am satisfied with my job (1-5)	3.783 [0.147]	3.773 [0.131]	3.989 [0.083]	0.957	0.192	0.146
My job is well suited to my abilities (1-5)	4.033 [0.134]	4.030 [0.122]	4.156 [0.086]	0.987	0.424	0.390

Notes: The table reports the average value of the responses to the questions (or level of agreement with a statement) in the post-intervention survey, by assigned treatment condition of the participating CSRs. For each question, we report the scale or value range of the possible answers. For some of the answers, we report different aggregation of the responses. For example, the statement "Do you remember receiving the email(s)?" had three possible answers: Yes (1), I am not sure (2) and No (3). In addition to reporting the average of these three values, we also aggregated the answers to create an indicator for those who respondent Yes (1) versus those who gave a different answer (0), and a similar one for those who responded No. Standard errors of the means are in brackets. The three rightmost columns report the p-values of the estimated pairwise mean differences between conditions. Details on the survey instrument are available from the authors.

4. Discussion and conclusions

This study evaluated the effect of providing performance feedback to public-sector customer service representatives (CSRs) whose tasks include enrolling residents in the organ donor registry. Theories in economics and behavioral science predict that performance feedback can affect employee performance even for activities that are not directly rewarded with explicit incentives. However, the direction of the effects is theoretically ambiguous, and the evidence context-specific. Our results indicate that while increasing the salience of organ donor registrations with an email that simply reminded CSRs of the importance of asking customers did improve signup performance, providing performance feedback, with or without a reference benchmark, was more effective. Specifically, signup performance for CSRs who received performance feedback in addition to the standard reminder increased by about 25 percent over the signups of those who only received the standard reminder (i.e., information and encouragement but no performance feedback). Furthermore, the addition of performance feedback led to a stronger persistence of the effect than the simple reminder treatment. Finally, and reassuringly, in our context reminders and the two types of feedback produced positive effects among both high-performing and low-performing CSRs.

The post-intervention survey responses indicated that CSRs recalled receiving the intervention e-mails, that they felt that enrolling customers into the donor registry added value to their job, and that asking people to sign up can be effective. However, CSRs in the individual feedback condition reported a higher perception of feeling pressured to ask, whereas those who also received benchmark information perceived less pressure and a stronger motivation effect than colleagues who only observed only individual feedback.

Despite the success of our light-touch, easy-to-implement, and relatively inexpensive interventions to achieve a statistically significant and relatively large 25% increase in daily organ donor registrations compared to otherwise equivalent encouragements and reminders, absolute signup levels remain low. Thus, implementing these types of feedback nudges alone should not be expected to drastically reduce the organ transplant shortage. According to the Global Observatory on Donation and Transplantation (2021), just over 153,000 transplants were performed across all eighty-two member countries in 2019, meeting less than 10% of the estimated need. In Ontario, only about 35% of the population is currently registered, and with significant medical and practical

limitations restricting under what circumstances donations can occur after death, it seems unlikely that incremental registration improvements alone will meet the annual need of some 1,500 people waiting for a transplant in the Province.[5] Redirecting efforts to focus on system-level (rather than individual-level) policy frameworks may be more impactful (Chater and Loewenstein 2022).[6] However, system-level changes can be hard and slow to implement, whereas our study shows that there are marginal improvements that organizations can make which are within the immediate realm of possibilities and do not require large resources and time, to save or at least enhance the quality of life of more residents.

[5] https://www.giftoflife.on.ca/en/publicreporting.htm.

[6] A possible system-wide reform is the introduction of economic rewards to organ donors in an attempt to encourage a massive increase in living kidney donation and possibly donations of kidneys and other organs by deceased individuals (Becker and Elias 2007, Elias et al. 2019, Taylor 2005).

References

Allcott, H. (2011), Social norms and energy conservation. *Journal of Public Economics*, 95(9–10), 1082-1095.

Ball, S., Eckel, C., Grossman, P. J., & Zame, W. (2001). Status in markets. *The Quarterly Journal of Economics*, *116*(1), 161-188.

Bandiera, O., Barankay, I. and Rasul, I., 2013. Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association*, 11(5), 1079-1114.

Bandiera, O., Larcinese, V., & Rasul, I. (2015). Blissful ignorance? A natural experiment on the effect of feedback on students' performance. Labour Economics, 34, 13-25.

Becker, G. S. & Elías, J.J. (2007). Introducing incentives in the market for live and cadaveric organ donations. *Journal of Economic Perspectives*, 21(3), 3–24.

Charness, G., & Grosskopf, B. (2001). Relative payoffs and happiness: an experimental study. *Journal of Economic Behavior & Organization*, *45*(3), 301-328.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The quarterly journal of economics*, *117*(3), 817-869.

Charness, G., Masclet, D. and Villeval, M.C., 2011. Competitive preferences and status as an incentive: Experimental evidence. *CIRANO-Scientific Publications* 2011s-07.

Charness, G., Masclet, D. and Villeval, M.C., 2014. The dark side of competition for status. *Management Science*, 60(1), pp.38-55.

Chater N. & Loewenstein, G. (in press). The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray. *Behavioral and Brain Sciences,* 2022 Sep 5:1-60.

Croson, R., Handy, F., & Shang, J. (2009). Keeping up with the Joneses: The relationship of perceived descriptive social norms, social information, and charitable giving. *Nonprofit Management and Leadership*, *19*(4), 467-489.

DellaVigna, S. and Linos, E. (2022). RCTs to Scale: Comprehensive Evidence from Two Nudge Units. *Econometrica*, 90 (1), 81–116.

Duflo, E., & Saez, E. (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. *Journal of public Economics*, *85*(1), 121-148.

Edelman, B. and Larkin, I. (2015). Social comparisons and deception across workplace hierarchies: Field and experimental evidence. *Organization Science*, *26*(1): 78-98.

Elias, J. J., Lacetera, N., & Macis, M. (2019). Paying for kidneys? A randomized survey and choice experiment. *American Economic Review*, 109(8), 2855-2888.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives don't work to modify behavior. *Journal of Economic Perspectives*, 25(4), 191–210.

Global Observatory on Donation and Transplantation (GODT), (2021). Organ Donation and Transplantation Activities 2019 Report. See http://www.transplant-observatory.org/wp-content/uploads/2021/08/ImpactCOVID19-2.pdf, accessed: October 13, 2021.

Kessler, J. B., & Roth, A. E. (2012). Organ allocation policy and the decision to donate. *American Economic Review*, 102(5), 2018-2047.

Kessler, J. B., & Roth, A. E. (2014). Loopholes undermine donation: An experiment motivated by an organ donation priority loophole in Israel. *Journal of Public Economics*, 114, 19-28.

Larkin, I., Pierce, L. and Gino, F. (2012). The psychological costs of pay-for-performance: Implications for the strategic compensation of employees. *Strategic Management Journal*, 33(10): 1194-1214.

Munshi, K. and Myaux, J., 2006. Social norms and the fertility transition. *Journal of Development Economics*, *80*(1), pp.1-38.

Reiff, J.S., Zhang, J.C., Gallus, J., Dai, H., Pedley, N.M., Vangala, S., Leuchter, R.K., Goshgarian, G., Fox, C.R., Han, M. and Croymans, D.M. (2022). When peer comparison information harms physician well-being. *Proceedings of the National Academy of Sciences*, *119*(29), p.e2121730119.

Robitaille, N., Mazar, N. Tsai, C.I., Haviv, A.M. and Hardy, E. (2021). Increasing Organ Donor Registrations with Behavioral Interventions: A Field Experiment, *Journal of Marketing*. https://journals.sagepub.com/doi/full/10.1177/0022242921990070.

Taylor, J. S. (2005). Stakes and Kidneys: Why Markets in Human Body Parts are Morally Imperative. *Live Questions in Ethics and Moral Philosophy*. Routledge.

Tran, A., & Zeckhauser, R. (2012). Rank as an inherent incentive: Evidence from a field experiment. *Journal of Public Economics*, *96*(9-10), 645-650.

Trillium Gift of Life Network (2017). Trillium Gift of Life Network Annual Report 2016/17. Available at https://www.giftoflife.on.ca/resources/pdf/Trillium_AR_16-17_ENG_access_10_12_2017.pdf. Accessed on 10/5/2021.

Trillium Gift of Life (2021). Public reporting. Available at: https://www.giftoflife.on.ca/en/publicreporting.htm#registration3yrhistory-cal. Accessed on 10/5/2021.

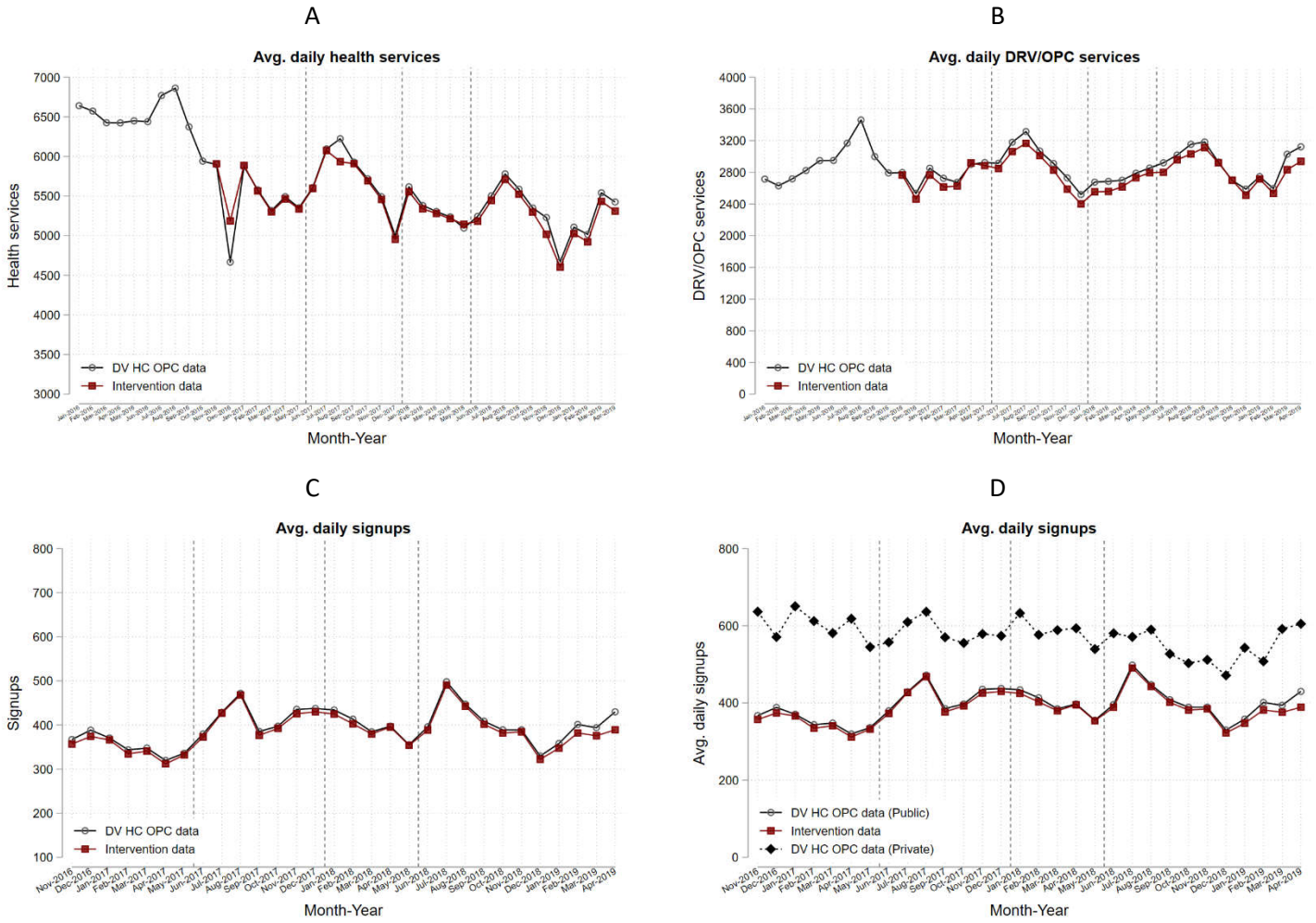
Nudging the Nudger:

Performance Feedback and Organ Donor Registrations

Julian House, Nicola Lacetera, Mario Macis, and Nina Mazar

ONLINE APPENDIX

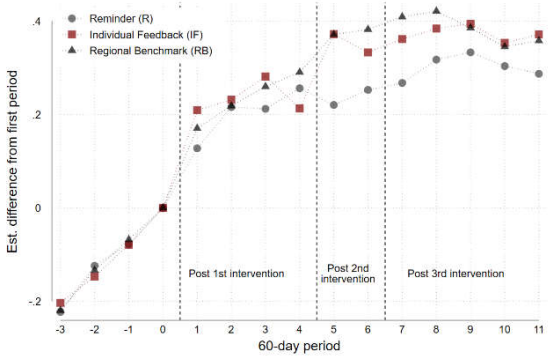
Figure A1: Comparison between ServiceOntario transaction data compiled by Ministries and data used for the interventions



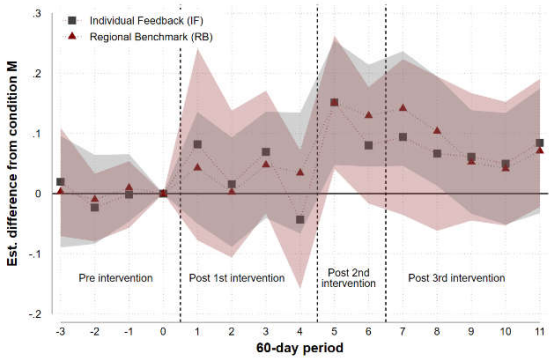
Notes: The graphs report the daily average services at ServiceOntario offices, by month, as recorded by the Ministry of Health (Health-related services: Panels A, C, and D) and the Ministry Of Transportations (Drivers' license and Ontario Photo Card: Panel B), indicated as "DV HC OPC data"; and the same information as collected at ServiceOntario on a daily or service-level basis that was used for our interventions, indicated as "Intervention data."

Figure A2: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups by sub-periods within each intervention – alternative time controls.

A. Within-treatment effects



B. Differences-in-differences from Reminder conditions



Notes: Panel A reports the estimated average changes in daily organ donor signup per CSR in sub-periods of sixty to seventy days within each post-intervention period, compared to the average daily signups in subperiod immediately preceding the first wave of treatment (subperiod 0 on the x-axis), for each single condition. The estimate in correspondence of point 1 (5, 7) on the x-axis indicates the average performance change between approximately two months after the first (second, third) intervention date and two months before. The values at point 5 represent the estimated difference between the first two months after the second intervention. Panel B displays the same estimates as differences of the IF and RB conditions from the R condition. The estimates are from an alternative specification of time controls in the regressions, and are reported in column 2 of Table S6 above. Shaded areas represent 95% confidence intervals, with standard errors clustered both at the CSR level and at the level of office-intervention period. Each intervention consisted of an e-mail whose content differed according to the experimental conditions to which a CSR was randomly assigned. The e-mails were sent on June 20, 2017 (1st intervention), January 29, 2018 (2nd intervention) and June 15, 2018 (3rd intervention).

Table A1: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups, in each intervention period: Regression: estimates for subperiods

Specification	Differences from own condition at subperiod 0	Differences in differences with respect to condition R, per subperiod
	(1)	(2)
R:Pre(-3)	-0.223***	-0.223***
	(0.054)	(0.054)
R:Pre(-2)	-0.124***	-0.124***
	(0.034)	(0.034)
R:Pre(-1)	-0.078***	-0.078***
	(0.025)	(0.025)
R:1st int.(1)	0.128***	0.128***
	(0.041)	(0.041)
R:1st int.(2)	0.216***	0.216***
	(0.039)	(0.039)
R:1st int.(3)	0.212***	0.212***
	(0.044)	(0.044)
R:1st int.(4)	0.256***	0.256***
	(0.058)	(0.058)
R:2nd int.(1)	0.220***	0.220***
	(0.056)	(0.056)
R:2nd int.(2)	0.253***	0.253***
	(0.059)	(0.059)
R:3rd int.(1)	0.268***	0.268***
	(0.072)	(0.072)
R:3rd int.(2)	0.318***	0.318***
	(0.078)	(0.078)
R:3rd int.(3)	0.333***	0.333***
	(0.084)	(0.084)
R:3rd int.(4)	0.304***	0.304***
	(0.089)	(0.089)
R:3rd int.(5)	0.287***	0.287***
	(0.098)	(0.098)
IF:Pre(-3)	-0.203***	0.019
	(0.043)	(0.046)
IF:Pre(-2)	-0.147***	-0.023
	(0.030)	(0.029)
IF:Pre(-1)	-0.079***	-0.002
	(0.023)	(0.028)
IF:1st int.(1)	0.209***	0.082
	(0.076)	(0.081)
IF:1st int.(2)	0.232***	0.016
	(0.055)	(0.062)
IF:1st int.(3)	0.281***	0.069
	(0.054)	(0.052)
IF:1st int.(4)	0.213***	-0.043
	(0.050)	(0.059)
IF:2nd int.(1)	0.372***	0.152***
	(0.064)	(0.056)
IF:2nd int.(2)	0.333***	0.080
	(0.064)	(0.049)

(continues on next page)

(continues from previous page)

IF:3rd int.(1)	0.362***	0.094
	(0.081)	(0.066)
IF:3rd int.(2)	0.384***	0.067
	(0.084)	(0.065)
IF:3rd int.(3)	0.394***	0.061
	(0.086)	(0.054)
IF:3rd int.(4)	0.354***	0.050
	(0.092)	(0.052)
IF:3rd int.(5)	0.372***	0.084
	(0.093)	(0.054)
RB:Pre(-3)	-0.220***	0.003
	(0.044)	(0.047)
RB:Pre(-2)	-0.134***	-0.010
	(0.040)	(0.038)
RB:Pre(-1)	-0.068**	0.010
	(0.027)	(0.028)
RB:1st int.(1)	0.170***	0.043
	(0.036)	(0.047)
RB:1st int.(2)	0.218***	0.002
	(0.043)	(0.046)
RB:1st int.(3)	0.260***	0.048
	(0.051)	(0.045)
RB:1st int.(4)	0.290***	0.034
	(0.056)	(0.051)
RB:2nd int.(1)	0.372***	0.151***
	(0.063)	(0.052)
RB:2nd int.(2)	0.382***	0.130***
	(0.058)	(0.043)
RB:3rd int.(1)	0.409***	0.141***
	(0.071)	(0.048)
RB:3rd int.(2)	0.421***	0.104**
	(0.075)	(0.046)
RB:3rd int.(3)	0.386***	0.053
	(0.081)	(0.044)
RB:3rd int.(4)	0.345***	0.041
	(0.086)	(0.047)
RB:3rd int.(5)	0.358***	0.071
	(0.098)	(0.053)
Constant	-0.123	-0.123
	(0.115)	(0.115)
CSR fixed effects	x	x
Observations	265,475	265,475
R-squared	0.297	0.297

Notes: The table reports estimates from linear regressions where the unit of observation is a CSR on a given day in which that CSR is active. The regressors are binary indicators for experimental conditions and interactions between experimental conditions and subperiods between 60 and 70 days within each intervention. R indicates the reminder e-mail condition, IF the reminder + individual feedback condition, and RB the reminder + individual feedback + regional benchmark condition. The estimated parameter on a given interaction term (e.g., RB: 2nd int.(2)) represents, in column 1, the estimated difference in daily signups between the period that the interaction term identifies (e.g., the second subperiod in the second intervention for condition RB), and the subperiod immediately before the first intervention for that same condition. In Column 2, the parameter estimates that refer to conditions IF and RB represent the estimated differences between the effects of the intervention in condition R in a particular intervention and subperiod and the effect in the same intervention and subperiod in either condition IF or RB. The regressions include variables that measure the number of days a CSR was present in the sample at any given date, its square, the total number of interactions of a CSR on a given day, and year and month fixed effects. The parameter estimates that Figure 3 of the main manuscript display graphically derive from those in this table (column 1 for Panel A, and column 2 for Panel B). Standard errors are clustered both at the CSR level and at the level of office-intervention period. * p<0.1, ** p<0.05, *** p<0.01.

Table A2: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups in each intervention period: Heterogeneous effects

Outcome variable:	Daily signups			
Sample:	Signup performance below average	Signup performance above average	Bottom 80% signup performance	Top 20% signup performance
	(1)	(2)	(3)	(4)
R:1st int.	0.079**	0.129**	0.107***	0.039
	(0.033)	(0.051)	(0.028)	(0.094)
R:2nd int.	0.083**	0.120	0.071**	0.187
	(0.036)	(0.074)	(0.034)	(0.122)
R:3rd int.	0.069	0.157	0.111**	0.130
	(0.049)	(0.097)	(0.052)	(0.131)
IF:1st int.	0.175***	0.245***	0.181***	0.163*
	(0.042)	(0.077)	(0.042)	(0.090)
IF:2nd int.	0.210***	0.188**	0.201***	0.184*
	(0.043)	(0.077)	(0.049)	(0.096)
IF: 3rd int.	0.180***	0.163	0.184***	0.140
	(0.059)	(0.099)	(0.065)	(0.129)
RB:1st int.	0.162***	0.190***	0.167***	0.221***
	(0.039)	(0.049)	(0.037)	(0.064)
RB:2nd int.	0.195***	0.262***	0.203***	0.327***
	(0.043)	(0.072)	(0.038)	(0.101)
RB:3rd int.	0.173***	0.192**	0.187***	0.238*
	(0.050)	(0.087)	(0.052)	(0.123)
Constant	-0.031	-0.372***	-0.156	-0.222*
	(0.095)	(0.128)	(0.100)	(0.126)
CSR fixed effects	x	x	x	x
Observations	168,566	147,113	212,069	103,610
R-squared	0.247	0.323	0.270	0.332

Notes: The table reports estimates from linear regressions where the unit of observation is a CSR on a given day in which that CSR is active. The regressors are binary indicators for experimental conditions and interactions between experimental conditions and interventions. R indicates the reminder e-mail condition, IF the reminder + individual feedback condition, and RB the reminder + individual feedback + regional benchmark condition. The estimated parameter on a given interaction term (e.g., RB: 2nd int.) represents the estimated difference in daily signups between the period that the interaction term identifies and the period before the first intervention for that same condition. The regressions include variables that measure the number of days a CSR was present in the sample at any given date, its square, the total number of interactions of a CSR on a given day, and year and month fixed effects. The parameter estimates displayed in Figure 4 of the main manuscript derive from those in this table. Standard errors are clustered both at the CSR level and at the level of office-intervention period. * p<0.1, ** p<0.05, *** p<0.01.

Table A3: Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups, in each intervention period: Main regression estimates with alternative time controls

Outcome variable:	Daily signups			*I* (Daily signups>0)
	(1)	(2)	(3)	(4)
IF	-0.050		-0.048	
	(0.087)		(0.071)	
RB	-0.114**		-0.109*	
	(0.058)		(0.058)	
R:1st int.	0.107*	0.095**	0.109***	0.031**
	(0.058)	(0.039)	(0.032)	(0.014)
R:2nd int.	0.131**	0.116**	0.141***	0.023
	(0.060)	(0.054)	(0.051)	(0.018)
R:3rd int.	0.254***	0.220***	0.265***	0.049**
	(0.073)	(0.068)	(0.060)	(0.021)
IF:1st int.	0.153	0.125**	0.152***	0.039**
	(0.139)	(0.054)	(0.056)	(0.017)
IF:2nd int.	0.292**	0.236***	0.282***	0.057**
	(0.120)	(0.064)	(0.058)	(0.023)
IF: 3rd int.	0.376***	0.295***	0.366***	0.075***
	(0.114)	(0.076)	(0.069)	(0.025)
RB:1st int.	0.129*	0.124***	0.122***	0.052***
	(0.070)	(0.044)	(0.033)	(0.013)
RB:2nd int.	0.304***	0.255***	0.301***	0.075***
	(0.083)	(0.060)	(0.052)	(0.019)
RB:3rd int.	0.378***	0.302***	0.363***	0.089***
	(0.083)	(0.066)	(0.064)	(0.022)
Constant	0.424***	-0.206*	0.362***	-0.044
	(0.083)	(0.116)	(0.094)	(0.042)
CSR fixed effects		x		x
Office fixed effects			x	
IF:1st int.-R:1st int.	0.046	0.031	0.043	0.008
RB:1st int.-R:1st int.	0.022	0.030	0.014	0.021
RB:1st int.-IF:1st int.	-0.024	-0.001	-0.030	0.013
IF:2nd int.-R:2nd int.	0.161	0.120**	0.141***	0.033*
RB:2nd int.-R:2nd int.	0.172**	0.139***	0.160***	0.052***
RB:2nd int.-IF:2nd int.	0.011	0.019	0.018	0.018
IF:3rd int.-R:3rd int.	0.123	0.076	0.101***	0.026
RB:3rd int.-R:3rd int.	.0.124*	0.082**	0.098***	0.040**
RB:3rd int.-IF:3rd int.	0.002	0.006	-0.003	0.014
Observations	265,475	265,475	265,475	265,475
R-squared	0.145	0.297	0.173	0.253

Notes: The table reports estimates from linear regressions where the unit of observation is a CSR on a given day in which that CSR is active. The regressors are binary indicators for experimental conditions and interactions between experimental conditions and interventions. R indicates the reminder e-mail condition, IF the reminder + individual feedback condition, and RB the reminder + individual feedback + regional benchmark condition. The estimated parameter on a given interaction term (e.g., RB: 2nd int.) represents the estimated difference in daily signups between the period that the interaction term identifies and the pre-intervention period for the same condition. The bottom part of the table reports relevant differences between estimated parameters. The regressions include variables that measure the number of days a CSR was present in the sample at any given date, its square, the total number of interactions of a CSR on a given day, and year and month fixed effects. The parameter estimates in Figure 2 in the main manuscript correspond to those reported in column 2 of this table. Standard errors are clustered both at the CSR level and at the level of office-intervention period. * p<0.1, ** p<0.05, *** p<0.01.

A1. Alternative computation choices

As an additional check of the reliability of our data, we focused on the calculation of the number of unique customers served, per day and in total, by each CSR. The calculation described in the manuscript relied on a particular interpretation of one of the variables the non-health-services datasets included. This column reported, at the CSR-day level, the number of "joint" health and non-health services that a CSR provided to customers on that date (the same customer may receive health and non-health services from the same CSR in a single visit if, for example, they renew their driver's license and register as an organ donor when prompted). In computing the number of daily interactions with unique customers, we assumed that the column indicating the joint transactions in the non-health services dataset indicated how many of the non-health services reported in the other columns also included a health-related service for a given customer and that the health-related dataset did not report these particular, "joint" customer interactions. As such, we did not include transactions reported in the "joint services" column of the non-health related services dataset when calculating the total number of non-health services provided in a day (to be divided by 1.3 to obtain an estimate of the number of unique customers served), to avoid double counting. An alternative assumption would be that these customers requiring joint transactions were also represented, as separate rows, in the health-transactions dataset. If this were the case, a more appropriate way to avoid double (or, in this case, triple) counting would have been to subtract the values in the joint services column from the computation of the total non-health transactions per day rather than simply excluding it from the summation. Although we opted for the former approach and relied on it both to compute the relevant statistics to communicate to the CSRs in the treatment emails and to perform our statistical analyses, we also ran all of these analyses based on the alternative computation of the total transactions and as a consequence of the total unique customer interactions. This different computation may affect the findings in two ways. First, the newly computed total daily interactions, used as control variables in the regressions or as the denominator of the outcome variables when signups are expressed as a percent of total daily customer interactions on the left-hand side, may alter the regression estimates. Second, the past performance we communicated to the CSRs in conditions IF and RB and the median and 80^{th} percentile regional benchmarks were also expressed as per hundred customer interactions, calculated without subtracting the values in the joint service column. As a consequence, it may be the case that some CSRs in condition RB saw in the email that they received that they were (for example) below the reported regional average, but, with respect to the number they saw as representing the regional average, according to the new calculation with the new estimate of total interaction they may have been above. This is a similar issue to the one described in the main text about the partial data transfer in the pre-intervention dataset. In Table A4, we report estimates for a subsample of our analyses corrected with these different calculations and show that no results described in the main text are meaningfully altered. Additional results are available from the authors upon request.

Table A4: Replications with different computation of daily interactions

Outcome variable:	Daily signups			*I* (Daily signups>0)	100*Daily signups/interactions
	(1)	(2)	(3)	(4)	(5)
IF	-0.057		-0.055		
	(0.085)		(0.070)		
RB	-0.118**		-0.114**		
	(0.059)		(0.057)		
R:1st int.	0.210***	0.173***	0.209***	0.056***	1.225***
	(0.057)	(0.040)	(0.033)	(0.014)	(0.285)
R:2nd int.	0.136**	0.129***	0.142***	0.035**	0.832**
	(0.056)	(0.049)	(0.046)	(0.016)	(0.362)
R:3rd int.	0.144**	0.141**	0.154**	0.032	0.569
	(0.073)	(0.065)	(0.059)	(0.020)	(0.478)
IF:1st int.	0.257*	0.203***	0.253***	0.063***	1.422***
	(0.133)	(0.050)	(0.054)	(0.016)	(0.349)
IF:2nd int.	0.298**	0.248***	0.283***	0.068***	1.343***
	(0.117)	(0.056)	(0.051)	(0.020)	(0.403)
IF: 3rd int.	0.270**	0.219***	0.259***	0.059**	0.954*
	(0.117)	(0.072)	(0.068)	(0.025)	(0.529)
RB:1st int.	0.236***	0.205***	0.226***	0.078***	1.596***
	(0.072)	(0.043)	(0.034)	(0.014)	(0.276)
RB:2nd int.	0.311***	0.267***	0.303***	0.087***	2.003***
	(0.078)	(0.051)	(0.043)	(0.016)	(0.368)
RB:3rd int.	0.272***	0.223***	0.254***	0.072***	1.542***
	(0.083)	(0.062)	(0.063)	(0.022)	(0.478)
Constant	0.411***	-0.260**	0.356***	-0.059	0.688
	(0.088)	(0.114)	(0.100)	(0.041)	(0.887)
CSR fixed effects		x		x	x
Office fixed effects			x		
IF:1st int.-R:1st int.	0.047	0.030	0.043*	0.008	0.197
RB:1st int.-R:1st int.	0.026	0.031	0.017**	0.022*	0.371
RB:1st int.-IF:1st int.	-0.020	0.001	-0.026	0.014	0.174
IF:2nd int.-R:2nd int.	0.161	0.119**	0.141***	0.033*	0.510
RB:2nd int.-R:2nd int.	0.175**	0.139***	0.161***	0.052***	1.171***
RB:2nd int.-IF:2nd int.	0.013	0.020	-0.020	0.019	0.660**
IF:3rd int.-R:3rd int.	0.126	0.077	0.105***	0.027	0.385
RB:3rd int.-R:3rd int.	0.128*	0.082**	0.100***	0.040**	0.973**
RB:3rd int.-IF:3rd int.	0.002	0.004	-0.005	0.013	0.587
Observations	265,475	265,475	265,476	265,475	265,475
R-squared	0.143	0.295	0.171	0.249	0.118

Notes: The table reports regression estimates from the same models as those in columns 1 through 3 of Table A1 and columns 2 and 3 of Table A4, with a different computation of daily interactions as described in the appendix section "Alternative computation choices" above. The bottom part of the table reports relevant differences between estimated parameters. The regressions include variables that measure the number of days a CSR was present in the sample at any given date, its square, the total number of interactions of a CSR on a given day, and year and month fixed effects. Standard errors are clustered both at the CSR level and at the level of office-intervention period. * p<0.1, ** p<0.05, *** p<0.01.