Journal of Health Economics 97 (2024) 102914

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhealeco

Nudging the nudger: Performance feedback and organ donor registrations

Julian House [a], Nicola Lacetera [b,e,*], Mario Macis [c], Nina Mazar [d]

[a] Behavioural Insight Unit, Ontario Public Service, Canada
[b] University of Toronto, Canada
[c] Johns Hopkins University, United States
[d] Boston University, United States
[e] University of Bologna, Italy

ARTICLE INFO

Keywords:
Public health
Prosocial behavior
Motivation
Public service
Field experiment
Organ donation

ABSTRACT

In a randomized controlled trial conducted in three waves 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 indicates 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. We also found suggestive evidence that performance feedback with benchmark information was the most motivating and created the least pressure to perform.

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, 2936 organ transplants were performed in 2022. However, by the end of that same year, 3777 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

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, University of Toronto, University of Genova, Ludwig Maximilians Universität in Munich, 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. 32,650) and the Johns Hopkins Homewood IRB (HIRB00005769). It was pre-registered at the AEA RCT Registry (No. 0,001,974).

* Corresponding author.
E-mail address: nicola.lacetera@utoronto.ca (N. Lacetera).

[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.

https://doi.org/10.1016/j.jhealeco.2024.102914
Received 6 November 2023; Received in revised form 27 June 2024; Accepted 29 June 2024
Available online 3 July 2024

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 mass 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, 2017), which through their customer service representatives (CSRs) provide a wide range of services to residents ranging from issuing 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 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 e-mail 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. 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 that 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 content, 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 three-wave experiment).

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 of 2.5 years from November 2016 to April 2019. We used these data both to compute the individual and regional-level statistics included in the intervention e-mails and to estimate the effects of the interventions. The final

[2] Performance evaluations relative to peers may also result in sabotage and other types of unethical behavior (Edelman and Larkin 2015; 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.

dataset for analysis includes 265,475 observations on 693 CSRs, operating in 82 offices on 745 distinct days.

Overall, adding performance feedback resulted in between 0.07 and 0.1 additional signups per CSR per day compared to reminders alone, corresponding to a roughly 11 % increase. 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 characterized both high- and low-performance CSRs. Finally, there was no difference between the effectiveness of individual performance feedback without regional benchmark information and individual performance feedback plus the regional benchmark information.

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 the CSRs had paid attention to our intervention e-mails and whether the feedback interventions had any effects on CSRs' motivations and perceptions. The great majority of CSRs recalled receiving the intervention e-mails, suggesting that they had paid consistent attention to the e-mails 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 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 encourage more sign-ups. This perception was stronger among the CSRs in the RB condition, suggesting that providing performance feedback with benchmark information may enhance CSRs' motivation to sign up more customers to the organ donor registry, reinforce their belief in their agency over this activity, and make the job more meaningful overall. Moreover, CSRs in the IF condition were more likely to report feeling pressure to signup donors. Although some of these differences are not statistically significant, they are consistent in their direction and align with existing recent evidence, such as Reiff et al., 2022. Similar to their study, we conclude that analyses of performance feedback should not only evaluate outcomes but also track measures of motivation and perceptions to ensure a holistic understanding of these interventions' effects.

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 (the most recent available with respect to the beginning of our intervention), ServiceOntario's in-person centers processed about 25 million transactions annually, 20 % of which were handled by their government-owned offices. In these offices across the province, ServiceOntario employs several hundred customer service representatives (CSRs) who 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. Thus there is no external motivation for the CSRs to increase organ donor registrations.

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 RCT 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 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).
- **"Regional 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).

[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. These offices were not available for inclusion in our study.

Fig. 1 shows an example of the RB condition e-mails. These e-mails were sent on June 20, 2017 (1st wave), January 29, 2018 (2nd wave), and June 15, 2018 (3rd wave).

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. This e-mail announcement did not specify the 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.[4] Note that our experiment does not include a pure control; condition R also received email communications encouraging them to sign up more organ donor registry members. Nonetheless, this condition serves as an informative baseline that effectively reflects a business-as-usual scenario. Moreover, it maintains the salience of the organ donation topic constant across CSRs. Following the RCT, we administered a survey to obtain individual characteristics of the CSRs (e.g., age, gender, and tenure), gauge whether they remembered the intervention e-mails from the previous eighteen months and how they perceived these e-mails, and to investigate any effect of the interventions on CSRs motivations and perceptions (Reiff et al. 2022). All CSRs who were still employed at the time of the survey and who were part of the trial 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) and kept that assignment fixed over the three waves. 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; in the pre-intervention period, about 34 % of CSRs did. We chose to assign each of these multisite employees to the office (and thus, condition) where they spent more time in the months immediately preceding the first intervention wave. ServiceOntario staff assisted in determining the condition assignment of these multisite workers.[5] 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). Additionally, the CSRs who joined Service Ontario after randomization are assigned to the condition that corresponds to their office of affiliation.[6]

When we assigned CSRs to experimental conditions, the dataset at our disposal (as of April 30, 2017) included 576 individual CSRs in 80 offices. There were 24 offices and 177 CSRs assigned to Condition R, 27 offices and 198 CSRs to Condition IF, and 28 offices and 190 CSRs to. 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. Constructing signup performance measures and statistics

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).

[4] Administrative data show that over 95% of recipients opened the e-mails almost immediately after reception.

[5] At the time of the randomization, about 66% of the CSR had operated, in the previous six months, in only one office, 25% in two, and 6% in three. Among those who operated in more than one office, the average share of daily observations in their prevailing office was 87%, and the median 94%.

[6] A small number of CSRs entered our sample between the end date of the period for which we calculated individual and benchmark statistics, and the date when Service Ontario sent the treatment emails (for example, with reference to the first intervention, between May 1 and June 20 2017,). As such, these CSRs received their first email in the subsequent intervention (in the example above, on Jan 20 2018). In the previous intervention, therefore, these CSRs received the Reminder email, and afterwards to the condition of their office of affiliation. To comply with our office-level randomization design, however, in the analysis we consider these CSRs as assigned to the condition of their office for the entire period. This makes the nature of our estimated treatment effects akin to an Intent to Treat approach, and, as such, more conservative (although this concerns only a handful of workers).

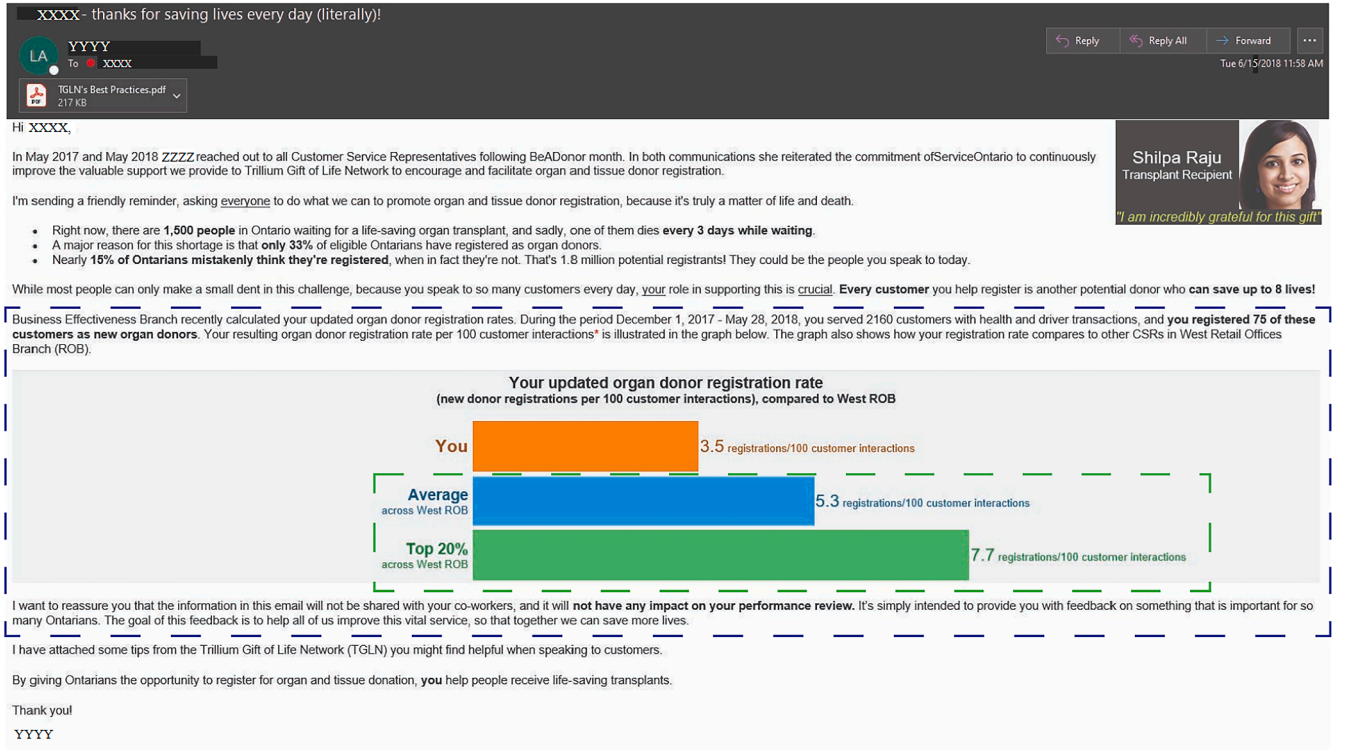

Fig. 1. Sample of RB condition e-mail sent to CSRs.
Notes: The figure reports a snapshot of the RB e-mails that the CSRs received. The content of the e-mails was the same for all CSRs (i.e. across all three conditions) except for the individual feedback and regional benchmark statistics, shown in dashed blue and green boxes, respectively. The names of the e-mail sender and receiver, as well as of the government executive who sent the information e-mails a few weeks before each intervention date, are redacted, and the date on the top right corresponds to the day on which CSRs received the e-mail (here: the third wave-date). All e-mails included an attached PDF "TGLN's Best Practices" – a tip sheet that Trillium Gift of Life Network, Ontario's organ and tissue donation agency, designed to help CSRs become more effective at prompting donations (e.g., by answering potential customer questions with accurate facts). We covered the face and name of the transplant recipient at the top right of the e-mail for confidentiality reasons, but both the picture and name were visible to the CSRs.

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 e-mails. 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 80th percentiles of the total individual signups (per 100 overall interactions) in each of Ontario's four regions.

2.5. 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 Fig. 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 first wave IF and RB e-mails and consisted in the exclusion of some transaction types for some CSRs, which resulted in us overestimating 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 (5206 CSR-day observations, roughly 2 % of the sample) does not affect the estimates meaningfully (Section 3.3 and Table 5 below).

A third check was made possible by the fact that the month-level transaction statistics compiled separately by the Ministries on whose behalf ServiceOntario provides services 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 (Appendix Fig. A1).

Finally, 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, Tables A4 and A5).

2.6. Data description and balance checks

The full dataset includes CSR-day-level information from November 1, 2016 through April 30, 2019, amounting to 295,884 CSR-day observations. The number of organ donor signups per day includes some outliers. Some of these are implausibly large values which are likely reporting errors. Others may indicate that on a given day, a CSR did not engage in any face-to-face activities with customers but worked on "mail-in" (i.e., bulk) activities such as processing received organ donor consent forms in paper form via postal mail. For this reason, in the analyses below, we winsorized the number of signups and the total number of daily customer interactions at the 99.9th percentile, and also excluded CSR-day observation that reported mail-in activities.

After excluding the observations that reported mail-in activities as explained above, the final dataset for analysis included 265,475 observations on 693 CSRs, operating in 82 offices on 745 distinct days. We winsorized both the daily signup and daily unique customer interaction counts at the 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 were 0.65 (range: 0–11) and 13.4 (range: 0–50.7), respectively (Fig. 2). Table 1 reports results from balance tests. In Panel A of Table 1, we report statistics at the office level, i.e. the unit of randomization. There are no statistically significant differences across any of the pre-intervention characteristics (from November 1, 2016 through June 15, 2017). In Panel B, we report balance tests at the CSR-level. Here, *t*-tests for the pre-intervention characteristics show that two out of nine comparisons are significantly different at the five percent level. Overall, we conclude that the randomization

process was effective. Furthermore, the inclusion of individual CSR fixed effects in our preferred regression specifications helps to control for time-invariant unobservable CSR characteristics, enhancing the robustness of our findings.[7]

3. Estimation and results

3.1. Overall effects

We employ a difference-in-difference (DiD) approach, motivated by three main reasons: First, condition R, which we use as baseline, is not a pure control in the standard sense (as we described in Section 2.2). This group also received email communications, making it essential to compare before-after differences in the performance feedback conditions with before-after differences in the R condition. Moreover, the data include a substantial period both before and after the intervention, and our intervention involved three emails distributed over a period of time, rather than a single occurrence. Finally, a DiD approach follows naturally from our experimental setup and, by making full use of the available data, it also increases statistical precision.

Our analysis begins by estimating the effects of the intervention on CSR signup performance by aggregating the three waves into one single post-intervention period. First, we estimate the parameters from a difference-in-difference model, where R, the provision of a reminder e-mail, is the control condition, and the treatment is the combination of the two experimental conditions that included individual performance feedback without and with regional benchmark information (i.e., IF and RB):

$$Y_{cdmy} = \alpha_R + \alpha_{IF,RB} I_{IF,RB} + I_{t>0}\left(\beta_R + \beta_{IF,RB} I_{IF,RB}\right) + \gamma X_{cdmy} + \mu_m + \nu_y + \eta_z + \varepsilon_{cdmy}. \quad (1)$$

In Model 1, Y, the outcome variable, is the number of signups by CSR c on day d, year y and month m. $I_{IF,\ RB}$ is a binary indicator for whether a CSR was in conditions IF or RB (value of one if they were, and zero if they were not). $I_{t>0}$ takes the value of one if an observation is in any day following the date of the first intervention. Therefore, the parameter $\alpha_{IF,RB}$ represents the difference between the average daily signups between CSRs in conditions IF and RB together as compared to those in condition R in the pre-intervention period; β_R is the average difference in signup performance for CSRs in condition R, before and after the first intervention; and $\beta_{IF,RB}$, the difference in difference parameter, represents the effect of the interventions on CSRs in the feedback conditions combined as compared to the R condition, net of any pre-intervention difference between these two groups. The estimate of $\beta_{IF,RB}$ thus measures any additional impact of providing feedback as opposed to just a reminder, over the full post-intervention period (about two years).

The vector X_{cdmd} 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. One limitation of our data is that we do not have administrative information on CSR tenure. We were only able to compute a measure of CSR experience *within our sample*: the number of days of activity since we began 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_z indicates, according to the specification, different additional fixed effects that we add to the model: region, office, or individual CSR level. As described in Section 2.5, 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 (this identifies the effects of interest from CSRs who did work in more than one office).

Our preferred specification, here as well as in the analyses below, includes fixed effects at the level of the individual CSR. This model allows us to account for any time-invariant unobserved heterogeneity among workers. Given the limits in our ability to control for observable characteristics because of data constraints and the presence of some individual-level differences between some of these variables across conditions in the pre-intervention period (Table 1, Panel B), controlling for unobserved individual heterogeneity provides more reliable (and likely more conservative) estimates of the parameters of interest.

We account for within-office and CSR correlation of the remaining variability by clustering standard errors at the office and CSR level. The estimates of the main parameters of interest in Model 1, from linear regressions, are in columns 1 through 4 in Table 2. The remaining four columns provide estimates from a variation of Model 1 where we separately estimate differences from condition R of each of the two feedback conditions (i.e., IF and RB), separately:

$$Y_{cdmy} = \alpha_R + \alpha_{IF} I_{IF} + \alpha_{RB} I_{RB} + I_{t>0}(\beta_R + \beta_{IF} I_{IF} + \beta_{RB} I_{RB}) + \gamma X_{cdmy} + \mu_m + \nu_y + \eta_c + \varepsilon_{cdmy}. \quad (2)$$

The estimates in Table 2 indicate that the effect of adding any type of feedback to the standard reminder e-mails, was, over the three waves combined, between 0.07 and 0.1 additional daily organ donor signups per CSR. Taking the average daily signup for all CSRs in condition R over the pre-intervention period, 0.65, as the baseline, the most conservative estimate (from CSR-level fixed-effect specifications) indicate an additional of 0.07/0.65 = 10.7 % increase in daily signup performance from providing performance feedback (no matter if only individual level or also regional benchmark) over the increase from providing only a reminder e-mail.

[7] Appendix Figure A2 reports, for each office, the share of CSRs with a signup performance above the regional average (Panel A) and in the top 20th percentile (Panel B), as well as the corresponding confidence intervals. There is some variation across offices; however, for the large majority of offices, the shares of interest are not significantly different from the overall mean. Moreover, there are no significant differences across the three experimental conditions.

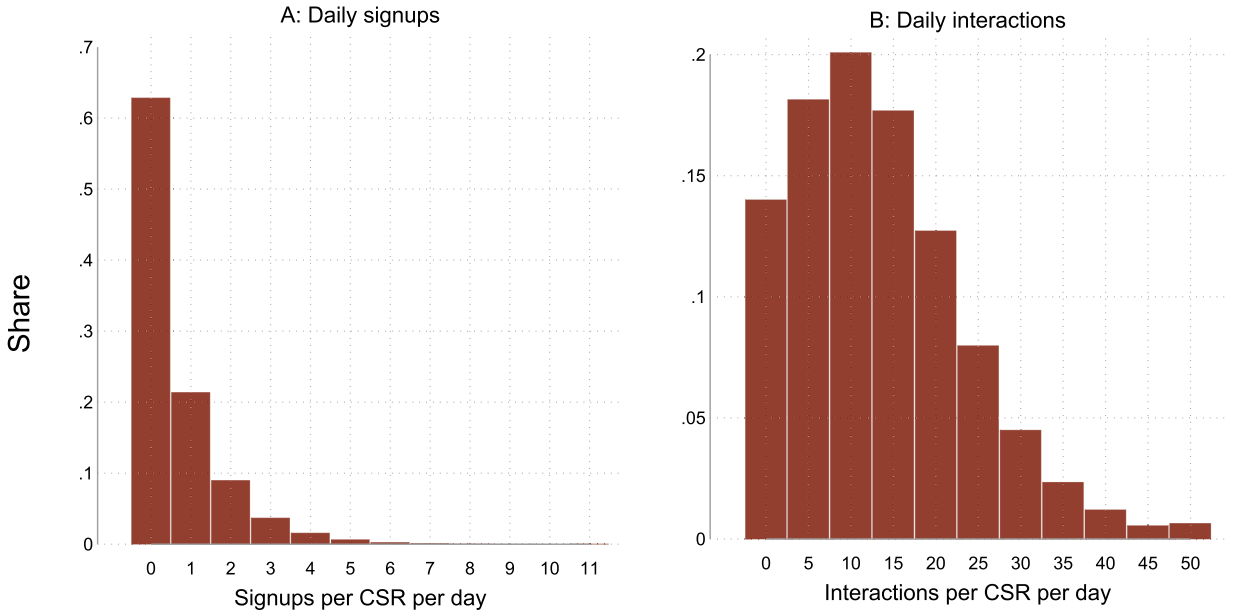

Fig. 2. Distribution of the numbers of daily signups and customer interactions per CSR over the entire period of observation.
Notes: The figures report the empirical distribution of signups (panel A) and customer interactions (panel B) per day and CSR between November 1, 2016 and April 30, 2019. The 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 99.9th percentile and rounded the values of daily interactions to the closest integer.

Table 1
Descriptive statistics and balance tests.

A. Office-level

		Conditions			p-values from pairwise *t*-tests		
		R	IF	RB	IF - R	RB-R	RB-IF
Organ donors signups	Mean	3.584	3.394	2.377	0.884	0.120	0.380
	SE	[0.652]	[1.081]	[0.434]			
	N	24	27	28			
Customer interactions	Mean	74.762	77.370	61.497	0.911	0.478	0.463
	SE	[14.264]	[17.973]	[12.049]			
	N	24	27	28			
Transactions	Mean	96.887	102.517	80.499	0.857	0.504	0.449
	SE	[18.513]	[24.267]	[16.029]			
	N	24	27	28			
Number of CSRs	Mean	5.461	5.444	4.775	0.991	0.545	0.612
	SE	[0.878]	[1.107]	[0.720]			
	N	24	27	28			

B. CSR-level

		Conditions			p-values from pairwise *t*-tests		
		R	IF	RB	IF - R	RB-R	RB-IF
Organ donors signups	Mean	0.640	0.631	0.545	0.874	0.083	0.146
	SE	[0.035]	[0.042]	[0.041]			
	N	187	208	193			
Customer interactions	Mean	13.127	13.981	12.500	0.201	0.317	0.025
	SE	[0.447]	[0.487]	[0.439]			
	N	187	208	193			
Transactions	Mean	16.983	18.523	16.313	0.085	0.420	0.013
	SE	[0.585]	[0.662]	[0.588]			
	N	187	208	193			

Notes: The table reports mean estimates of available CSR characteristics and activities at the level of each office (Panel A) and the individual CSR level (Panel B), limited to the pre-intervention period, together with *p* values from two-tailed test of equality between each pair of treatment conditions. R indicates the reminder e-mail condition, IF the individual feedback condition, and RB the regional benchmark condition.

Table 2
Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups over the three waves combined.

Outcome variable:	Daily signups							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IF+RB	−0.081	−0.073		−0.079				
	(0.063)	(0.055)		(0.058)				
IF					−0.051	−0.042		−0.051
					(0.087)	(0.073)		(0.070)
RB					−0.114*	−0.105*		−0.109*
					(0.058)	(0.058)		(0.059)
Post 1st intervention	0.179***	0.183***	0.161***	0.186***	0.181***	0.185***	0.161***	0.188***
	(0.059)	(0.060)	(0.052)	(0.057)	(0.061)	(0.060)	(0.052)	(0.058)
IF+RB X Post 1st. Intervention	0.106***	0.100***	0.071**	0.089***				
	(0.031)	(0.031)	(0.031)	(0.030)				
IF X Post 1st. Intervention					0.108***	0.102***	0.068	0.093**
					(0.037)	(0.037)	(0.041)	(0.037)
RB X Post 1st. Intervention					0.102**	0.095**	0.074**	0.084**
					(0.039)	(0.038)	(0.030)	(0.035)
Constant	0.388***	0.376***	−0.279**	0.330**	0.387***	0.375***	−0.279**	0.328**
	(0.089)	(0.095)	(0.127)	(0.127)	(0.091)	(0.097)	(0.127)	(0.126)
Fixed effects	Year, Month	Year, Month, Region	Year, Month, CSR	Year, Month, Office	Year, Month	Year, Month, Region	Year, Month, CSR	Year, Month, Office
Clusters	CSR, Office	CSR, Office	CSR, Office	CSR, Office	CSR, Office	CSR, Office	CSR, Office	CSR, Office
Observations	265,475	265,475	265,475	265,475	265,475	265,475	265,475	265,475
R-squared	0.144	0.150	0.297	0.172	0.144	0.150	0.297	0.172

Notes: The table reports estimates from linear regressions where the unit of observation is a CSR on a given day on which that CSR is active. The regressors are binary indicators for experimental conditions and interactions between experimental conditions and a binary indicator for whether an observation is for a period following the first intervention date (June 20, 2017). R indicates the reminder e-mail condition, IF the individual feedback condition, and RB the regional benchmark 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 customer interactions of a CSR in a given day. Standard errors are clustered both at the CSR and office level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Columns 5 though 8 show that the effects were not different between the IF and the RB treatments.

3.2. Effects over time and dynamics

The extended collaboration with Service Ontario made it possible to repeat the interventions over multiple waves and to have data over a long period (2.5 years) to study both short-term and longer-term impacts of the interventions. We designed our experiment to evaluate both the differential effect between each condition in each given wave, and the differences in signup rates within a condition, between the different waves.

Fig. 3 reports all these effects in one graph; the values that the various markers report are the "β" coefficient estimates from the following regression equation:

$$Y_{cdmy} = \alpha_0 + I_{IF}\alpha_{IF} + I_{RB}\alpha_{RB} + I_{t=1}(\beta_{R1} + I_{IF}\beta_{IF1} + I_{RB}\beta_{RB1}) + I_{t=2}(\beta_{R2} + I_{IF}\beta_{IF2} + I_{RB}\beta_{RB2}) + I_{t=3}(\beta_{R3} + I_{IF}\beta_{IF3} + I_{RB}\beta_{RB3}) + \gamma X_{cdmy} + \mu_m + \nu_y + \eta_c + \varepsilon_{cdmy}. \quad (3)$$

$I_{t=1}$, $I_{t=2}$ and $I_{t=3}$ take the 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 $\widehat{\beta}_{IF2}$ 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, $\widehat{\beta}_{RB3} - \widehat{\beta}_{RB2}$ 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. $\widehat{\beta}_{RB1} - \widehat{\beta}_{IF1}$, for instance, is the estimate of 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. Consistently with the multi-wave intervention design, we cluster the

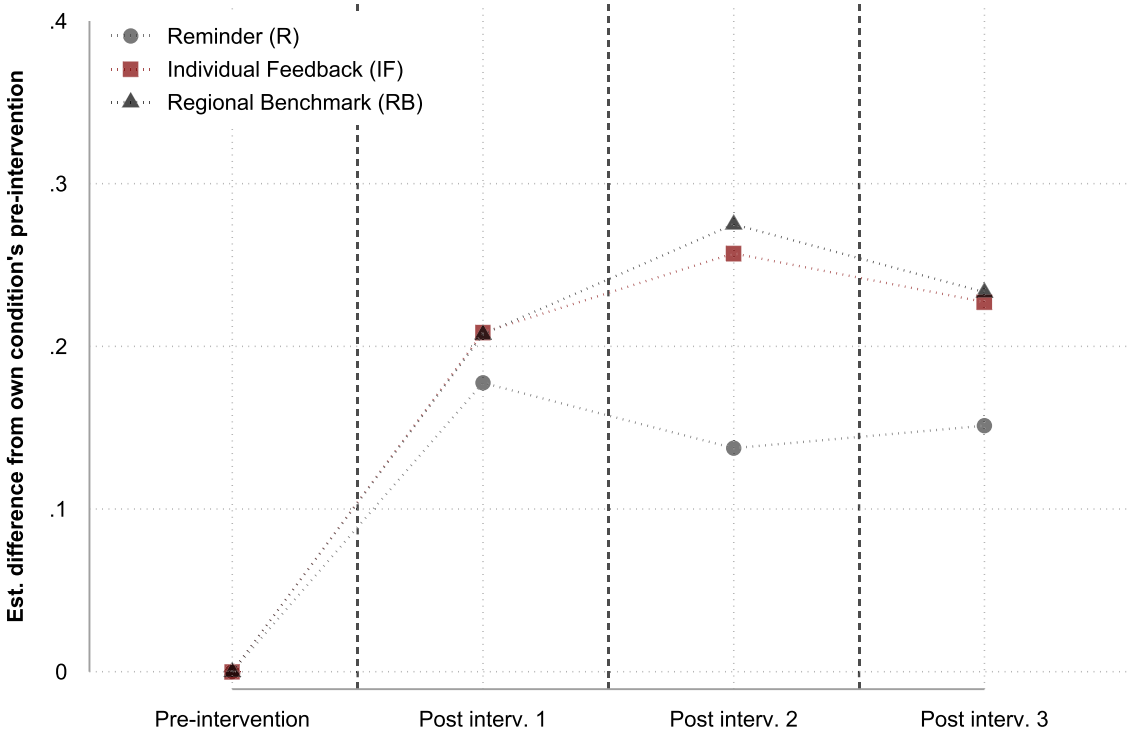

Fig. 3. Effects of Reminder, Individual Feedback, and Regional Benchmark communications on the number of daily signups in each intervention period.

Notes: The figure reports the estimated average changes in daily organ donor signup per CSR after each of the three intervention periods, 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 3. Each intervention consisted of an e-mail whose content differed according to the experimental conditions to which a CSR was randomly assigned to. The e-mails were sent on June 20, 2017, January 29, 2018, and June 15, 2018.

standard errors at the office-intervention and CSR level. The estimates in Fig. 3 are from column 2 of Table 3 (Panel A), where we report the results from a CSR fixed effect specification of Model 3.

The other two columns in Panel A of Table 3 show that alternative specifications (with only year and month fixed effects, or with the addition of office-level instead of CSR level fixed effects) convey similar estimates (slightly larger than the CSR fixed-effect estimates). Panel B of the table displays the p-values from tests of equality between different pairs of coefficients.

Although no significant differences emerge between conditions in the first wave, 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 0.65 daily signups per CSR as the reference, this represents an increase of over 20 % compared to the standard reminder e-mail that provided 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 regional benchmark performance feedback does not lead to different outcomes than just providing individual performance feedback.

The fact that the differences between the reminder and the feedback conditions emerge in the second wave, and not in the first, reinforces our interpretation that these effects are driven specifically by feedback information. If the main driver of the behavioral change was just the increased salience of the organ donation signup activity initiated by the performance feedback e-mails, and not its actual informational content/value, one should have expected larger differences from the Reminder condition to emerge in the first wave when, arguably, the salience or novelty effect should have been stronger.

In Fig. 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. With this additional analysis, we explore the dynamics of the reactions to the conditions within each period; in particular, we assess whether the effects occurred smoothly over time, took some time before materializing, or concentrated right after the reception of the e-mails. The choice of the length of each period, which we did not pre-register, was in part determined by the overall time between subsequent interventions. In particular, the period between the second and third wave turned out to be shorter than originally planned (equivalently, the period between the first two waves was longer); this left us with limited room on how to further segment the time between interventions, without making each segment too short. A range of about two months was a compromise between these different needs.

The response to the e-mails concentrated mainly in the few weeks immediately after the first wave 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 condition

Table 3
Effects of reminder, individual feedback, and regional benchmark communications on the number of daily signups in each intervention period: By intervention.

A: 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)
Fixed effects	Year, Month	Year, Month, CSR	Year, Month, Office
Clusters	CSR, Office-intervention	CSR, Office-intervention	CSR, Office-intervention
Observations	265,475	265,475	265,475
R-squared	0.145	0.297	0.172

B: Difference-in-Dfference estimates			
	(1)	(2)	(3)
IF-R 1st	0.047	0.031	0.044*
RB-R 1st	0.023	0.030	0.014*
RB-IF 1st	−0.024	−0.001	−0.030
IF-R 2nd	0.162	0.120**	0.142***
RB-R 2nd	0.172**	0.138***	0.159***
RB-IF 2nd	0.010	0.018	0.017
IF-R 3rd	0.124	0.076	0.103***
RB-R 3rd	0.126*	0.082**	0.099***
RB-IF 3rd	0.001	0.006	−0.004
R: 1st-0	0.217***	0.178***	0.215***
R: 2nd-1st	−0.070*	−0.040	−0.063***
R: 3rd-2nd	0.009	0.014	0.014
IF: 1st-0	0.263*	0.209***	0.260***
IF: 2nd-1st	0.045	0.049	0.034*
IF: 3rd-2nd	−0.028	−0.030	−0.025
RB: 1st-0	0.239***	0.207***	0.229***
RB: 2nd-1st	0.079	0.068**	0.082***
RB: 3rd-2nd	−0.037	−0.042	−0.046*

Notes: Panel A reports estimates from linear regressions where the unit of observation is a CSR on a given day on which that CSR is active. The regressors are binary indicators for experimental conditions and interactions between experimental conditions and binary indicators for each intervention wave. R indicates the reminder e-mail condition, IF the individual feedback condition, and RB the regional benchmark condition. The estimated parameter for a given interaction term (e.g., RB: 2nd int.) represents the estimated difference in daily signups for CSR in a given condition, between the period that the interaction term identifies, and the pre-intervention period for the same condition. Panel B shows p-values from test of differences between 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 customer interactions of a CSR in a given day, and year and month fixed effects. Each intervention consisted of an e-mail whose content differed according to the experimental condition to which a CSR was randomly assigned. The e-mails were sent on June 20, 2017, January 29, 2018, and June 15, 2018. The parameter estimates in Fig. 3 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

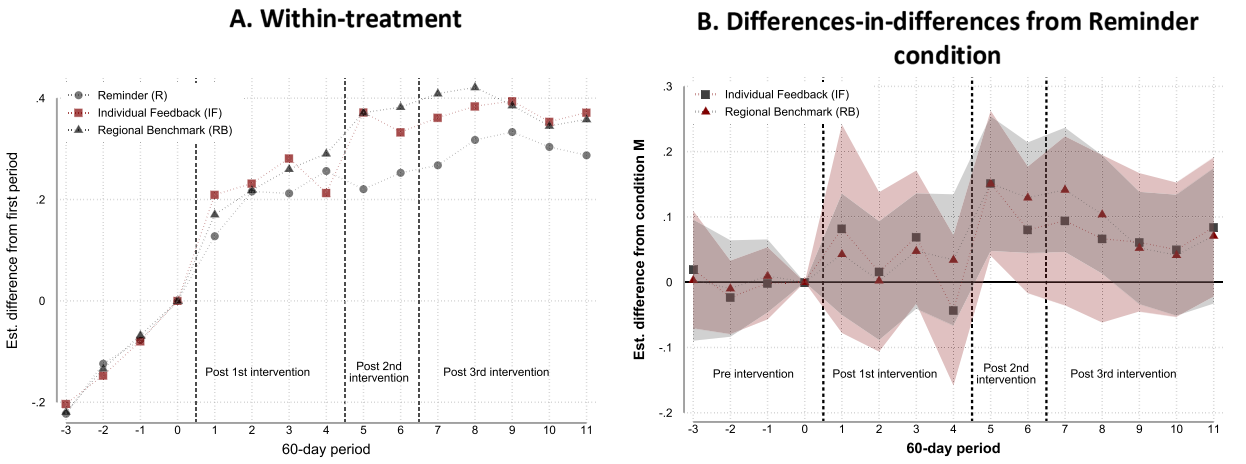

Fig. 4. Effects of Reminder, Individual Feedback, and Regional Benchmark communication on the number of daily signups by sub-periods.
Notes: Panel A reports the estimated average changes in daily organ donor signup per CSR in sub-periods of sixty to seventy days within each wave, compared to the average daily signups in the subperiod immediately preceding the first wave of treatment (subperiod 0 on the x-axis), for each 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 first 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 Table A1 in the Appendix. Shaded areas in Panel B 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 condition to which a CSR was randomly assigned. The e-mails were sent on June 20, 2017, January 29, 2018, and June 15, 2018.

R. To better gauge the differential impact of the two feedback conditions compared to the simple e-mail reminder, Panel B of Fig. 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 a reminder e-mail.

There is a small attenuation of the effect after the third email. This suggests that CSRs may have grown accustomed to the repeated messages, resulting in a reduced impact over time. Additionally, ceiling effects are also a possible explanation, where further improvements are hard to achieve after a certain level of response has been obtained by the initial emails. Both of these factors could contribute to the observed slight decline in effectiveness with successive communications.

3.3. Additional analyses and robustness checks

Tables 4 and 5 report the estimates from additional analyses to investigate other potential effects of the interventions and assess the robustness of the results from our main specification.

The parameter estimates in Table 4 are from the model described in Eq. (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 3 above. Column 3 of Table 4 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 feedback interventions 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.

In columns 1 through 5 of Table 5, 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 2.5 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 3.

Table 4
Effects of Reminder, Individual Feedback, and Regional Benchmark emails on the number of daily signups after each intervention wave: Alternative outcomes.

A: Regression estimates

Outcome variable:	100*Daily signups/interactions		Signups>0	Daily transactions
	(1)	(2)	(3)	(4)
R:1st int.	1.183***	1.229***	0.057***	1.130*
	(0.270)	(0.278)	(0.014)	(0.629)
R:2nd int.	0.843**	0.884**	0.038**	1.164
	(0.348)	(0.367)	(0.016)	(0.873)
R:3rd int.	0.631	0.630	0.035*	0.004
	(0.455)	(0.472)	(0.020)	(1.073)
IF:1st int.	1.415***	1.444***	0.065***	0.641
	(0.338)	(0.358)	(0.016)	(1.001)
IF:2nd int.	1.381***	1.378***	0.072***	−0.207
	(0.384)	(0.405)	(0.020)	(1.199)
IF: 3rd int.	1.035**	1.037*	0.062**	−0.118
	(0.506)	(0.533)	(0.024)	(1.325)
RB:1st int.	1.541***	1.584***	0.078***	1.142*
	(0.259)	(0.272)	(0.014)	(0.633)
RB:2nd int.	1.948***	1.989***	0.090***	1.125
	(0.347)	(0.361)	(0.017)	(0.805)
RB:3rd int.	1.595***	1.631***	0.076***	0.963
	(0.453)	(0.475)	(0.021)	(1.116)
Constant	0.441	1.595*	−0.074*	31.146***
	(0.844)	(0.845)	(0.041)	(1.579)
Fixed effects	Year, Month, CSR	Year, Month, CSR	Year, Month, CSR	Year, Month, CSR
Clusters	CSR, Office-intervention	CSR, Office-intervention	CSR, Office-intervention	CSR, Office-intervention
Observations	265,475	265,475	2,65,475	265,475
R-squared	0.120	0.118	0.252	0.357

B: Difference-in-Dfference estimates

	(1)	(2)	(3)	(4)
IF-R 1st	0.231	0.215	0.008	−0.489
RB-R 1st	0.357	0.356	0.021	0.012
RB-IF 1st	0.126	0.141	0.013	0.501
IF-R 2nd	0.538*	0.494	0.033*	−1.372
RB-R 2nd	1.106***	1.105***	0.051***	−0.039
RB-IF 2nd	0.567*	0.611*	0.018	1.333
IF-R 3rd	0.404	0.406	0.027	−0.123
RB-R 3rd	0.965***	1.000***	0.040**	0.959
RB-IF 3rd	0.560*	0.594*	0.014	1.082
R: 1st-0	1.183***	1.229***	0.057***	1.130*
R: 2nd-1st	−0.341	−0.345	−0.019*	0.034
R: 3rd-2nd	−0.212	−0.253	−0.003	−1.160*
IF: 1st-0	1.415***	1.444***	0.065***	0.641
IF: 2nd-1st	−0.034	−0.066	0.007	−0.849
IF: 3rd-2nd	−0.346	−0.341	−0.010	0.089
RB: 1st-0	1.541***	1.584***	0.078***	1.142*
RB: 2nd-1st	0.408*	0.405*	0.011	−0.017
RB: 3rd-2nd	−0.353	−0.358	−0.014	−0.162

Notes: Panel A reports estimates from linear regressions where the unit of observation is a CSR on a given day on which that CSR is active. The regressors are binary indicators for experimental conditions and interactions between experimental conditions and binary indicators for each intervention wave. R indicates the reminder e-mail condition, IF the individual feedback condition, and RB the regional benchmark condition. The estimated parameter for a given interaction term (e.g., RB: 2nd int.) represents the estimated difference in daily signups for CSR in a given condition, between the period that the interaction term identifies, and the pre-intervention period for the same condition. Panel B shows p-values from tests of differences between parameters. The regressions include variables that measure the number of days a CSR was present in the sample at any given date, its square, and 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). Each intervention consisted of an e-mail whose content differed according to the experimental condition to which a CSR was randomly assigned. The e-mails were sent on June 20, 2017, January 29, 2018, and June 15, 2018. Standard errors are clustered both at the CSR level and at the level of office-intervention period. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

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 signup activity (Appendix Fig. A1), this does not seem to be the case, but we take this as a further robustness check.

Table 5
Effects of reminder, individual feedback, and regional benchmark emails on the number of daily signups after each intervention wave: Robustness to 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
Model specification:	Linear	Linear	Linear	Linear	Linear	Poisson
	(1)	(2)	(3)	(4)	(5)	(6)
R:1st int.	0.133***	0.156***	0.180***	0.187***	0.198***	0.165***
	(0.034)	(0.054)	(0.040)	(0.042)	(0.055)	(0.033)
R:2nd int.	0.125***	0.110*	0.140***	0.140***	0.218***	0.129***
	(0.044)	(0.057)	(0.048)	(0.050)	(0.065)	(0.041)
R:3rd int.	0.133**	0.121	0.158**	0.148**	0.219**	0.144***
	(0.060)	(0.075)	(0.064)	(0.068)	(0.088)	(0.051)
IF:1st int.	0.221***	0.202***	0.210***	0.225***	0.247***	0.187***
	(0.053)	(0.054)	(0.048)	(0.047)	(0.059)	(0.040)
IF:2nd int.	0.235***	0.221***	0.265***	0.306***	0.335***	0.245***
	(0.054)	(0.064)	(0.055)	(0.062)	(0.072)	(0.049)
IF: 3rd int.	0.206***	0.247***	0.235***	0.249***	0.296***	0.226***
	(0.068)	(0.079)	(0.069)	(0.073)	(0.096)	(0.063)
RB:1st int.	0.204***	0.200***	0.206***	0.210***	0.127**	0.214***
	(0.044)	(0.054)	(0.043)	(0.060)	(0.053)	(0.036)
RB:2nd int.	0.253***	0.259***	0.274***	0.298***	0.266***	0.275***
	(0.053)	(0.061)	(0.051)	(0.064)	(0.062)	(0.043)
RB:3rd int.	0.216***	0.209***	0.235***	0.253***	0.313***	0.252***
	(0.060)	(0.070)	(0.061)	(0.082)	(0.081)	(0.055)
No interactions in a given day			−0.190***			
			(0.031)			
Constant	−0.295**	−0.404**	−0.214*	−0.219*	−0.342**	
	(0.116)	(0.163)	(0.112)	(0.122)	(0.157)	
Fixed effects	Year, Month, CSR	Year, Month, CSR	Year, Month, CSR	Year, Month, CSR	Year, Month, CSR	Year, Month, CSR
Clusters	CSR, Office-intervention	CSR, Office-intervention	CSR, Office-intervention	CSR, Office-intervention	CSR, Office-intervention	CSR, Office-intervention
Observations	243,190	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 linear 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 binary indicators for each intervention wave. R indicates the reminder e-mail condition, IF the individual feedback condition, and RB the regional benchmark condition. The estimated parameter for a given interaction term (e.g., RB: 2nd int.) represents the estimated difference in daily signups for CSR in a given condition, between the period that the interaction term identifies, and the pre-intervention period for the same condition. Panel B shows p-values from test of differences between parameters. The regressions include variables that measure the number of days a CSR was present in the sample at any given date, its square, and 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). Each intervention consisted of an e-mail whose content differed according to the experimental condition to which a CSR was randomly assigned. The e-mails were sent on June 20, 2017, January 29, 2018, and June 15, 2018. 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 shows estimates from the main analyses discussed above, but from a model with month-year indicators, combined, as regressors. Appendix Fig. A3 replicates Fig. 3, again with estimates from this alternative model specification. On the one hand, the estimates imply a slightly more contained increase in signups after the first intervention wave, and a more continuous increase, with less tapering, in later waves. 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.

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 feedback information may 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

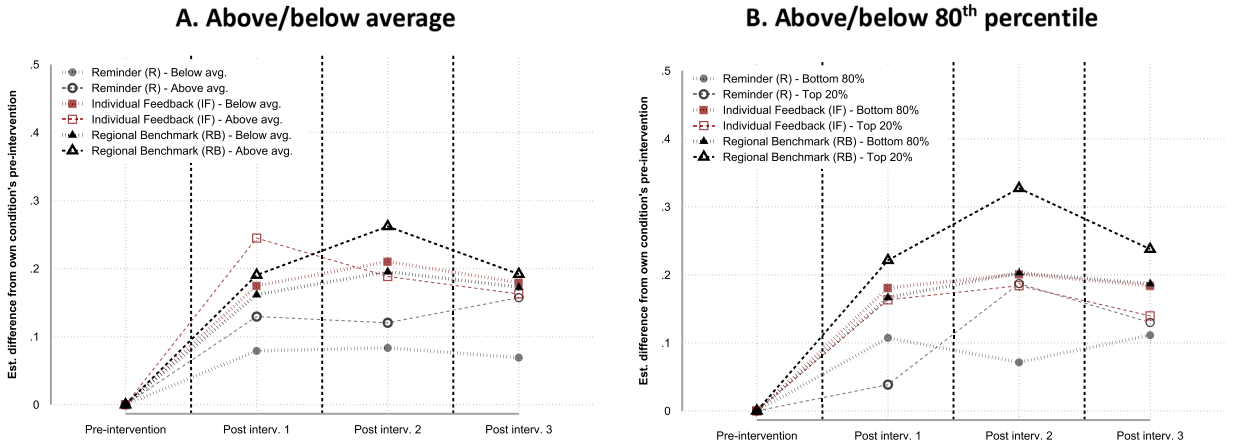

Fig. 5. Effects of Reminder, Individual Feedback, and Regional Benchmark emails on the number of daily signups, by signup performance.
Notes: The figures report the estimated average changes in daily organ donor signup per CSR after each of the three intervention waves, compared to the average daily signups in the pre-intervention periods for each single condition. In panel A, the estimates per condition are separated between CSRs whose performance in the period immediately before a given wave was above the regional average (Appendix Table A2, column 2), and those with a performance below average (Table A2, column 1). In Panel B, the separation is between the CSRs with pre-intervention performance in the top 20 % (Table A2, column 4), and those with performance in the bottom 80 % in a given region (Table A2, column 3). Each intervention consisted of an e-mail whose content differed according to the experimental condition to which a CSR was randomly assigned. The e-mails were sent on June 20, 2017, January 29, 2018, and June 15, 2018.

Fig. 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.[8]

3.5. *Post-intervention survey*

Table 6 reports the average values of the responses to the post-intervention survey questions (or level of agreement with a statement), by assigned condition of the participating CSRs. For each question, we report the scale or value range of the possible answers. To test for systematic differences in responses between conditions, we calculate both the row p-values from pairwise *t*-tests, and p-values corrected for multiple hypotheses testing, because of the large number of outcomes that we consider.[9]

In the survey, 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 e-mails, regardless of their content, and that the signup differences that we observed are due to the specific content of the messages.[10]

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.

Once we adjust the p-values for multiple-hypothesis testing, the differences across experimental conditions are mostly not statistically significant. However, the statistically significant differences consistently suggest that providing performance feedback with benchmark information (RB) may enhance CSRs' motivation to sign up more customers to the organ donor registry, reinforce their belief in their agency over this activity, and make the job more meaningful overall. In contrast, CSRs in the individual feedback condition without benchmarks (IF) reported the lowest agreement on these topics among the three conditions. Moreover, CSRs in the IF condition were more likely to report feeling pressure to signup donors, although the statistical significance of these differences did not survive the multiple hypothesis correction. Overall, the evidence suggests that providing performance feedback with benchmarks might be preferable both to encourage CSRs and to increase their overall job satisfaction. Providing a frame of reference through the

[8] Unfortunately, our data does not include administrative information on CSR tenure, which prevents us from exploring treatment heterogeneity by CSR experience.

[9] In addition to our questions of interest, the survey included questions that we did not use for our study but that were of interest to Service Ontario management and which we include as a service as part of our collaboration.

[10] About 34% of the respondents (77 out of 217) responded "agree" or "strongly agree" to the statement "I discussed the e-mail(s) with my colleagues". The percentage is higher in the conditions with performance feedback (about 38% in both of them) than in the "reminder" condition (23%), although the numbers are relatively small. The fact that there was some conversation (at least as we can infer from the survey responses) suggests that having the same experimental condition within an office was an appropriate design choice to avoid confusion and spillover effects.

Table 6
Responses to the post-intervention survey.

	Mean per conditions [std. error]			p-values from pairwise *t*-tests (mult. Hp. testing correction)		
	R	IF	RB	IF - R	RB-R	RB-IF
Customers more likely to signup if asked (1–5)	3.586	3.452	3.686	0.378	0.481	0.100
	[0.107]	[0.107]	[0.092]	(1.000)	(1.000)	(0.969)
Signing up donors makes job more meaningful (1–5)	3.800	3.479	4.020	0.049	0.117	0.000
	[0.113]	[0.116]	[0.086]	(0.864)	(0.990)	(0.000)
Most want to be donors, but didn't sign up (1–5)	3.029	3.027	3.127	0.994	0.524	0.508
	[0.110]	[0.103]	[0.104]	(0.991)	(1.000)	(1.000)
By remembering to ask, I'd sign up more people (1–5)	4.086	3.973	4.402	0.485	0.009	0.002
	[0.101]	[0.125]	[0.071]	(1.000)	(0.390)	(0.117)
By explaning registry, I'd sign up more people (1–5)	3.943	3.795	4.029	0.337	0.532	0.098
	[0.106]	[0.112]	[0.089]	(1.000)	(1.000)	(0.971)
I don't remember receiving the email(s) (0–1)	0.029	0.042	0.059	0.675	0.358	0.617
	[0.020]	[0.024]	[0.023]	(1.000)	(1.000)	(1.000)
Email(s) included new information (1–5)	3.284	3.101	3.247	0.187	0.763	0.237
	[0.095]	[0.099]	[0.076]	(0.998)	(1.000)	(1.000)
Content of email(s) was accurate (1–5)	3.672	3.493	3.677	0.140	0.960	0.106
	[0.086]	[0.084]	[0.075]	(0.991)	(1.000)	(0.970)
I discussed email(s) with colleaugues (1–5)	2.896	2.942	3.022	0.788	0.430	0.645
	[0.109]	[0.134]	[0.110]	(1.000)	(1.000)	(1.000)
Email(s) motivated me to register donors (1–5)	3.433	3.087	3.527	0.030	0.474	0.003
	[0.098]	[0.123]	[0.086]	(0.738)	(1.000)	(0.204)
Email(s) made me think it is possible to sign up more (1–5)	3.478	3.130	3.667	0.018	0.085	0.000
	[0.089]	[0.115]	[0.067]	(0.579)	(0.962)	(0.000)
Email(s) gave me tips to sign up more (1–5)	3.448	3.116	3.247	0.019	0.114	0.312
	[0.096]	[0.102]	[0.082]	(0.498)	(0.983)	(1.000)
Email(s) me think managers value more signups (1–5)	3.448	3.188	3.548	0.078	0.456	0.008
	[0.103]	[0.104]	[0.087]	(0.961)	(1.000)	(0.395)
Email(s) increased how important I see signups (1–5)	3.373	3.014	3.398	0.037	0.868	0.019
	[0.108]	[0.131]	[0.099]	(0.796)	(1.000)	(0.587)
Email(s) made me feel in competition (1–5)	2.552	2.725	2.989	0.348	0.007	0.139
	[0.111]	[0.144]	[0.110]	(1.000)	(0.242)	(0.992)
Email(s) made me feel pressured to ask (1–5)	2.612	3.000	2.559	0.044	0.740	0.013
	[0.118]	[0.149]	[0.104]	(0.857)	(1.000)	(0.537)
I would like to receive more email(s) (0–1)	0.656	0.708	0.733	0.535	0.320	0.738
	[0.061]	[0.057]	[0.048]	(1.000)	(1.000)	(1.000)
I register (less, same, more) donors than avg. of my colleagues (1–3)	2.000	2.015	2.077	0.882	0.415	0.531
	[0.066]	[0.076]	[0.063]	(1.000)	(1.000)	(1.000)
I register (less, about same, more) donors than 80 % of my colleagues (1–3)	1.721	1.712	1.747	0.933	0.800	0.710
	[0.085]	[0.071]	[0.062]	(1.000)	(1.000)	(1.000)
I register (less, about same, more) donors now than before June 2017 (1–3)	2.339	2.209	2.304	0.265	0.734	0.356
	[0.079]	[0.084]	[0.063]	(1.000)	(1.000)	(1.000)
My job increases welfare in the community (1–5)	3.783	3.848	3.789	0.689	0.972	0.669
	[0.130]	[0.100]	[0.094]	(1.000)	(1.000)	(1.000)
I am satisfied with my job (1–5)	3.783	3.773	3.989	0.957	0.192	0.146
	[0.147]	[0.131]	[0.083]	(1.000)	(1.000)	(0.996)
My job is well suited to my abilities (1–5)	4.033	4.030	4.156	0.987	0.424	0.390
	[0.134]	[0.122]	[0.086]	(1.000)	(1.000)	(1.000)

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 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 aggregations of the responses. For example, the statement "I remember receiving the e-mail(s)?" had three possible answers: Yes (1), I am not sure (2) and No (3). We also aggregated the answers to create an indicator for those who responded No (1) versus those who gave a different answer (0). Standard errors of the means are in brackets. The three rightmost columns report the p-values of the estimated pairwise mean differences between conditions. We computed the multiple-hypothesis correction with the Stata command *mhtreg* (List et al. 2019; Steinmayr, 2020). Details on the survey instrument are in the Appendix.

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 statistically significant difference between respondents in the three conditions in their reported interest in receiving similar messages in the future (the share of CSRs indicating interest in receiving similar messages in the future was relatively high, 64 – 73 %).

4. Discussion and conclusions

This study evaluated the effect of providing performance feedback to public-sector customer service representatives (CSRs) who are encouraged to enroll 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 e-mail that simply reminded CSRs of the importance of asking customers did improve signup performance, adding performance feedback, with or without a reference benchmark, was even more effective. Specifically, signup performance for CSRs who received individual performance feedback in addition to the standard reminder increased by about 10 percent over the signups of those who only received the standard reminder (i.e., information and encouragement but no performance feedback), and to about 20 percent after repeated interventions. 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 fact that performance feedback, both with and without benchmark information, had similar effects could potentially be explained by the markedly pro-social nature of organ donation and the absence of monetary rewards, which might enhance the intrinsic motivation of CSRs, thus influencing the effects of both types of feedback for both high and low performers. Additionally, both types of feedback provided CSRs with new information about their performance, which, too, could explain the similar effectiveness of the two interventions. The simple reminder was insufficient possibly because it lacked the personalized and comparative insights that could more effectively engage and motivate CSRs, emphasizing the importance of targeted feedback in achieving behavioral change. Another possible explanation is that providing feedback, with or without benchmarks, creates a sense that one's behavior is being observed, which has been shown to promote prosocial behavior in other contexts (see for example Bateson et al., 2013). Although our study is limited in its ability to identify precise mechanisms, to our knowledge, our findings are novel. Other studies of performance feedback in medical contexts did not include feedback treatments with and without benchmarks (e.g., Gauri et al., 2021). For instance, Meeker et al. (2016) focused on comparing antibiotic prescribing rates among physicians to those of top performers, demonstrating the influence of peer benchmarks on reducing inappropriate prescriptions. Similarly, Reiff et al. (2022) studied how lists of top performers affected primary care physicians' recommendations for preventive care measures, yet did not test the impact of individual feedback without comparative benchmarks. These studies often found varying impacts. Further evidence is needed, and understanding the mechanisms fully requires considering evidence from multiple contexts.

The post-intervention survey responses suggest that feedback without benchmark information might have increased the perception of pressure to perform among CSRs. Specifically, CSRs who received individual feedback reported a higher perception of feeling pressured to ask customers to sign up as organ donors compared to those who also received benchmark information. While this finding was not robust to multiple hypothesis testing, it aligns with evidence from Reiff et al. (2022) that certain types of performance feedback can lead to increased pressure.

Our light-touch, easy-to-implement, and relatively inexpensive interventions achieved a statistically significant and relatively large 25 % increase in actual daily organ donor registrations compared to an otherwise equivalent encouragement and reminder e-mail. However, despite the success, 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 1500 people waiting for a transplant in the Province.[11] Redirecting efforts to focus on system-level (rather than individual-level) policy frameworks may be more impactful (Chater and Loewenstein 2022).[12] 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 a few more residents.

Author statement

The study received approval from the Research Ethics Board of the University of Toronto (Protocol no. 32,650) and the Johns Hopkins Homewood IRB (HIRB00005769).

Julian House, Nicola Lacetera, Mario Macis and Nina Mazar have no financial or personal conflicts of interest to report regarding this manuscript.

CRediT authorship contribution statement

Julian House: Writing – review & editing, Methodology, Investigation, Conceptualization. **Nicola Lacetera:** Writing – review &

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

[12] 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).

editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Mario Macis:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Nina Mazar:** Writing – review & editing, Methodology, Investigation, Conceptualization.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jhealeco.2024.102914.

References

Ball, S., Eckel, C., Grossman, P.J., Zame, W., 2001. Status in markets. Q. J. Econ. 116 (1), 161–188.

Bandiera, O., Barankay, I., Rasul, I., 2013. Team incentives: evidence from a firm level experiment. J. Eur. Econ. Assoc. 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. Econ. 34, 13–25.

Bateson, M., Callow, L., Holmes, J.R., Redmond Roche, M.L., Nettle, D., 2013. Do images of 'watching eyes' induce behaviour that is more pro-social or more normative? A field experiment on littering. PLoS One 8 (12), e82055.

Becker, G.S., Elías, J.J., 2007. Introducing incentives in the market for live and cadaveric organ donations. J. Econ. Perspect. 21 (3), 3–24.

Charness, G., Grosskopf, B., 2001. Relative payoffs and happiness: an experimental study. J. Econ. Behav. Organ. 45 (3), 301–328.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Q. J. Econ. 117 (3), 817–869.

Charness, G., Masclet, D., Villeval, M.C., 2011. Competitive Preferences and Status As an Incentive: Experimental Evidence. CIRANO-Scientific Publications, pp. 2011s–20107.

Charness, G., Masclet, D., Villeval, M.C., 2014. The dark side of competition for status. Manage Sci. 60 (1), 38–55.

Chater, N., Loewenstein, G., 2022. The i-frame and the s-frame: how focusing on individual-level solutions has led behavioral public policy astray. Behav. Brain Sci. 1–60. Sep 5.

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 Manag. Leadership 19 (4), 467–489.

Duflo, E., Saez, E., 2002. Participation and investment decisions in a retirement plan: the influence of colleagues' choices. J. Public Econ. 85 (1), 121–148.

Edelman, B., Larkin, I., 2015. Social comparisons and deception across workplace hierarchies: field and experimental evidence. Org. Sci. 26 (1), 78–98.

Elias, J.J., Lacetera, N., Macis, M., 2019. Paying for kidneys? A randomized survey and choice experiment. Am. Econ. Rev. 109 (8), 2855–2888.

Gauri, V., Jamison, J., Mazar, N., Ozier, O., 2021. Motivating bureaucrats through social recognition: external validity—A tale of two states. Organ. Behav. Hum. Decis. Process. 163 (March), 117–131.

Gneezy, U., Meier, S., Rey-Biel, P., 2011. When and why incentives don't work to modify behavior. J. Econ. Perspect. 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. Am. Econ. Rev. 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. J. Public Econ. 114, 19–28.

List, J.A., Shaikh, A.M., Xu, Y., 2019. Multiple hypothesis testing in experimental economics. Exp. Econ. 22, 773–793.

Meeker, D., Linder, J.A., Fox, C.R., Friedberg, M.W., Persell, S.D., Goldstein, N.J., Knight, T.K., Hay, J.W., Doctor, J.N., 2016. Effect of behavioral interventions on inappropriate antibiotic prescribing among primary care practices: a randomized clinical trial. JAMA 315 (6), 562–570.

Munshi, K., Myaux, J., 2006. Social norms and the fertility transition. J. Dev. Econ. 80 (1), 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., Croymans, D.M., 2022. When peer comparison information harms physician well-being. Proc. Natl. Acad. Sci. 119 (29), e2121730119.

Robitaille, N., Mazar, N., Tsai, C.I., Haviv, A.M., Hardy, E., 2021. Increasing organ donor registrations with behavioral interventions: a field experiment. J. Mark. https://journals.sagepub.com/doi/full/10.1177/0022242921990070.

Steinmayr, A. (2020). MHTREG: stata module for multiple hypothesis testing controlling for FWER. https://ideas.repec.org/c/boc/bocode/s458853.html.

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. J. Public Econ. 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.