Beyond bonuses: the incentive effect of a prosocial initiative on bankers *

Nava Ashraf[†] Oriana Bandiera[‡] Alexia Delfino[§] Matteo Fossi[¶]

March 4, 2026

Preliminary draft. Please do not circulate without permission.

Abstract

We study whether a credible prosocial intervention can shift organizational culture in a high-powered incentive environment. We partner with one of the world's largest banks and embed a field experiment in the roll-out of a new volunteering initiative, randomly varying exposure across more than 700 offices and 35,000 employees at the office, team, and individual levels. The intervention increases volunteering, improves performance, reduces disengagement, and lowers turnover. Effects are substantially larger at the office level than at the individual level, consistent with strategic complementarities. Survey evidence shows that the program increases perceived cultural fit and social responsibility, indicating shifts in shared beliefs. Leadership participation is central: when senior employees do not visibly engage, cultural and performance gains may even reverse. The results suggest that credible interventions can reshape organizational culture and improve economic outcomes.

*We thank our partners at the bank for sharing the data. We thank Pascale Bourquin, Beatrice Montano, Jonathan Old, Gabriela Diaz-Pardo, Zijun Lin, Somei Miyashita and Joe Basford for excellent research assistance. All errors are our own.

[†]STICERD and Department of Economics, LSE

[‡]STICERD and Department of Economics, LSE

[§]Department of Economics, Bocconi University, LEAP and IGIER

[¶]Department of Economics, University of Southern California

1 Introduction

Organizations routinely rely on high-powered incentives to motivate performance. By tying pay and promotion to measurable output, firms can increase effort and productivity (Lazear and Moore (1984); Lazear (2000); Gneezy et al. (2003); Bandiera et al. (2007); Blanes i Vidal and Nossol (2011); Bandiera et al. (2013); Friebel et al. (2017); DellaVigna and Pope (2018)). Yet a long-standing concern is that such incentives may also distort behavior: a strong focus on individual rewards may erode cooperation, discretionary effort, and shared norms (Kandel and Lazear (1992); Prendergast (1999); Fehr and Schmidt (2000); Fehr and Falk (2002); Falk and Kosfeld (2006); Breza et al. (2018); Lazear (2018)). In environments where production depends on coordination and trust, incentives can therefore generate a fundamental tension between individual performance and collective culture.

This paper asks whether this tension is inevitable. Can a firm operating under strong performance incentives deliberately reshape its internal culture in a way that strengthens cooperation and prosocial norms, without weakening productivity? We study this question in partnership with one of the world's largest banks. The setting is deliberately demanding: banking is characterized by strong individual incentives, intense competition, and persistent concerns about cultural dysfunction (Cohn et al., 2014; Thakor, 2015; Beunza, 2019; Egan et al., 2019). We embed a large-scale field experiment in the roll-out of a new corporate volunteering initiative covering approximately 35,000 employees across more than 700 offices. Crucially, the intervention does not alter pay, monitoring, or formal incentives. Instead, it provides a visible signal of the firm's prosocial orientation and reduces the cost of engaging in prosocial activities.[1]

Studying cultural change inside organizations poses three central empirical challenges. First, culture is a group-level object: it consists of shared beliefs and norms within teams or offices. Identifying cultural change therefore requires exogenous variation involving a large number of organizational units, not just individuals. Second, culture evolves slowly. Generating credible shifts requires tracking outcomes over time. Third, culture affects behavior through social interactions. When effort and cooperation are strategically complementary, the impact of a cultural shift depends on how many others are exposed. In such settings, individual treatment effects need not coincide with equilibrium effects, and experiments that vary exposure only at the individual level may miss amplification operating at higher levels of aggregation.[2]

[1] Globally, businesses invest substantial resources in programs designed to build a positive corporate culture, with employee volunteering programs among the most prominent of these efforts (Cycyota et al., 2016). A leading provider estimates that firms recently supported more than 99 million hours of volunteering and $18.5 billion in community contributions (Benevity Inc., 2024). Surveys indicate that 70–80% of large firms offer formal employee volunteering programs, and median annual community investments exceed $20 million per firm (Basil et al., 2009; Brudney and Gazley, 2006; CECP, 2024; Perez et al., 2014).

[2] A growing literature studies interventions that enhance prosociality and social skills and documents improvements in individual behavior and workplace outcomes (Borghans et al., 2008; Ashraf and Bandiera, 2017; Deming, 2017; Kosse and Tincani, 2020; Alan et al., 2023). We complement this work by examining

Our experimental design addresses these challenges. We implement a large-scale randomized encouragement design embedded in the roll-out of the bank's new prosocial initiative. Exposure is randomized at three levels—office, team, and individual—creating exogenous differences in treatment intensity across organizational units. Offices vary in the share of exposed employees, while individual exposure remains independently assigned. This structure allows us to distinguish direct individual responses from spillover effects operating through the local cultural environment. Crucially, the design enables us to test for equilibrium amplification. If culture affects the marginal return to discretionary effort, then performance effects should increase disproportionately with treatment density within an office. By comparing individual-level effects to office-level effects under random variation in exposure, we can distinguish purely individual responses from equilibrium changes driven by shifts in the office-level cultural environment. We combine this design with comprehensive personnel records for the universe of UK employees—spanning all major functions and seniority levels—and repeated employee surveys, allowing us to track culture, performance, and turnover jointly over time.

We find that the intervention increases volunteering, improves performance, and reduces turnover and disengagement. Importantly, effects are substantially larger at the office level than at the individual level. This pattern is consistent with equilibrium amplification: as more employees internalize prosocial norms, the marginal return to discretionary effort rises for all workers. The effectiveness of this mechanism, however, depends critically on credibility. Consistent with interpreting the initiative as a signal of the organization's prosocial orientation, managerial participation plays a central role: when senior employees visibly engage, cultural shifts and performance gains are significantly stronger; when they do not, the intervention can backfire, undermining credibility and reversing the intended cultural effects.

To guide the empirical analysis, we first develop a tractable framework in which corporate culture is modeled as a shared belief about the organization's prosocial orientation (Van den Steen, 2010b; Dessein and Prat, 2022). Workers choose effort in an environment with strategic complementarities: the marginal return to exerting discretionary effort depends positively on the fraction of coworkers who internalize prosocial norms, consistent with economic theories of identity, prosocial motivation, and leadership-driven coordination (Akerlof and Kranton, 2000, 2005; Bénabou and Tirole, 2006; Bolton et al., 2013). The volunteering initiative operates as a credibility shock that increases the perceived alignment between the firm's actions and its prosocial orientation, shifting the cultural updating process and raising the equilibrium share of prosocial types within the office. Because effort is complementary in culture, even modest shifts in beliefs can generate amplified effects on performance and retention at the office level (Gartenberg et al., 2019; Bandiera et al., 2020). Finally, the model also highlights two key margins: the opportunity cost of time, which may differ across hierarchical levels, and the role of leadership participation in reinforcing

how a firm-wide intervention shifts shared cultural beliefs and generates spillover effects.

or weakening the credibility of the cultural shift.

In the empirical part, we document five main findings. First, the encouragement design substantially increases volunteering. Treated employees are six percentage points more likely to volunteer, a roughly 25 percent increase over the control mean. Importantly, take-up rises not only at the individual level but also with treatment intensity within offices, consistent with social reinforcement at the office level.

Second, the intervention improves on-the-job performance and reduces disengagement. Employees exposed to the policy receive higher performance and values ratings and are significantly less likely to be classified as low performers. The probability of "quiet quitting" — sustained low performance without exit — declines meaningfully, and turnover falls as well. These effects are economically meaningful: low performance fall by approximately one-tenth of the control mean, and turnover declines by 2 percentage points relative to a control mean of 15 percent. Taken together, these results indicate that the initiative does not come at the expense of performance or retention; if anything, it strengthens both. The improvement in performance operates primarily through within-person changes in effort rather than selection into exit. In particular, decomposing total effects into within-person changes and compositional shifts reveals that most of the gains are driven by behavioral responses among employees who stay. Selection plays a limited role, implying the initiative alters how employees behave rather than simply reshuffling the workforce.

Third, treatment effects grow sharply with the level of aggregation. Office level estimates are several times larger than individual-level effects. In particular, the share of low performers at the office level falls by roughly one-quarter relative to the baseline mean, and office level turnover declines by about one-half relative to baseline. This pattern is consistent with the strategic complementarities emphasized in the model. Individual treatment primarily affects the recipient's own beliefs and behavior, generating a modest first-round effect. By contrast, greater treatment intensity within an office shifts the equilibrium level of cultural alignment, increasing the marginal return to effort for all workers. The resulting amplification resembles an equilibrium multiplier: as the share of exposed employees rises, performance effects grow disproportionately.

Fourth, we provide direct evidence that changes in cultural perceptions are central to these outcomes. Using survey measures, we construct an index of perceived cultural fit and a measure of perceived social responsibility among colleagues. Both increase significantly in response to treatment, with effects ranging from 0.07 to 0.11 standard deviations depending on the specification, and with particularly strong and robust gains on measures related to customer orientation. As with the previous outcomes, the magnitude of the coefficients grows when moving from individual-level perceptions to office-level cultural measures. We find that offices experiencing larger increases in cultural alignment also exhibit larger improvements in performance. Additionally, we document suggestive evidence of convergence in beliefs at the office level, with dispersion in cultural perceptions declining more in more treated offices. Consistent with the model's predictions, these results suggest that a credibility shock induced by the volunteering program can indeed shift offices toward a new

cultural equilibrium.

Finally, we find that leadership plays a decisive role, in line with the predictions of our model. Senior employees face higher opportunity costs of time due to greater reliance on performance-based compensation and participate in volunteering at lower rates. We construct a measure of the junior–senior volunteering share gap and show that larger gaps are associated with weaker shifts in cultural beliefs. Moreover, treatment effects depend heavily on the baseline cultural alignment of senior employees. The impact of the intervention is significantly attenuated in offices with a higher share of seniors who have low perceived cultural fit at baseline. These findings align with the model's emphasis on credibility: leadership behavior shapes whether employees interpret the initiative as a genuine shift in organizational priorities and is therefore central to its success.

Taken together, our results suggest that the trade-off between high-powered incentives and cooperative culture is not inevitable. When leadership behavior makes cultural change credible, prosocial initiatives can reshape beliefs, boost effort through strategic complementarities, and improve a firm's economic outcomes. In this view, and in line with a long tradition in the organizational economics literature, culture is not a soft add-on to incentives, but a central component of organizational success (Guiso et al., 2015).

This paper contributes to several strands of the literature. First, we contribute to the personnel economics literature on workplace incentives (Gibbons and Waldman, 1999; Lazear, 2000, 2009; Blanes i Vidal and Nossol, 2011; Larkin, 2014; Dohmen and Falk, 2011; Holmström, 2017; DellaVigna and Pope, 2018), and in particular to work on prosocial and social incentives (Bénabou and Tirole, 2006; Ashraf et al., 2014; Ashraf and Bandiera, 2018; Breza et al., 2018; DellaVigna et al., 2022). We show that a large-scale prosocial organizational initiative, without altering pay schemes, can raise performance and reduce turnover even in a high-powered incentive environment such as banking. Relatedly, we contribute to the literature on peer effects and social interactions in firms (Bandiera et al., 2005; Mas and Moretti, 2009; Bandiera et al., 2013; Cornelissen et al., 2017) by documenting equilibrium amplification: as exposure increases within offices, effects grow disproportionately, consistent with strategic complementarities in effort.

Second, we contribute to the economics of culture (Bisin and Verdier, 2001; Guiso et al., 2006), and more specifically to the literature on corporate culture (Van den Steen, 2010b; Guiso et al., 2015; Gartenberg et al., 2019; Dessein and Prat, 2022; Graham et al., 2022). We provide large-scale experimental evidence that a prosocial organizational initiative can causally shift shared beliefs about the firm's priorities, increase perceived cultural alignment, and reduce dispersion in cultural perceptions within offices.

Third, we contribute to research on leadership and the transmission of values within firms (Bertrand and Schoar, 2003; Bandiera et al., 2007; Bloom et al., 2012; Lazear et al., 2015; Bolton et al., 2013; Bandiera et al., 2020; Hoffman and Tadelis, 2021). We show that cultural change is inherently mediated by leaders' behavior: treatment effects depend strongly on senior employees' attitudes and active participation, underscoring the central role of leadership in enabling, or obstructing, the diffusion of organizational values.

Finally, we add to the growing literature on field experiments in organizations (Blasco et al., 2019; Gosnell et al., 2020; Sandvik et al., 2020; Alan et al., 2023; Ashraf et al., 2025) by embedding a multi-level randomized encouragement design inside a large, complex organization.

The remainder of the paper is organized as follows. Section 2 presents the theoretical framework. Section 3 describes the partner bank. Section 4 introduces the data and key variables. Section 5 details the intervention and randomization design, and reports tests of experimental validity and first-stage effects. Section 6 presents the main intent-to-treat results on workplace performance and turnover. Section 7 examines cultural change and improved cultural fit as underlying mechanisms. Section 8 analyzes heterogeneity by seniority and the role of senior employees in driving cultural change. Finally, section 9 concludes.

2 Conceptual Framework

This section develops a tractable model in which corporate culture is represented as a state of shared beliefs that reduces coordination frictions and shapes organizational performance and retention (Van den Steen, 2010a,b). The model further embeds cultural evolution dynamics that combine prestige-based learning and conformist transmission to characterize convergence to a new cultural equilibrium following a credibility shock (Savitskiy et al., 2025).

Setup

Consider a single office populated by a unit mass of workers indexed by $i \in [0, 1]$. Time is discrete, $t = 0, 1, 2, \ldots$. Effort is imperfectly contractible and includes discretionary prosocial actions toward clients or coworkers that cannot be fully specified, observed, or enforced through formal contracts. In each period, workers choose binary effort $e_{it} \in \{0, 1\}$.

At $t = 0$ the firm implements a randomized prosocial initiative. Let $T \in \{0, 1\}$, where $T = 0$ denotes the pre-initiative period and $T = 1$ the post-initiative period. The realization of T is publicly observed and common knowledge within the office, and serves as a public signal of the organization's prosocial orientation. The initiative does not alter monetary incentives, such as wages or promotions, nor does it affect informational frictions or monitoring technologies.

Each worker i at time t has a binary cultural state $s_{it} \in \{H, L\}$. Type H represents internalization of a prosocial norm, the belief that the organization values discretionary actions that benefit others even when imperfectly contractible, while type L reflects a transactional view under which only explicitly incentivized actions are expected.[3]

[3]Throughout, we use the term "culture" to refer to the alignment between workers' beliefs and the firm's prosocial orientation.

Within each office, culture is summarized by the office-level state $p_t \in [0,1]$, defined as the fraction of workers in the high cultural state, $p_t = \Pr(s_{it} = H)$. The state p_t is common knowledge and enters payoffs through a coordination channel: a higher p_t increases the expected return to exerting effort by raising the likelihood that coworkers reciprocate and match discretionary actions. Thus culture affects individual incentives and generates office-level spillovers in performance.

Static Decisions

We now discuss the decision-making process of agent i in a generic period t.

Effort. In each period t, workers choose binary effort $e_{it} \in \{0,1\}$. Worker i's within-period payoff depends on her effort choice and on the office-level cultural state p_t, which enters payoffs through a coordination channel.

If worker i exerts effort, her payoff is given by

$$U(1) = b + a \cdot \mathbf{1}\{s_{it} = H\} + \gamma p_t - c, \tag{1}$$

where b captures baseline effort-dependent private returns (e.g. career incentives, reputational gains, etc.) that are independent of culture; $a > 0$ captures an additional intrinsic payoff from exerting effort for workers in the high cultural state; and $c > 0$ denotes the private cost of exerting effort. The term γp_t captures coordination benefits. In particular, $\gamma > 0$ captures strategic complementarities in effort, so that the return to exerting effort increases with p_t, the share of workers in the high cultural state. A higher value of p_t makes effort more likely to be reciprocated and matched by coworkers, thereby raising its expected payoff.

If worker i does not exert effort, her payoff is normalized to $U(0) = 0$. Hence, the net payoff from exerting effort is given by Equation (1). Worker i exerts effort if and only if

$$e_{it} = \begin{cases} 1 & \text{if } \Delta U_{it}(p_t) \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Individual output is proportional to effort. Specifically, define worker i's output in period t as

$$y_{it} = A e_{it}, \tag{2}$$

where $A > 0$ is a scale parameter, which we assume to be common across workers.

Aggregate office output is obtained by integrating individual output across the unit mass of workers:

$$Y_t = \int_0^1 y_{it}\, di = A \int_0^1 e_{it}\, di. \tag{3}$$

Thus, output is proportional to the mass of workers exerting effort. Since effort decisions depend on the cultural state p_t through Equation (1), culture affects performance by shifting individual effort incentives and, in turn, aggregate output.

Exit. In each period t, workers decide whether to remain in the firm or to exit. Let $V_i(p_t)$ denote worker i's continuation value from staying in the office when the cultural state is p_t, and let U_0 denote the outside option associated with exiting, which we take as given and assume common across workers for simplicity. Worker i remains in the office if and only if

$$V_i(p_t) \geq U_0.$$

We assume that $V_i(p_t)$ is weakly increasing in contemporaneous utility. Since Equation (1) implies that higher p_t weakly increases realized utility, it follows that $V_i(p_t)$ is weakly increasing in p_t.[4]

Leadership as a Credible Signal Leadership behavior shapes workers' beliefs about the organization's underlying objectives. Consequently, the credibility of the prosocial initiative depends on whether senior leaders visibly participate.

Suppose senior leaders differ in an underlying commitment type $\tau \in \{A, I\}$, reflecting "authentic" versus "inauthentic" alignment with the firm's prosocial orientation. Participation in the initiative requires time and attention and therefore entails an opportunity cost k_τ, with $k_A < k_I$. Each senior leader chooses a binary, publicly observable action $x_j \in \{0, 1\}$, where $x_j = 1$ denotes visible participation in the initiative ("walking the talk") and $x_j = 0$ denotes no visible participation.

Workers observe leadership participation but not leaders' types. Because participation is more costly for inauthentic leaders, observing $x_j = 1$ increases the posterior probability that leadership is authentically aligned with the prosocial shift. Participation therefore acts as a credibility-enhancing signal.

Let $\delta \equiv \delta(T, x)$ denote worker i's perceived credibility of the prosocial shift as a function of initiative status $T \in \{0, 1\}$ and observed leader participation office share $x \in [0, 1]$.[5]

Assumption 1 (Credibility ordering). *The credibility parameter satisfies the ordering*

$$\delta(1, 1) > \delta(1, 0) \geq \delta(0, 1) > \delta(0, 0).$$

That is, both the initiative and leadership participation raise credibility relative to baseline.

Assumption 2 (Credibility complementarity). *Credibility is weakly supermodular in* (T, x)*:*

$$\delta(1, 1) - \delta(0, 1) \;\geq\; \delta(1, 0) - \delta(0, 0).$$

That is, leadership participation amplifies the credibility effect of the intervention.

[4] The timing of decisions in period t is as follows: (1) worker i observes the state variables (p_t, T, x); (2) she chooses effort e_{it}; (3) payoffs are realized; and (4) she decides whether to remain in the office. Cultural updating then takes place, determining p_{t+1} only after these decisions have been made.

[5] We are implicitly imposing three important assumptions in this part of the model: (1) Leadership participation is a one-time, publicly observable decision made at the introduction of the initiative; (2) Any heterogeneity in leaders' underlying commitment to the prosocial orientation is reflected in the participation decision.

Dynamics: Cultural Convergence

The static section treats p_t as an exogenous state variable inherited from period $t-1$. We now endogenize its evolution by specifying a parsimonious updating rule for p_t based on two social learning mechanisms: (i) prestige-based learning and (ii) conformist transmission. The former captures the idea that new norms are more likely to spread if they come from high-status/leadership figures, while the latter embodies standard horizontal diffusion dynamics among peers.

Update probability and aggregate law of motion. Suppose that in each period t, cultural updating follows a Calvo-style adjustment process: a fraction $\lambda \in (0,1)$ of workers reconsider their cultural state through social learning, while the remaining fraction $1-\lambda$ retain their prior state.

Suppose worker i is visited by the "Calvo fairy." Before updating, she samples her coworkers and observes their cultural traits. Let $\hat{p}_{it}$ denote the share of sampled coworkers in the high cultural state H in worker i's sample. The probability that worker i adopts the prosocial norm is a convex combination of the credibility parameter $\delta(T,x)$ and the sampled peer share $\hat{p}_{it}$:

$$\Pr(s_{i,t+1} = H \mid \hat{p}_{it}, \delta) = (1-\eta)\,\psi(\alpha_0 + \alpha_1\delta) + \eta\,\phi(\hat{p}_{it}), \tag{4}$$

where

- $\psi(z) = \frac{1}{1+e^{-z}}$ is the logistic function capturing prestige-based learning, with $\alpha_1 > 0$ so that higher credibility δ increases the probability of adopting the prosocial norm.
- $\phi(q) = \frac{q^\rho}{q^\rho + (1-q)^\rho}$, with $\rho > 1$, captures conformist transmission, so that workers disproportionately adopt the majority trait.
- $\eta \in [0,1]$ governs the relative importance of conformity versus prestige-based updating. When $\eta = 0$, updating is only credibility-driven; when $\eta = 1$, it is purely conformist.

In a large population with random sampling, the sampled share coincides with the population share in the limit, so that $\hat{p}_{it} = p_t$ for aggregating purposes. The aggregate cultural state therefore evolves according to

$$p_{t+1} = (1-\lambda)p_t + \lambda\,G(p_t;\delta), \tag{5}$$

where

$$G(p;\delta) \equiv (1-\eta)\,\psi(\alpha_0 + \alpha_1\delta) + \eta\,\phi(p). \tag{6}$$

and the function $G(p;\delta)$ is increasing in both p and δ.

Definition 1 (Cultural equilibrium). *Given a credibility level δ, a cultural equilibrium is a fixed point $p^*(\delta)$ satisfying*

$$p^*(\delta) = G\big(p^*(\delta);\delta\big). \tag{7}$$

Theoretical Results

Proposition 1 (Equilibrium culture is increasing in credibility). *Let $p^*(\delta)$ denote the largest fixed point of $G(\cdot;\delta)$ on $[0,1]$. Suppose $G(p;\delta)$ is continuous in p and, for every fixed $p \in [0,1]$, weakly increasing in δ. Then $p^*(\delta)$ is weakly increasing in δ.*

Proof. See Appendix A. □

Corollary 1.1 (Comparative statics in credibility). *Given Proposition 1 and Assumptions 1 and 2, the following statements hold:*

(i) The intervention raises equilibrium culture:

$$p^*\big(\delta(1,x)\big) \geq p^*\big(\delta(0,x)\big) \quad \text{for any fixed } x.$$

(ii) Leadership participation raises equilibrium culture:

$$p^*\big(\delta(T,1)\big) \geq p^*\big(\delta(T,0)\big) \quad \text{for any fixed } T.$$

(iii) Intervention and leadership are complementary in equilibrium culture:

$$p^*\big(\delta(1,1)\big) - p^*\big(\delta(0,1)\big) \;\geq\; p^*\big(\delta(1,0)\big) - p^*\big(\delta(0,0)\big).$$

Proof. See Appendix A. □

Proposition 2 (Effort is increasing in culture). *For any worker i, holding her cultural type fixed at $s_{it} \in \{H, L\}$, if $p_{t+1} \geq p_t$, then*

$$e_{i,t+1} \geq e_{it}.$$

Proof. See Appendix A. □

Corollary 2.1 (Individual-level implications of cultural change). *Given Proposition 2, if $p_{t+1} \geq p_t$, then for any worker i:*

(i) Individual output increases:

$$y_{i,t+1} \geq y_{it}.$$

(ii) The continuation value increases, thus exit becomes weakly less likely:

$$V_i(p_{t+1}) \geq V_i(p_t).$$

Proof. See Appendix A. □

Corollary 2.2 (Office-level amplification). *Given Proposition 2, suppose $p_{t+1} > p_t$ and there exists a measurable set $S \subseteq [0,1]$ with positive Lebesgue measure $\mu(S) > 0$ such that, holding workers' cultural types fixed, for all $i \in S$,*

$$e_{i,t+1} > e_{it}, \qquad V_i(p_{t+1}) > V_i(p_t).$$

Then:

1. *Aggregate output strictly increases:*

$$Y_{t+1} - Y_t = A \int_0^1 (e_{i,t+1} - e_{it})\, di \;\geq\; A\,\mu(S) \;>\; 0.$$

2. *Turnover*

$$\tau_t = \int_0^1 \mathbf{1}\{V_i(p_t) < U_0\} di$$

strictly decreases:

$$\tau_{t+1} - \tau_t < 0.$$

Proof. See Appendix A. □

Proposition 3 (Threshold Effects and Multiple Equilibria under Strong Conformity). *Consider the aggregate law of motion*

$$p_{t+1} = (1-\lambda)p_t + \lambda G(p_t; \delta),$$

defined in equation (5), with $\lambda \in (0,1)$ and continuous $G(\cdot;\delta)$. For parameter values of (η, ρ) that generate an S-shaped aggregate map (i.e., sufficiently strong conformity), the equation $G(p;\delta) = p$ admits three fixed points

$$p_L^*(\delta) < p^c(\delta) < p_H^*(\delta),$$

where $p^c(\delta)$ denotes the critical threshold share.[6]

Suppose that

$$G(p;\delta) > p \quad \text{for } p \in (p^c, p_H^*),$$
$$G(p;\delta) < p \quad \text{for } p \in (p_L^*, p^c),$$

so that p_L^* and p_H^* are locally stable and p^c is unstable. Then:

1. *If $p_0 > p^c(\delta)$, the sequence (p_t) converges to $p_H^*(\delta)$.*
2. *If $p_0 < p^c(\delta)$, the sequence (p_t) converges to $p_L^*(\delta)$.*

Proof. See Appendix A. □

Corollary 3.1 (Convergence to the High Equilibrium under Pure Conformity). *Under the conditions of Proposition 3, suppose:*

1. *Updating is purely conformist ($\eta = 1$).*
2. *There exists a unique stable interior fixed point $p_H^* \in (\frac{1}{2}, 1)$ satisfying*

$$\phi(p_H^*) = p_H^*, \quad \phi(p) > p \text{ for } p \in (\tfrac{1}{2}, p_H^*), \quad \phi(p) < p \text{ for } p \in (p_H^*, 1).$$

[6] This requires assuming $\rho > 1$ and that η places sufficient weight on the conformist component so that the slope of $G(p;\delta) = \phi(p)$ exceeds one in a neighborhood of the midpoint.

3. $p_0 \in (\frac{1}{2}, p_H^*)$.

Then the sequence (p_t) increases monotonically and converges to p_H^*.[7]

Proof. See Appendix A. □

Corollary 3.2 (Declining Cultural Dispersion along the Convergence Path). *Under the conditions of Corollary 3.1, the within-office variance*

$$\mathrm{Var}(s_{it}) = p_t(1 - p_t)$$

strictly decreases along the convergence path.

Proof. See Appendix A. □

Testable Implications

Empirical Prediction 1 (Equilibrium Culture and Leadership Complementarity). *Greater exposure to the intervention raises the equilibrium level of cultural alignment within the office. Leadership participation amplifies this effect: offices with more engaged leaders experience systematically larger shifts in cultural alignment for a given level of exposure.*

Empirical Prediction 2 (Performance and Retention Effects). *Higher equilibrium culture increases performance and reduces exit; due to strategic complementarities, these effects are stronger at the office level than at the individual level.*

Empirical Prediction 3 (Cultural Convergence). *Along the convergence path toward the high-culture equilibrium, within-office cultural dispersion declines.*

Empirical Prediction 4 (Effort-Based Performance Gains). *Increases in equilibrium culture raise discretionary effort at the individual level, so performance improvements arise primarily from within-worker behavioral changes rather than workforce reallocation.*

3 Context

Our partner is a leading global bank with over 200,000 employees worldwide, of whom approximately 35,000 work in the UK, which account for 13% of total revenues (as of 2020). We collaborate with the UK division. Since the bank is headquartered in the UK,

[7] The conformist updating function

$$\phi(p) = \frac{p^\rho}{p^\rho + (1-p)^\rho}, \qquad \rho > 1,$$

is symmetric around $p = \frac{1}{2}$. Under pure conformity, this symmetry implies $\phi(\frac{1}{2}) = \frac{1}{2}$; moreover, when $\rho > 1$, we have $\phi(p) > p$ for $p > \frac{1}{2}$ and $\phi(p) < p$ for $p < \frac{1}{2}$. Hence, $\frac{1}{2}$ is the unstable threshold separating the two regions of convergence.

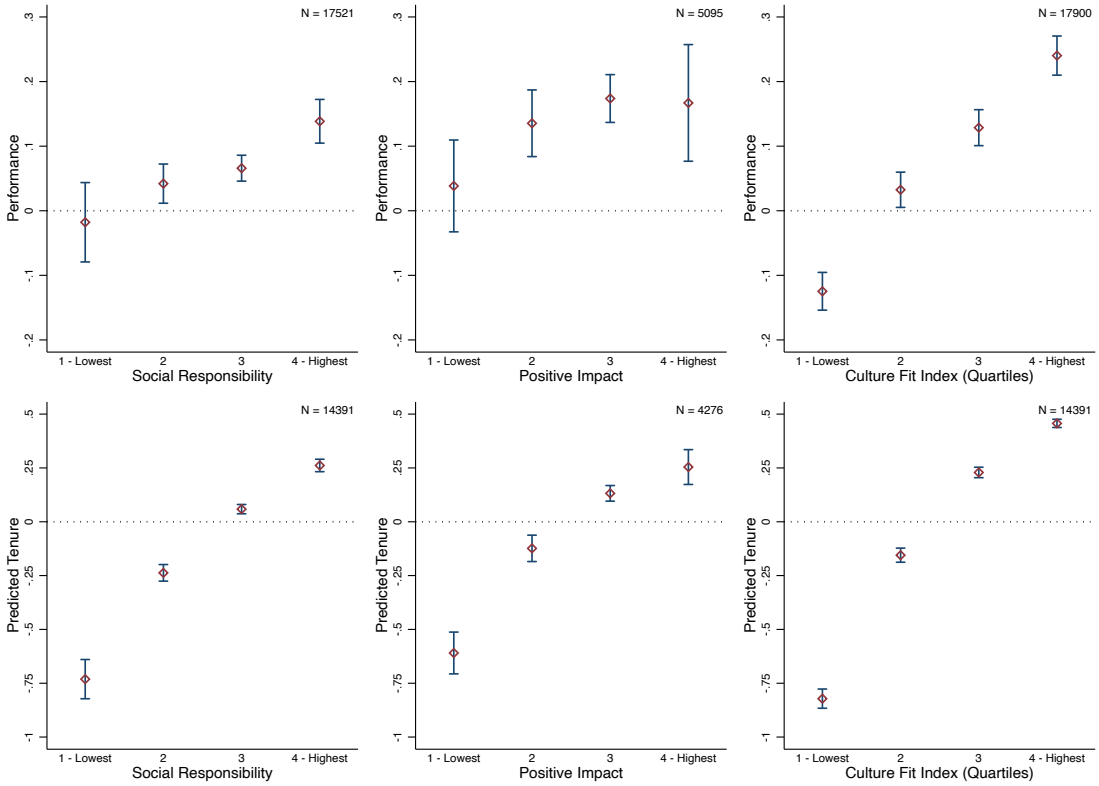
this gives us access to a broad set of employees, from domestic commercial banking to global banking functions based at the headquarters. Figure C1 shows the distribution of the bank's roughly 700 offices across the country. Figure C2 in the Appendix, shows the distribution of employees by area codes.

The partner bank has an organizational and incentive structure similar to other large global financial institutions. It comprises more than ten business divisions, the largest being retail and commercial banking. Performance is measured through a scorecard system designed to align the bank's strategic objectives with incentives at different hierarchical levels, including the board and executive directors. Scorecards are set at a high level and cascaded to business lines and regions, where country-division managers define the specific components to reflect the long-term strategy of the bank; each employee's scorecard is then tailored to their functional role. We will discuss the two main performance measures in Section 4.

Employees are classified in 9 different hierarchical levels ("career bands"), which differ in managerial responsibilities as well as pay grade. The top career bands represent around 5% of the total workforce and 40% of employees belong to the two bottom entry levels. As documented in Cullen and Perez-Truglia (2023) and Bircan et al. (2022) – and similarly to many other multinational companies – women represent over half of entry level positions, but become a minority in top levels (less than one-third in our context). Earnings in all career bands have both fixed and variable components, and the level and share of variable pay increase as employees move up the career ladder. We discuss seniority in more detail in Section 8.

The partnership was born from the bank's interest in improving its culture and reconnecting bankers to the communities they serve by leveraging and encouraging prosocial actions, rather than providing an additional perk. This reflects anecdotal evidence from management that, especially in the aftermath of the global financial crisis, bankers had become detached from the social purpose of their work, stifling motivation and client relationships. Figure 1 confirms this pattern: correlational evidence between survey answers, performance, and predicted tenure shows a significant positive relationship between social responsibility, perceived impact, cultural fit and these economic measures.

Figure 1: Predicted Tenure and Performance by Social Impact and Cultural Fit



Notes: *Each panel plots average outcomes (y-axis) by baseline cultural attitudes (x-axis). The top row reports standardized performance ratings, and the bottom row reports standardized predicted tenure. Performance and predicted tenure are normalized to have mean zero and unit variance in the estimation sample. On the x-axis, social responsibility and positive impact are measured on four-point scales (1 = lowest, 4 = highest), while the culture fit index corresponds to quartiles of the standardized beliefs measure. The sample is restricted to pre-treatment periods (timeID* < 3*), and outcomes are averaged at the individual level prior to plotting. 95% confidence intervals.*

4 Data

We have access to a rich set of confidential survey and administrative data covering the entire population of UK employees.[8] These data come from three main sources:

[8] We exclude pilot observations (706) and a list of employees whom the bank classified as "social impact champions" prior to the policy and were not included in the randomization (1,647). This reduces our sample from 36,851 to 34,498 employees.

1. *HR data between 2018 and 2020* – This data covers the universe of UK employees and contains measures of performance, turnover, career band, tenure and job role, as well as individual, team, and office identifiers. This data is at the individual-semester level, updated twice a year in June and December. It thus covers a total of six semesters, from January 2018 to June 2020.

2. *Employee survey data between 2017 and 2020* – Every year the bank collects employee surveys to measure workforce sentiment towards the bank, including alignment with organizational values, intent to stay, job satisfaction and perceptions of coworkers or managers. Additionally, it contains self-reported indicators for the gender, ethnicity and age of the respondents. The average yearly response rate is approximately 40%.

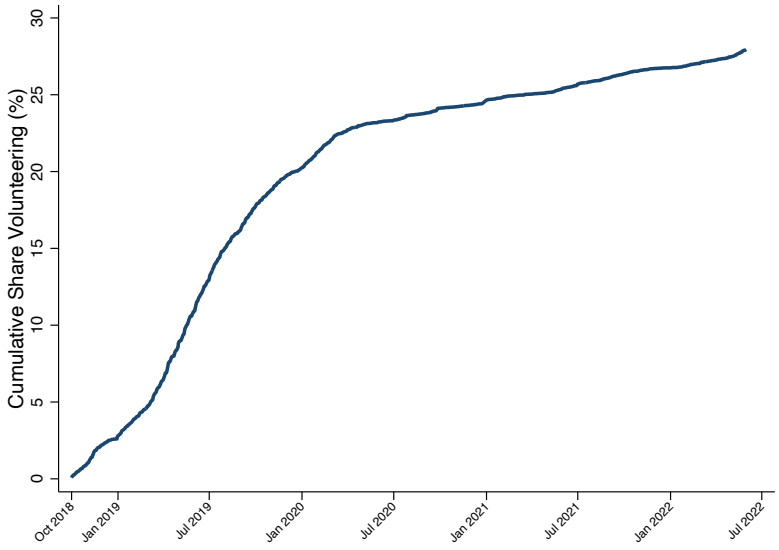
3. *Volunteering data between 2018 and 2022* – Individual-level records on volunteering capture each sign-up and participation, including the date, hours contributed (personal and work), host organization, type of activity, whether it was done alone or in a team, and whether it was employee organized or centrally organized.

For our main analysis, we merge the three sources into a single panel dataset at the individual–semester level; Appendix B.1 describes the construction of the panel in detail. Additionally, we construct team- and office-level datasets. Offices are defined based on postal codes, and teams are defined as departments (e.g., IT, Administration) within each office. The construction of these identifiers is detailed in Section B.2 in the Appendix. We also construct a cross-sectional dataset for the post-treatment period, which we use to study outcomes such as turnover.

Main outcome variables Our main outcomes belong to three different categories:

1. Volunteering – For every individual in our data, our main outcome variable for volunteering is an indicator variable equal to one if the person ever engaged in volunteering. Figure 2 shows that by the end of the study period approximately 30% people in our sample had volunteered at least once in our study period.

Figure 2: Volunteering Take Up



2. On-the-job Performance – We consider the following main outcome variables, both measured on a scale 1 (min) to 4 (max).

   (a) *Performance rating:* this is the line manager's objective assessment of an employee's performance relative to the expectations for their role.

   As mentioned in Section 3, performance is organized around a scorecard system. Each year (or bi-annually), senior management at the division level sets strategic goals, which are translated into concrete objectives in a scorecard and cascaded to business lines and regions. Expectations for each role are written down as specific targets in the employee's scorecard, and line managers then rate employees based on how well they meet these targets.[9]

   Given that it is based on achieving specific, written down goals that are meant to capture value creation in each role, this is a relatively hard and objective measure of performance and, at least in the bank's intention, is aimed at capturing how productive an employee is within their division and seniority level.

[9] For instance, employees managing clients' assets have performance objectives tied to financial indicators such as portfolio profitability relative to a benchmark index, net new assets, and fee income. By contrast, employees managing mortgage portfolios have objectives linked to net lending, lending quality, and default rates. Back office employees are evaluated on operational targets such as delivering specific digital process upgrades on time and within budget, reducing processing times or error rates, and resolving a given share of assistance requests within agreed service level standards.

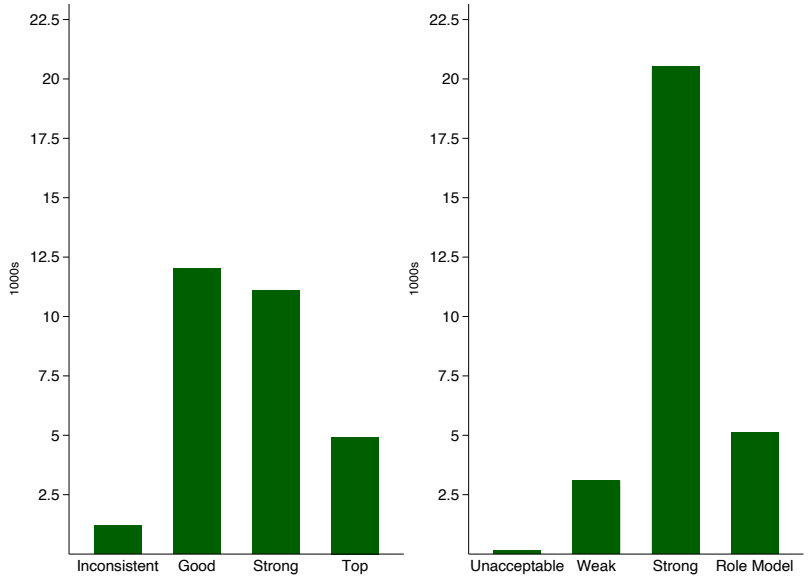
(b) *Values rating:* this is the line manager's subjective assessment of an employee and is aimed at capturing whether the employee displays qualities such as integrity, cooperation, and connection with customers (two in five employees are in client facing roles). This measure relies on the line manager's opinion and likely also reflects the line manager's own perceptions. We interpret this performance metric as a measure of people skills, in line with Hoffman and Tadelis (2021).

(c) *Low Performance*: Indicator equal to one if the employee's performance rating is classified as "Inconsistent" (1/4).

(d) *Low Values*: Indicator equal to one if the employee's values rating is classified as "Unacceptable" or "Weak" (1 or 2/4).[10]

Figure 3 shows histograms of these measures at baseline. For performance, most employees are in the middle categories "Good" and "Strong," with a sizable group rated "Top" and only a small fraction rated "Inconsistent." By contrast, values show much less variation and are much more skewed.

Figure 3: Performance (Left - Performance Rating, Right - Values rating)



3. Turnover and Disengagement

[10]We include both categories, unlike for Low Performance, because very few employees are rated 1/4 on values, and using only that category would leave too little variation to perform meaningful inference.

(a) *Quiet quitting*: We construct three distinct measures: (i) an indicator equal to one if an employee has a performance rating of at most 2 (on the 1–4 scale) in 3 out of the 4 semesters post-treatment; (ii) an indicator equal to one if an employee has a values rating of at most 2 (on the 1–4 scale) in 3 out of the 4 semesters post-treatment; and (iii) is the union of (i) and (ii).

While the Low Performance and Low Values indicators capture low on-the-job performance in a given semester, the quiet quitting measures are designed to identify employees who are persistently disengaged while neither leaving the firm voluntarily nor being dismissed.[11]

(b) *Exit*: indicator equal to one if an employee left the bank.[12]

4. Organizational Culture – The employee surveys ask workers to rate their agreement with 17 statements (e.g., "I am proud to say I work for this bank") on a 1–5 Likert scale (1 = strongly disagree, 5 = strongly agree). The full list of questions is reported in Appendix B.3.

(a) *Social responsibility:* The statement "Where I work, people have a strong sense of responsibility to society" is our closest direct measure of prosociality in the organization, so we keep it separate.[13]

(b) *Indexes:* Using 16 of the 17 measures (excluding social responsibility), we construct our culture index following the standardization procedure in Kling et al. (2007): (i) standardize each question using the baseline mean and standard deviation; (ii) take the simple average across the standardized items; and (iii) standardize this average again using the baseline mean and standard deviation. Additionally, for the same 16 items, we also take the first principal component as a separate measure.[14]

Figure 4 plots the beliefs index at baseline (before standardization, on the 1–5 scale). Most employees are clustered between about 3 and 4.25, with a mean around 3.6 and a substantial left tail.

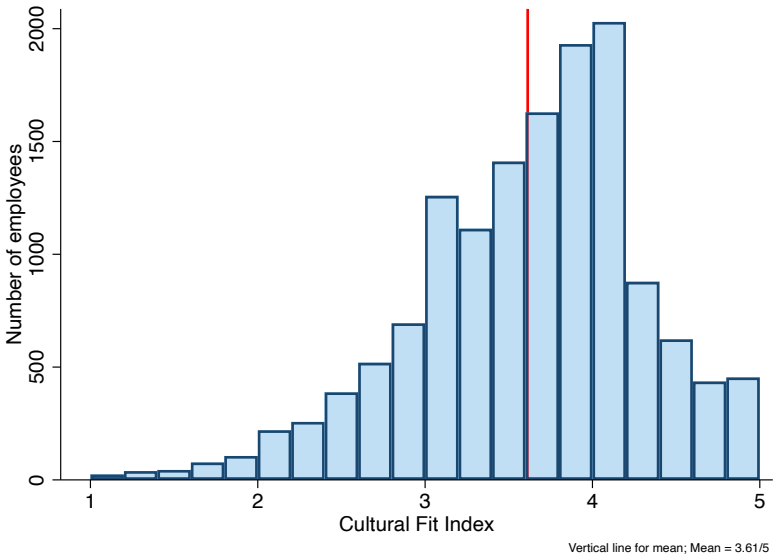
[11] Disengagement is a major issue in large firms due mainly to monitoring constraints and rigid labor market regulations making it difficult to dismiss low effort employees. Gallup's *State of the Global Workplace 2025* reports that 21% of employees are disengaged, with associated productivity losses of about $438 billion globally. During our study period (2018–2020), the global disengagement rate fluctuated between 19% and 22%, almost twice its pre–Global Financial Crisis level.

[12] HR data distinguish between voluntary and involuntary exits. However, based on discussions with the bank's HR, we were advised that these indicators are unreliable and may not accurately reflect the true reason for leaving. For instance, a departure recorded as "voluntary" may in practice follow strong pressure from the manager to find another job, and we cannot systematically distinguish such cases from the cases in which the employee independently decided to leave the bank.

[13] This question was not asked in the fourth survey wave, so we observe it once post-treatment.

[14] Pairwise correlations among the 16 beliefs items range from about 0.30 to 0.80, with most lying between 0.40 and 0.65. The correlation between the first PCA factor and the standardized index is above 0.99.

Figure 4: Cultural Fit Index (Not Standardized)



In the Appendix, we report semester-average summary statistics for the aforementioned variables at the individual, team, and office level (Table B7). Additionally, Table B8 reports summary statistics for all the main survey measures, including the list of culture fit questions and demographics (notably, 85% of our sample is white, 14% are born outside the UK, 52% are female, and 45% are in customer-facing roles).[15]

5 Empirical Design

5.1 Intervention

In 2018, the bank introduced a new volunteering policy with two main components:

1. The bank set up a centralized online platform that makes it easy for employees to sign up for volunteering activities or propose their own initiatives. This reduces search and coordination costs.

2. The bank grants employees two days of paid leave per year to be used for volunteering,

[15]HR data does not contain demographics, so the only demographic measures we have are self-reported in the employee survey. However, since these characteristics are time invariant and approximately 80% of employees respond to the survey at least once over the four year period, we observe demographics for most of the sample.

subject to approval by their line manager. This reduces the opportunity cost of volunteering, since employees can participate without sacrificing their annual leave.

In practice, participating in the program typically works as follows. First, an employee registers on the platform. Secondly, she browses the available volunteering opportunities. After selecting an activity that requires X hours of volunteering, the employee requests approval from their line manager. If the manager approves, the employee is allowed to take time off to participate in the activity during her working hours.

Employees may sign up for any activities listed on the platform, and they can split their volunteering across multiple events of any duration (full day, half day, 2 hours, etc.). As long as the cumulative volunteering time in a given year does not exceed 14 hours (equivalent to two working days), the bank treats this as additional paid leave for volunteering and the employee does not have to use their regular annual leave. Once the 14-hour threshold is exceeded, any further volunteering time must be taken from the employee's standard annual leave conditions.

The policy is a large investment for the bank and a meaningful incentive for employees. Even leaving aside the cost of building and running the online platform, our back-of-the-envelope calculations suggest that the funds allocated to the policy amount to about 1% of the annual wage bill (as of 2020). From the employee's point of view, the two extra days of paid volunteering leave correspond to approximately 10% of the average annual leave endowment.

Unlike many workplace initiatives, such as team-building exercises or managerial training programs that target culture only indirectly, this policy was explicitly designed to foster prosocial behavior by lowering its costs and responding to employees' own demand for meaningful opportunities to "do good." According to interviews with managers at the partner bank, the policy was motivated by the anecdotal observation that employees who perceived a stronger positive impact of their work on society tended to report higher job satisfaction, collaborate more effectively with colleagues, and exhibit lower turnover. In parallel, focus groups we conducted within the bank revealed systematic misperceptions about coworkers' prosocial motives. On average, employees underestimate how much their colleagues care about contributing to society, whether through personal volunteering outside of work or through socially oriented decisions on the job (for example, screening clients or projects based on social impact). This pattern is consistent with the evidence in Ashraf et al. (2020) and with the correlational evidence documented in Figure 1.

Take up is high. In particular, as shown in Figure 2, approximately 30% of employees volunteer at least once, and the average time spent volunteering is just above 9 hours, with clear bunching around 7 hours, or one working day (Figure C3). A non-trivial share of volunteers go beyond 14 hours, indicating that some employees are willing to use their own annual leave to participate in the program.[16]

[16] According to the Chief Executives for Corporate Purpose report (CECP, 2024), based on data from 219 major firms worldwide, the median employee volunteering take up rate in 2023 was 23%, with an average

5.2 Randomization

We embed a randomized controlled trial in the bank's initial roll out of the volunteering policy in the UK. We implement an encouragement design that randomizes exposure by sending to treated employees an email that introduces the new volunteering policy, explains how to participate through the platform, and highlights the two additional days of paid leave that can be used for volunteering activities recorded on the platform. The encouragement is effective at increasing policy take up, as we will show. This is noteworthy since control employees also had access to the platform in the exact same way.

The randomization was implemented in three stages by varying exposure to the policy at the office, team, and individual level (Appendix B.4 discusses sample stratification):[17]

Step 1: **Office level (July 2018)** – Each office was randomly assigned to one of five treatment arms, defined by the target share of treated teams within the office: 0%, 25%, 50%, 75%, or 100%.

Step 2: **Team level (July 2018)** – Conditional on the office assignment, we randomly selected teams up to the treatment share defined in Step 1. All employees in the treated teams received the encouragement email in July.

Step 3: **Individual level (October 2018)** – 40% of the employees who had not yet been treated were randomly selected to receive the encouragement email.[18]

Overall, 68% of employees received the treatment email in either July 2018 or October 2018. We construct treatment variables at the individual, team, and office level as follows. At the individual level, treatment is an indicator equal to one if the employee received the randomized email about the volunteering policy. At the team and office level, treatment is a continuous variable equal to the share of employees in that team or office (as defined in June 2018) who received the randomized email in either July or October.

This multi-tiered design creates exogenous variation in exposure to the volunteering policy at the office, team, and individual level, which we leverage to measure spillovers at the team and office level. Figure 5 confirms that the experiment successfully generated substantial variation in treatment shares across both offices and teams, with larger variation at the office level.

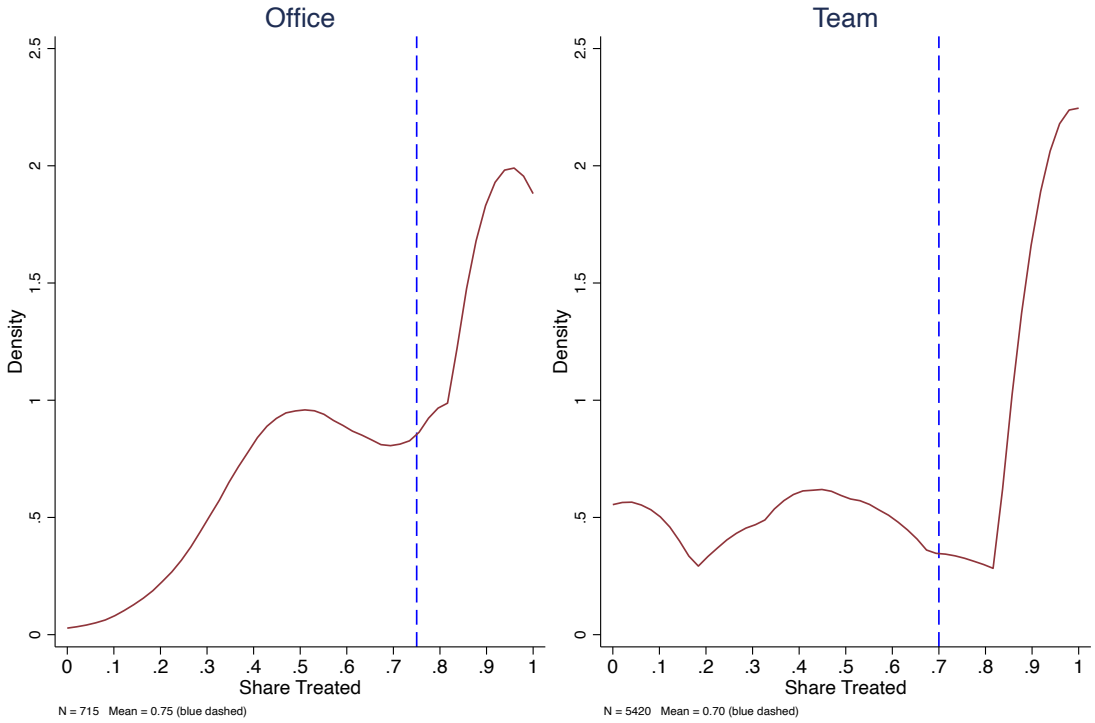
---

of 1.8 volunteer hours per employee, both substantially lower than in our partner bank.

[17] To illustrate, suppose that Office X is randomly assigned to the 50% arm and contains four teams. In July, we randomly select two of the four teams for treatment and send the encouragement email to all employees in those two teams; the remaining two teams form the control group at this stage. Among the people belonging to untreated teams, in October we randomly select approximately 40% of them and send them the encouragement email as well.

[18] Our partner bank was initially reluctant to treat more than half of the employees, but decided to increase the treatment share after observing preliminary results on the email effectiveness. This allowed us to exploit an additional layer of variation, at the individual level, which we leverage in our analysis.

Figure 5: Share Treated in Offices and Teams



5.3 Experimental Validity

We now confirm the empirical validity of our multi-step randomization approach. First, Figure C4 in the Appendix shows that the individual-level randomization implemented in October 2018 preserves the monotonicity of treatment assignment across offices in different density groups. In particular, we confirm that if Office A was assigned a higher exposure probability than Office B, then after the individual-level randomization, Office A continues to have a (weakly) higher share of treated employees than Office B.

Additionally, we test whether the October individual-level assignment is consistent with random assignment within offices. To do so, we follow the approach of Lavy and Schlosser (2011). We implement this diagnostic using Monte Carlo simulations. For each office, holding fixed office size and the observed share of treated employees, we generate 1,000 hypothetical treatment assignments by randomly reallocating treatment across individuals within the office. For each simulated assignment, we compute the standard deviation of individual treatment status. We then compare the resulting simulated distribution of standard deviations to the one observed in the data.

Figure C5 in the Appendix reports the office-level Lavy and Schlosser (2011) diagnostic

and provides strong evidence of the randomness of the October assignment. In particular, the observed and simulated distributions of within-office treatment dispersion closely overlap, with very similar shapes and support, and a difference in their means of 0.03 standard deviations. As a robustness check, Figure C6 reports the same diagnostic at the team level. Specifically, we compare the distribution of treatment dispersion generated under random assignment within teams to the distribution observed in the data and confirm the evidence found at the office-level.[19]

Finally, we conduct balance tests at the individual, team, and office levels. For each level of aggregation, p-values are adjusted jointly across all baseline characteristics (independent of the table in which they are reported) using the Romano–Wolf stepdown procedure of List et al. (2019), which controls the family-wise error rate under arbitrary dependence across outcomes. We rely on these adjusted p-values because balance tests involve a large number of correlated baseline comparisons; without correction, conventional p-values would mechanically generate false positives. The regression specification mirrors equation (8), except that pre-treatment measures are used as outcomes. Appendix Table B1 reports balance tests for the main outcome variables, Table B2 reports balance tests for additional variables from HR records, and Table B3 reports balance tests for additional survey measures.[20]

At the office level, all tested variables are balanced with List-adjusted p-values above 5%, with the sole exception of tenure. At the team level, we detect additional imbalance in values, while at the individual level turnover differs across treatment status. Importantly, the broad set of cultural fit measures is very well balanced, with only one question exhibiting imbalance at the team level. Given the large number of baseline outcomes tested, this limited patterns suggests that the randomization was broadly successful in providing a balanced sample. Nevertheless, to address any residual concern that these imbalances could affect our estimates, we control for baseline outcome measures throughout the analysis and conduct robustness checks using richer sets of covariates.

5.4 First-stage Effects on Volunteering

We now test whether our encouragement design was effective in increasing volunteering take-up. In particular, we estimate the ITT effect of the randomized email on participation using cross-sectional regressions. The outcome variables are indicators for ever registering on the platform and ever volunteering (or the corresponding shares at the team and office level). For each $u \in \{i, t, o\}$ (i = individual, t = team, o = office), we estimate:

$$y_{u,s} = \alpha + \beta T_u + \lambda_s + \varepsilon_{u,s}, \tag{8}$$

[19] Note that the office- and team-level analyses use different samples. The standard deviation of treatment assignment is only defined for units in which some individuals are treated and others are not. Because treatment shares vary less within teams than within offices, this restriction excludes more teams than offices from the simulations.

[20] For completeness, Tables B4, B5, and B6 report the corresponding estimates with conventional (not adjusted) p-values.

where $y_{u,s}$ is an indicator at the individual level or a share at the team and office level, T_u is an indicator for being treated at the individual level (and the corresponding treatment share at the team and office level), and λ_s denotes strata fixed effects. Standard errors are clustered by office for regressions at the individual and team level.

Table B10 in the Appendix reports the effects on registration: treated employees are 11 p.p more likely to register on the platform than controls, relative to a control mean of 48.4%. The corresponding effects at the team and office level are of similar magnitude. More importantly, when we turn to actual volunteering in Table 1, we find that treated individuals are about 6 percentage points more likely to volunteer, which is roughly a 25% increase over the control mean of 23.6%. The effect sizes are even larger when outcomes are aggregated to the team and office level.

Notably, take up varies broadly across offices, with some reaching participation rates above 50%, while others hardly 10% even several years after the policy is introduced. This variation, despite a common corporate policy and financial incentives, suggests that office specific factors such as communication, coordination, and local managerial practices play an important role in shaping volunteering behavior (Appendix, Figure C7).

Table 1: Outcome Variable: Volunteering

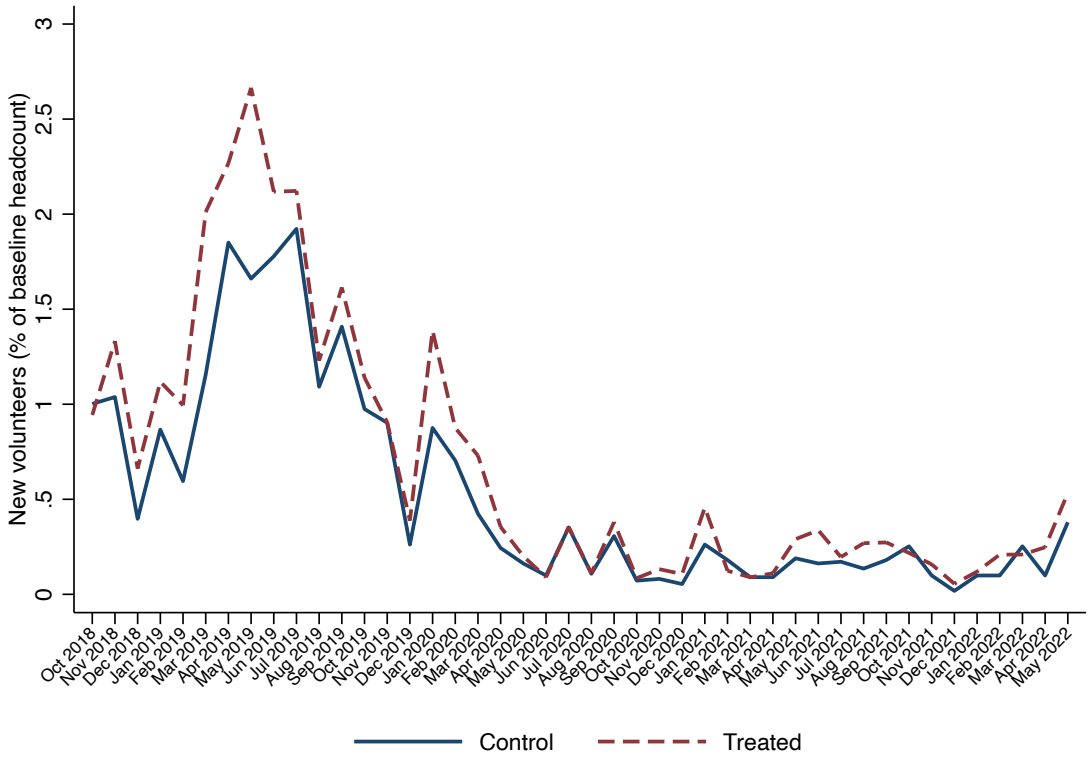
	T(SE)[P]	$\overline{Y}(SD)$	N
Individual	0.0594***	0.2364	33720
	(0.0154)	(0.4249)	
	[0.0001]		
Team	0.0733***	0.3036	5367
	(0.0209)	(0.3658)	
	[0.0005]		
Office	0.0841**	0.4456	713
	(0.0340)	(0.2260)	
	[0.0137]		

Notes: *Each column reports estimates from regressions of the volunteering indicator on treatment assignment. Individual and team regressions include strata fixed effects and cluster standard errors at the office level. Office-level regressions include strata fixed effects and report heteroskedasticity-robust standard errors.* $\overline{Y}$ *reports the control group mean at the individual level and the sample mean at the team and office level; standard deviations are reported in parentheses. P-values are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

Dynamics Figure 6 plots monthly take-up of volunteering, measured as the share of baseline employees in each treatment group who volunteer for the first time. From October 2018 through April 2020, the treated group displays higher first-time volunteering rates than the control group in every single month. Thereafter, the two series converge, likely reflecting the sharp drop in overall participation following the first UK lockdown in late March 2020 and during the subsequent pandemic period.

As shown in Figure C8 in the Appendix, this pattern translates into a large and persistent cumulative gap between the two groups. By the end of the study period, the volunteering rate in the treatment group is approximately one-third higher than in the control group.

Figure 6: Volunteering: Monthly Take Up by Treatment Group



To provide causal estimates on the dynamics of the ITT effect on volunteering over time, we turn to our panel dataset. First, we estimate a regression in which the treatment indicator is interacted with post-treatment semester dummies. Appendix Figure C9 plots the coefficient on treatment in semester 3 and, for semesters 4–6, the linear combination of this coefficient with the corresponding treatment by semester interaction terms, showing positive ITT estimates over the time horizon.

Since the volunteering variable in the panel, $Y_{i\tau}$, is an indicator that switches from 0 to 1 in the first semester in which an employee volunteers, increases in volunteering over time are partly mechanical. To strip out this mechanical component, we define

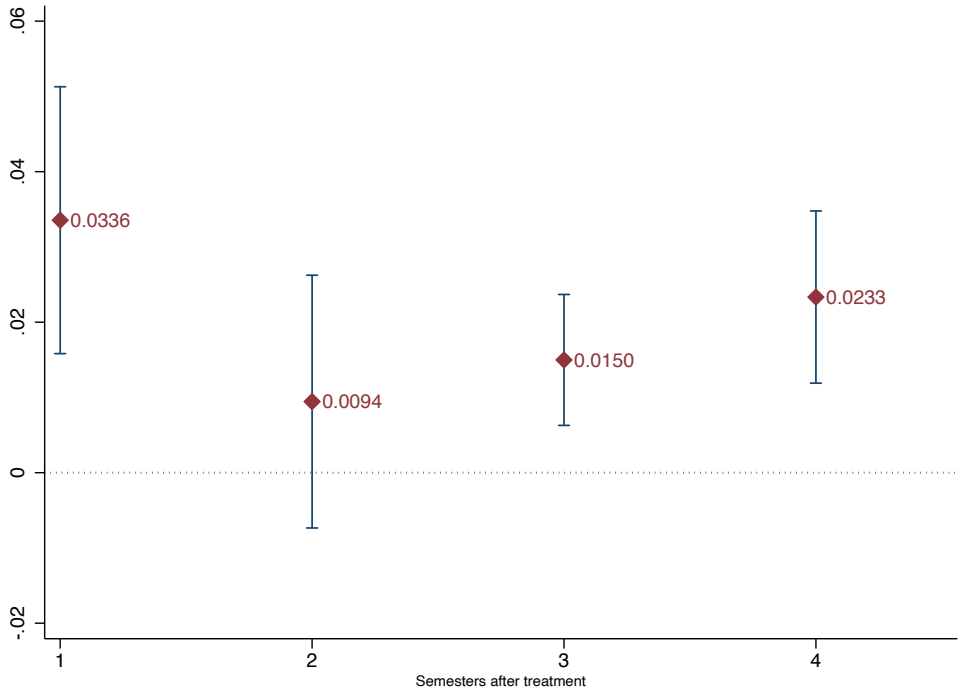
$$\Delta Y_{i\tau} = Y_{i\tau} - Y_{i,\tau-1},$$

where $Y_{i\tau}$ is the ever–volunteered indicator. By construction, $\Delta Y_{i\tau} = 1$ in the semester of first volunteering and 0 otherwise. For each post–treatment semester τ, we restrict the sample to employees with $Y_{i,\tau-1} = 0$ (those who have not yet volunteered) and regress $\Delta Y_{i\tau}$ on treatment, controlling for strata and semester fixed effects. The resulting coefficients can be interpreted as ITT effects on the hazard rate, that is, on the conditional probability that an employee who has not yet volunteered starts volunteering in semester τ.[21]

Figure 7 shows the results of this estimation. Treated employees are about 3.4 percentage points more likely than controls to volunteer for the first time in the first post treatment semester. In post-treatment semesters 2, 3, and 4, among employees who have not yet volunteered, the probability of first volunteering is still higher in the treatment group by roughly 0.9, 1.5, and 2.3 percentage points, respectively. The estimate for semester 2 is imprecise, while the effects in semesters 3 and 4 are positive and statistically significant ($p < 0.05$). We conclude that the encouragement design strongly raises the probability of first volunteering immediately after the email and continues to increase the hazard of first volunteering over the rest of the study period. Yet the results are stronger in the first semester after treatment, which is when the policy has the highest informational value.

[21] This can only be done at the individual level because $\Delta Y_{i\tau}$ is defined using the timing of each worker's first adoption. Once we aggregate to teams or offices, we only observe shares volunteering in each period and can no longer identify who is a new adopter.

Figure 7: Volunteering (Hazard)



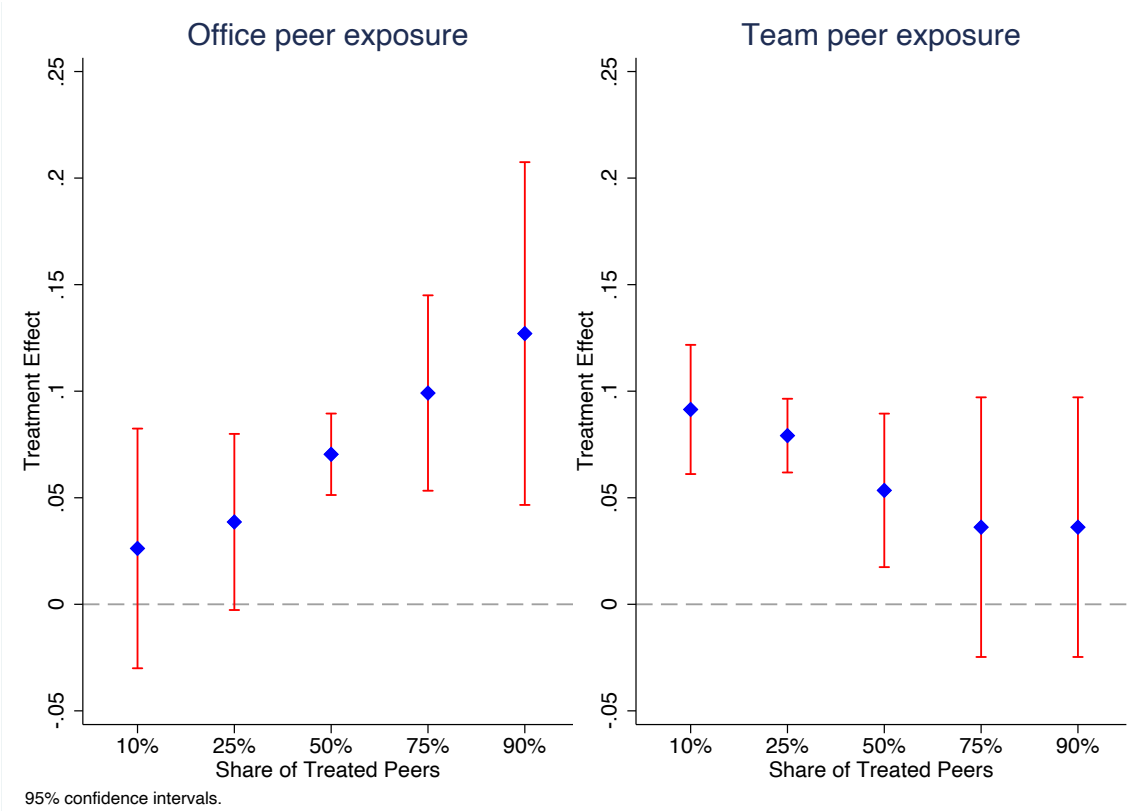
Notes: *The figure reports ITT estimates on the hazard of first-time volunteering. The outcome is an indicator equal to one if an employee who has not yet volunteered by semester $\tau - 1$ volunteers in semester τ. Estimates are obtained from regressions of this outcome on treatment interacted with semester indicators (semesters 3–6), including strata fixed effects. The sample is restricted to employees "at risk of first adoption" in each semester. Standard errors are clustered at the office level. Vertical bars represent 95% confidence intervals.*

Spillovers To examine whether the effect of the encouragement depends on exposure to treated peers, we construct leave-one-out measures of peer treatment intensity at the office and team level. For each employee i in stratum s, we compute the share of treated coworkers in her office (or team), excluding herself. We estimate

$$y_{i,s} = \alpha + \beta T_i + \gamma S_i + \pi \left(T_i \times S_i\right) + \lambda_s + \varepsilon_{i,s}, \tag{9}$$

where T_i is the individual treatment indicator, S_i is the leave-one-out share of treated peers, and λ_s are strata fixed effects. The interaction coefficient π captures whether the treatment effect varies with peer exposure. Appendix Table B9 reports the estimates.

Figure 8: Volunteering Spillovers



Notes: *The figure reports heterogeneous treatment effects on volunteering by peer exposure. Peer exposure is defined as the leave-one-out share of treated peers in an individual's office (left panel) or team (right panel). For each percentile (10th, 25th, 50th, 75th, 90th) of the peer exposure distribution, we estimate a linear model of volunteering on individual treatment, the share of treated peers, and their interaction, including randomization strata fixed effects and clustering standard errors at the office level. The plotted coefficients correspond to the implied treatment effect evaluated at each percentile of peer exposure. Vertical bars represent 95% confidence intervals.*

These results indicate that spillovers operate primarily at the office level. The individual treatment effect increases with the share of treated coworkers in the office, consistent with strategic complementarities: when more colleagues are exposed to the initiative, volunteering becomes more socially reinforced and individually attractive. By contrast, spillovers at the team level are much weaker and, if anything, they go in the opposite direction.

6 Main Results

This section presents the main results for on-the-job outcomes, turnover, and disengagement, described in Section 4. Section 6.1 discusses time varying outcomes estimated in the

panel. Section 6.2 presents results for cross-sectional outcomes.

6.1 On-the-job Performance

In the panel, for each unit $u \in \{i, t, o\}$, stratum s, and semester $\tau \geq 3$, we estimate

$$y_{u,\tau,s}^{\text{post}} = \alpha + \beta T_u + \gamma y_u^{\text{pre}} + \lambda_s + \delta_\tau + \varepsilon_{u,\tau,s}, \tag{10}$$

where $y_{u,\tau,s}^{\text{post}}$ is the post-treatment outcome, y_u^{pre} is the baseline performance/values rating, T_u is treatment assignment, λ_s are strata fixed effects, and δ_τ are semester fixed effects. At the team and office level, all variables are shares. Standard errors are clustered by office identifier.

Table 2 reports ITT estimates from equation (10). The first two rows use the 1–4 categorical ratings for performance and values as outcomes, while the last two rows report effects on the "Low Performance" and "Low Values" indicators. In addition to conventional p-values (in brackets), the table reports List-adjusted p-values (in braces) correcting for multiple hypothesis testing across the four outcomes within each level of aggregation.

Concerning the categorical ratings, the estimates show that treatment raises both performance and values at all three levels: coefficients are statistically significant at the 1% level for individuals, 5% for offices, and 10% for teams. Effect sizes are small at all levels, even though the point estimates become somewhat larger when we aggregate to teams and offices. This is expected and is consistent with the fact that line managers rate employees primarily relative to their peers within the scorecard framework, which makes large average shifts in these categorical ratings unlikely.

For Low Performance, the ITT at the individual level is -0.0038 on a control mean of 0.0344, an 11% reduction ($p < 5\%$). At the team level, the coefficient is -0.0090 on a pre treatment mean of 0.0389 ($p < 10\%$). At the office level, the ITT is -0.0205 on a pre treatment mean of 0.0787, roughly a 26% reduction for an office moving from 0% to 100% treated ($p < 5\%$). For Low Values, the coefficient magnitudes are similar, albeit not statistically significant at conventional levels.

Robustness In Table B11, we drop all employees who, at any point during the panel, are based in an office outside our baseline sample.[22] In Table B12, we restrict the sample to employees who respond to the employee survey at least once. In both cases, the ITT estimates are of same sign and very similar in magnitude and precision to those in Table 2.

Additionally, Table B13 implements a PDS LASSO specification following Belloni et al. (2014). We allow for a rich set of baseline covariates and their interactions.[23] At the office

[22] Because the panel spans six semesters, some employees who are initially in a baseline office subsequently transfer to offices not included in the baseline sample.

[23] The following HR-based covariates are included in the LASSO selection procedure: function, managerial status, seniority, full-time employment status, performance, values, semester fixed effects, and strata fixed effects.

level, the ITT coefficients in Table B13 remain similar to those in Table 2: performance and values continue to be positive and precisely estimated, and the reductions in low performance and low values are of comparable magnitude and statistically significant at the 5% level. At the team and individual levels, point estimates retain the same sign as in the baseline specification but are no longer statistically significant, with the exception of the treatment effect on individual-level values.

Finally, we address multiple hypothesis testing following List et al. (2019). Table B14 reproduces Table 2 using p-values adjusted separately within each level of aggregation (four outcomes per level). Table B15 instead reports p-values adjusted jointly across all twelve hypotheses (three levels × four outcomes). Conceptually, the within-level adjustment treats each level of aggregation as a distinct economic margin—individual, team, and office—reflecting that these correspond to different equilibrium objects in the model. The joint adjustment, by contrast, treats all outcomes across all levels as belonging to a single family of hypotheses, thereby imposing a stricter standard that accounts for dependence across both outcomes and aggregation levels. The results remain largely robust to both corrections. All coefficients continue to be statistically significant, with two exceptions: team-level low performance, and office-level low performance under the within-level adjustment in Table B14 ($p \approx 0.10$).

6.2 Retention and Disengagement

In the cross-section, for each unit $u \in \{i, t, o\}$ and stratum s, we estimate

$$y_{u,s}^{\text{post}} = \alpha + \beta T_u + \gamma y_u^{\text{pre}} + \lambda_s + \varepsilon_{u,s}, \tag{11}$$

where $y_{u,s}^{\text{post}}$ is the post-treatment outcome, y_u^{pre} is the corresponding baseline value of the same outcome, T_u is treatment assignment, and λ_s are strata fixed effects. At the team and office level, all variables are shares. Standard errors are clustered at the office level for individual and team regressions.[24]

Table 3 reports ITT estimates from these regressions. Concerning turnover, at the individual level, treated employees are 1.95 percentage points less likely to leave the bank over the post-treatment period, roughly a 13% reduction ($p<0.10$). We do not find any effect at the team level. At the office level, the ITT estimate is -0.0607 relative to a sample mean in the post-treatment period of 0.1261 ($p<0.01$).

The bottom part of the table looks at the measures of disengagement discussed in Section 4. We find significant decreases at all levels of aggregation. At the individual level, for the performance based measure, treated employees are 2.07 p.p. less likely to be classified as quiet quitters, approximately a 5% reduction. For the values based indicator, we find a reduction of 1.52 p.p., or an 8% reduction. Finally, for the combined "performance

[24] For turnover at the individual level, we do not observe a baseline value, so specifications with turnover as the outcome omit y_u^{pre}.

Table 2: On-the-job Performance

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_{preT}}(SD)$	N	T(SE)[P]	$\overline{Y_{preT}}(SD)$	N
Performance	0.0170***	2.7031	105733	0.0307*	2.7140	14358	0.0866***	2.5334	2716
	(0.0063)	(0.7919)		(0.0168)	(0.6015)		(0.0322)	(0.3717)	
	[0.0069]			[0.0683]			[0.0073]		
Values	0.0229***	3.0550	105165	0.0216*	3.1077	14331	0.0572**	3.0375	2716
	(0.0064)	(0.5318)		(0.0111)	(0.4077)		(0.0236)	(0.2414)	
	[0.0004]			[0.0512]			[0.0155]		
Low Performance	-0.0038**	0.0344	105733	-0.0090*	0.0389	14358	-0.0205**	0.0787	2716
	(0.0016)	(0.1822)		(0.0052)	(0.1450)		(0.0100)	(0.1074)	
	[0.0191]			[0.0828]			[0.0422]		
Low Values	-0.0070	0.1055	105165	-0.0088	0.0872	14331	-0.0219	0.1336	2716
	(0.0049)	(0.3072)		(0.0087)	(0.2097)		(0.0155)	(0.1460)	
	[0.1557]			[0.3107]			[0.1575]		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. All specifications include semester fixed effects and randomization strata fixed effects and control for baseline values of the outcome (measured in semester 2; if unavailable, semester 1). Control group means and standard deviations at the individual level are computed over the post-treatment period (timeID $\geq$ 3). At the team and office levels, reported means and standard deviations correspond to pre-treatment averages. Low performance and low values are indicator variables. Standard errors are clustered at the office level and reported in parentheses. Significance stars are based on conventional cluster-robust p-values. * p<0.10, ** p<0.05, *** p<0.01.*

or values" measure, we find a 1.39 p.p. reduction. At higher levels of aggregation the effects remain negative and precisely estimated, with coefficient magnitudes larger at the team level and four-to-five times as large at the office level ($p<0.05$ at all levels of aggregation and for all outcomes, except for the combined measure at the team level where $p \approx 0.053$).

Robustness In Table B16, we exclude employees who at any point transfer to offices outside the baseline sample. The estimates are very similar to those in Table 3: the office-level reduction in exit remains large and statistically significant ($p < 5\%$), and the negative effects on quiet quitting (performance, values, and the combined measure) remain negative and of similar magnitude across all three levels of aggregation. In Table B17, where we restrict the sample to employees who respond to the survey at least once, the office-level effects on exit and quiet quitting continue to be statistically significant. By contrast, exit at the individual level and quiet quitting in values are no longer statistically significant.

Turning to the PDS LASSO specification in Table B18, the office-level treatment effects on exit and all measures of quiet quitting remain statistically significant and similar in

Table 3: Turnover and Disengagement

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N
Exiter	-0.0195*	0.1466	32456	-0.0007	0.0348	5317	-0.0607***	0.1261	713
	(0.0109)	(0.3537)		(0.0102)	(0.0922)		(0.0207)	(0.1211)	
	[0.0757]			[0.9482]			[0.0035]		
Disengagement: Quiet Quitting									
Performance	-0.0207***	0.4527	31174	-0.0257**	2.7087	5226	-0.0857***	0.5190	713
	(0.0051)	(0.4978)		(0.0110)	(0.6013)		(0.0274)	(0.2027)	
	[0.0000]			[0.0195]			[0.0019]		
Values	-0.0152**	0.1809	31058	-0.0276***	3.1072	5220	-0.0516**	0.2014	713
	(0.0066)	(0.3850)		(0.0100)	(0.4100)		(0.0232)	(0.1526)	
	[0.0212]			[0.0061]			[0.0262]		
Performance or Values	-0.0139***	0.4567	31043	-0.0206*	5.8162	5217	-0.0800***	0.5220	713
	(0.0052)	(0.4981)		(0.0106)	(0.8989)		(0.0272)	(0.2029)	
	[0.0082]			[0.0528]			[0.0034]		

Notes: *SE clustered by office for individual and team level regressions; heteroskedasticity robust SE for office-level regressions. Strata FE included. $\bar{Y}$ is the control mean (individual level) or the sample mean (team/office level, where treatment is a share). Controls for baseline values in qq-values, baseline performance in qq-performance and baseline values + performance in qq-performance-values. Control for baseline turnover in team and office level regressions. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.*

magnitude to the baseline estimates.[25] In contrast, at the individual and team levels, coefficients remain of similar sign and magnitude but are mostly no longer statistically significant.

Finally, Tables B19 and B20 address multiple hypothesis testing following List et al. (2019), implementing corrections across four hypotheses per level and across all twelve hypotheses jointly, respectively. After correcting for multiple testing, the treatment effect on exit at the individual level is not statistically significant ($p \approx 15\%$).

Interpreting the performance and turnover results in light of the model. The increasing magnitude of effects at higher levels of aggregation is consistent with the strategic complementarities embedded in the model and with Testable Implication 2. In the framework, effort depends positively on the office-level share of prosocial types, denoted p_t, so that increases in cultural alignment raise the marginal return to discretionary effort for all workers. Individual treatment generates a direct first-round effect on the treated employee. By contrast, greater treatment intensity within a team or office shifts the equilibrium level

[25] LASSO variables include function, managerial status, seniority, baseline exiter indicator, full-time-equivalent status, baseline performance and values ratings, age, and strata/semester indicators, together with their interactions.

of p_t, amplifying effort through complementarities (that is, when $\gamma > 0$). This produces an equilibrium multiplier: as treatment density rises, cultural alignment strengthens, leading to broader increases in effort and larger aggregate performance gains. The same mechanism applies to retention. Because the continuation value $V_i(p_t)$ is increasing in culture, shifts in p_t reduce exit probabilities more strongly in offices with higher treatment intensity, generating amplified turnover effects at higher levels of aggregation, in line with the model's second empirical prediction.

6.3 Change in behavior or change in composition?

Treatment improves on-the-job performance while reducing turnover. How do these two results interact with each other? Conceptually, the volunteering initiative can shape workplace culture through two distinct channels. The first is *selection* (change in composition): the policy makes the bank's prosocial values more salient, thus employees who do not share those values are encouraged and/or prefer to leave. The second is *effort* (change in behavior): among employees who stay, the policy may encourage them to internalize these values and, as their personal values become more closely aligned with the bank's, exert more effort. To distinguish these two channels, we proceed in three steps.

First, we abstract from treatment assignment and study how outcomes evolve between the pre- and post–treatment periods for the workforce as a whole. Exploiting the panel structure of the data, we follow Lazear (2000) and estimate two specifications. Let $y_{i\tau s}$ denote an HR outcome for individual i in stratum s and semester τ, and let post_τ be an indicator for post–treatment semesters. We estimate

$$y_{i\tau s} = \alpha + \theta^{\text{tot}} \, \text{post}_\tau + \lambda_s + \delta_\tau + \varepsilon_{i\tau s}, \tag{12}$$

where λ_s are strata fixed effects and δ_τ are semester fixed effects. The coefficient θ^{tot} captures the overall pre–post change in the observed sample, combining both selection and effort. Similarly, we estimate

$$y_{i\tau s} = \alpha_i + \theta^{\text{eff}} \, \text{post}_\tau + \delta_\tau + \varepsilon_{i\tau s}, \tag{13}$$

where α_i are individual fixed effects. In this specification, θ^{eff} measures the average within–person pre–post change (effort). Comparing θ^{tot} and θ^{eff} allows us to measure how much of the total change in outcomes is driven by within-person changes rather than composition.

Column (1) in Table 4 reports $\hat{\theta}^{\text{tot}}$ from equation (12), Column (2) reports $\hat{\theta}^{\text{eff}}$ from equation (13), and Column (3) reports the share of the total pre–post change attributable to selection (in percent). For performance, the overall pre–post increase is 0.0724 (col. 1), while the within-person increase is 0.0616 (col. 2), implying that 15% of the improvement is due to selection and 85% to effort. For low performance, the total decline is -0.0135 relative to a within-person decline of -0.0085, so selection accounts for roughly 37% of the

reduction in the share of low performers. By contrast, for values and low values the selection shares are negative (-6.6% and -24.4%), indicating that composition changes actually work against the overall increase in values. Taken together, the results in Table 4 are consistent with most of the pre–post change in performance and values being driven by within-person changes rather than by differential selection through exit.

Table 4: Decomposition Pre-Post: Changes in Composition (Selection) and Behavior (Effort)

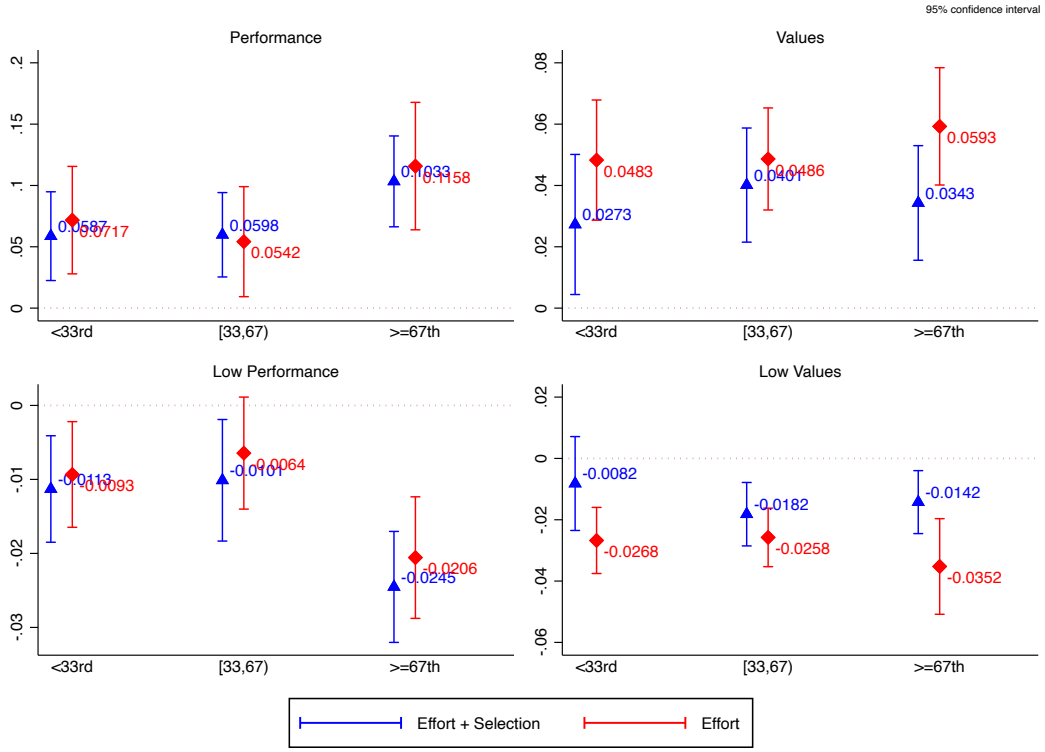
	(1)	(2)	(3)	(4)	(5)
	Post		Selection		
	Strata FE	Ind. FE	Share	$\overline{Y}$(SD)	N
Performance	0.0724***	0.0616***	15.0%	2.7237	169,306/168,113
	(0.0157)	(0.0214)		(0.7951)	
	[0.0000]	[0.0041]			
Values	0.0448***	0.0477***	-6.6%	3.0907	168,466/167,235
	(0.0079)	(0.0073)		(0.5365)	
	[0.0000]	[0.0000]			
Low Performance	-0.0135***	-0.0085**	36.7%	0.0345	169,306/168,113
	(0.0039)	(0.0037)		(0.1825)	
	[0.0006]	[0.0202]			
Low Values	-0.0212***	-0.0263***	-24.4%	0.0943	168,466/167,235
	(0.0049)	(0.0045)		(0.2922)	
	[0.0000]	[0.0000]			

Notes: *Each column reports estimates from regressions of the indicated outcome on a post-treatment indicator. Column (1) includes strata and semester fixed effects and captures the total pre–post change (effort + selection). Column (2) includes individual and semester fixed effects and captures within-person changes (effort). Column (3) reports the share of the total effect attributable to selection, defined as* $1 - (2)/(1)$*, in percent. Standard errors are clustered at the office level and reported in parentheses; P-values in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

Second, we extend the decomposition to examine whether the effort–selection patterns discussed before vary with offices' treatment intensity. Let $q_o \in \{1, 2, 3\}$ index terciles of the office-level treatment share (bottom, middle, and top tercile, respectively). We estimate augmented versions of equations (12) and (13) that interact the post indicator with these tercile dummies, and we compare θ_g^{tot} and θ_g^{eff} by terciles $g = 1, 2, 3$.

Figure 9 summarizes the results. Across all four outcomes, offices with higher treatment shares exhibit larger pre–post increases in performance and values and larger declines in low performance and low values, consistent with the treatment effects documented in Section 6.1. Additionally, within each tercile the "effort + selection" (blue) and "effort" (red) estimates are very close, with overlapping confidence intervals, indicating that selection accounts for only a small share of the total change. Finally, across office treatment shares, the decomposition shows that selection matters primarily for performance, not for values.

Figure 9: Effort-Selection Decomposition by Tercile Office Treatment Share



Notes: *The figure reports pre–post changes in performance and values outcomes by terciles of office-level treatment share. For each outcome, we estimate specifications that decompose the total pre–post change ("Effort + Selection") and the within-person change ("Effort"). Offices are grouped into terciles based on their baseline share of treated employees (<33rd percentile, 33rd–67th percentile, ≥67th percentile). Blue markers correspond to total pre–post changes estimated with strata and semester fixed effects; red markers correspond to within-person changes estimated with individual and semester fixed effects. Standard errors are clustered at the office level. Vertical bars represent 95% confidence intervals.*

Third, we assess the extent to which the estimated treatment effects reflect genuine behavioral responses rather than changes in workforce composition. To do so, we re-estimate equation (10) on both the full post-treatment sample and a balanced panel that includes only employees who remain in the firm throughout the post-treatment period. Column (1) reproduces the individual-level ITT estimates from the unbalanced sample, matching Table 2. Column (2) reports the same specification on the balanced sample, thereby isolating within-person effects. Column (3) quantifies the contribution of compositional change by reporting the share of the baseline ITT estimates that is attributable to selection.

We find that selection contributes little to the treatment effects on performance and values: the selection shares are 1.5% and -5.5%, respectively. This indicates that the improvements in these outcomes as of the treatment are driven almost entirely by within-person changes, with compositional shifts playing at most a minor role and slightly attenuating the gains in values. By contrast, selection accounts for a larger share of the effect on low performance (14.1%), consistent with weaker performers being more likely to exit. Notably, restricting the sample to employees who remain throughout the post-treatment period (Column (2)) yields ITT estimates for performance and values that are nearly identical in magnitude and precision to those in the full, unbalanced sample.

In conclusion, Section 6.3 delivers three clear results. First, behavioral adjustment—not workforce sorting—is the primary channel through which the intervention operates. The improvements in performance and values are driven almost entirely by within-person changes, with selective retention playing only a minor role. Second, the main ITT estimates in Section 6.1 remain essentially unchanged when we hold composition fixed, confirming the effectiveness of our intervention in changing effort, in line with the theoretical predictions of our model. Third, selection operates asymmetrically across outcomes. For performance and low performance, the total pre–post change exceeds the within-person change, indicating that exit disproportionately removes weaker performers and thereby amplifies the gains. For values and low values, by contrast, the gap between total and within-person changes is small and slightly negative. This pattern is consistent with the firm's incentive structure.

Table 5: Decomposition ITT Estimates into Selection and Effort

	(1)	(2)	(3)	(4)	(5)
	Treat		Selection		
	Unbalanced	Balanced	Share	$\overline{Y_C}(SD)$	N
Performance	0.0170***	0.0168**	1.5%	2.6758	105,733/92,274
				(0.7988)	
	(0.0063)	(0.0066)			
	[0.0069]	[0.0116]			
Values	0.0229***	0.0242***	-5.5%	3.0611	105,165/91,266
				(0.5471)	
	(0.0064)	(0.0060)			
	[0.0004]	[0.0001]			
Low Performance	-0.0038**	-0.0033**	14.1%	0.0408	105,733/92,274
				(0.1979)	
	(0.0016)	(0.0016)			
	[0.0191]	[0.0442]			
Low Values	-0.0070	-0.0067	4.0%	0.1115	105,165/91,266
				(0.3148)	
	(0.0049)	(0.0045)			
	[0.1557]	[0.1392]			

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment status. All specifications include semester and strata fixed effects. Regressions with performance [values] and low performance [values] as outcomes additionally control for the corresponding baseline measure (semester 2; if unavailable, semester 1). Column (1) reports ITT estimates on the full (unbalanced) post-treatment sample. Column (2) reports ITT estimates on a balanced panel restricted to employees who remain throughout the post-treatment period. Column (3) reports the share of the absolute ITT in Column (1) attributable to selection, computed as* $100 \times \big(1 - |(2)/(1)|\big)$*. Column (4) reports the control group mean and standard deviation. Standard errors are clustered at the office level and reported in parentheses; P-values in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

7 Culture

We have shown that the intervention reduces turnover and quiet quitting and improves on-the-job outcomes (Sections 6.1 and 6.2), and that these effects operate mainly through changes in effort among employees who stay rather than through selective exit from the firm (Section 6.3). Moreover, performance effects are substantially larger at higher levels of aggregation, consistent with the strategic complementarities highlighted in the model. What explains this behavioral response and the observed amplification at the office level? In the framework, such complementarities arise when the intervention shifts the equilibrium share of prosocial types, increasing the marginal return to effort for all workers. In this section, we provide evidence that a central mechanism operates through changes in employees' perceptions of banking culture: the initiative signals a credible move toward a more prosocial organizational culture.

7.1 Individual-level Changes in Cultural Fit

Table 6 reports individual level estimates for the standardized culture fit index, the PCA index, the social responsibility measure, and the 16 individual survey questions composing the indexes (see Section B.3 for list of questions and Section 4 for the description of the variables). We estimate:

(1) Baseline post-treatment specification:

$$y_{is\tau} = \alpha + \beta^{(1)} T_i + \lambda_s + \delta_\tau + \varepsilon_{is\tau}, \qquad \tau \geq 3, \tag{14}$$

where T_i is the individual treatment indicator, λ_s are strata fixed effects, and δ_τ are semester fixed effects.

(2) We add baseline controls for the outcome, interacted with post-treatment semester dummies. Let y_i^{pre} be the baseline value of $y_{is\tau}$; we estimate:

$$y_{is\tau} = \alpha + \beta^{(2)} T_i + \phi\, y_i^{\text{pre}} + \sum_{\tau \geq 3} \kappa_\tau \big(y_i^{\text{pre}} \cdot \mathbf{1}\{\tau\} \big) + \lambda_s + \delta_\tau + \varepsilon_{is\tau}, \qquad \tau \geq 3. \tag{15}$$

(3) Same as (2), but weighted by stabilized inverse-probability weights (IPW) constructed from observable baseline characteristics to adjust for differential post-treatment survey response. This reweights the respondent sample to reflect the baseline population, to the extent that selection operates through observables.[26]

[26] In particular, let R_i be an indicator equal to one if employee i responds to the culture survey at least once in the post-treatment period, and zero otherwise. We estimate a logistic regression of R_i on a vector of baseline covariates X_i, including a big-city indicator, full-time equivalent status, grade dummies, baseline performance and values ratings, function dummies, tenure and age categories, baseline team and office size (in logs), strata fixed effects, and indicators for missing values. Let $\hat{p}_i = \widehat{\Pr}(R_i = 1 \mid X_i, s(i))$ denote the fitted probability of response. For each stratum s, let $\bar{p}_s$ denote the observed mean response rate. We define the stabilized weight as $w_i^{\text{IPW}} = \bar{p}_{s(i)} / \hat{p}_i$. These weights upweight respondents with lower predicted response probabilities within strata, thereby adjusting for differential survey participation driven by observable baseline characteristics.

(4) Individual fixed effects DiD on employees observed at least once before and once after treatment:

$$y_{is\tau} = \alpha_i + \delta_\tau + \beta^{(4)}\big(T_i \times \text{post}_\tau\big) + \varepsilon_{is\tau}, \tag{16}$$

where α_i are individual fixed effects and post_τ is an indicator for $\tau \geq 3$.

For the standardized culture index, column (1), which does not condition on pre-treatment values, reports an ITT of 0.113 standard deviations, $p < 0.01$. In column (2), adding controls for the pre-treatment outcome interacted with time identifier absorbs part of this effect, and the ITT falls to 0.070 sd, $p < 0.01$. The weighted regression in column (3) delivers virtually the same estimate, indicating that although post-treatment survey response is not balanced between treatment and control, differential non-response does not affect ITT estimates once we condition on pre-treatment culture. The individual level DiD in column (4) shows a very similar effect of 0.068 sd, $p < 0.05$. The PCA index shows a very similar pattern.

For the social responsibility measure, the ITT is 0.121 sd (p<0.01) in column (1), 0.11 sd (p<0.01) in columns (2) and (3), so pre-treatment controls attenuate the ITT much less than for the index, while the DiD estimate in column (4) is similar to the one for the standardized index at about 0.07 sd (p<0.05).

When we look at the individual components of the index, the ITT estimates in columns (1)–(3) are all positive, precisely estimated, and range between 0.04 and 0.12 standard deviations in column (1), and between 0.04 and 0.09 in columns (2) and (3). Once we net out selection in column (4), most coefficients remain positive and significant, but some no longer are, indicating that changes in the composition of respondents matter more for some measures than for others.

Most notably, across columns (1)–(4), the three questions that capture more closely the relationship with customers show some of the largest and most precisely estimated effects.[27] In column (1) coefficients range from 0.092 to 0.123 sd, in column (2) from 0.083 to 0.091 sd, in column (3) from 0.083 to 0.090 sd, and in column (4) from 0.040 to 0.071 sd. These estimates are statistically significant at $p < 0.01$ in columns (1)–(3) and at $p < 0.05$ in column (4). This evidence suggests that realigning workers' values toward prosociality not only improves overall cultural fit with the bank, but also strengthens bankers' connection with the customers they serve. These results line up with HTE in Appendix Table B24, which splits the sample between customer-facing employees and not, and shows that changes in the culture fit index as of treatment are concentrated among customer-facing employees.

[27] "Bank treats customers fairly" ("[Bank Name] is genuine in its commitment to delivering fair outcomes for its customers"); "Bank meets customers' needs" ("I feel confident that customers receive products and services that meet their needs."); "Can challenge customer policies" ("I can challenge policies or procedures that I think will be bad for customers.").

Table 6: Culture Fit

	Post-treatment regressions (1)–(3)									DiD (4)		
	(1) FE			(2) FE + Baseline			(3) FE + Baseline + IPW			(4) Indiv. FE		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_C}(SD)$	N	T#Post(SE)[P]	$\overline{Y_C}(SD)$	N
Culture index (std.)	0.1134***	0.1152	29578	0.0704***	0.0777	20492	0.0700***	0.0777	20492	0.0682**	0.0774	39813
	(0.0317)	(0.9666)		(0.0200)	(0.9846)		(0.0203)	(0.9846)		(0.0271)	(0.9847)	
	[0.0004]			[0.0005]			[0.0006]			[0.0121]		
Culture index (PCA)	0.3227***	-0.0903	29578	0.2180***	-0.1921	18016	0.2167***	-0.1921	18016	0.1943**	-0.1929	31550
	(0.0906)	(2.7685)		(0.0570)	(2.8488)		(0.0579)	(2.8488)		(0.0771)	(2.8486)	
	[0.0004]			[0.0001]			[0.0002]			[0.0120]		
Social responsibility	0.1034***	3.7324	14292	0.0944***	3.7281	9980	0.0956***	3.7281	9980	0.0588**	3.7279	25288
	(0.0331)	(0.8062)		(0.0325)	(0.8205)		(0.0324)	(0.8205)		(0.0260)	(0.8203)	
	[0.0019]			[0.0038]			[0.0033]			[0.0239]		
Can be productive in my job	0.0622**	3.3027	29578	0.0754***	3.2641	20375	0.0737**	3.2641	20375	0.0877**	3.2640	39551
	(0.0302)	(1.0630)		(0.0282)	(1.0835)		(0.0285)	(1.0835)		(0.0359)	(1.0833)	
	[0.0398]			[0.0077]			[0.0100]			[0.0148]		
Can achieve career goals	0.0816***	3.3516	29578	0.0592**	3.3316	20104	0.0591**	3.3316	20104	0.0314	3.3317	38943
	(0.0278)	(1.0194)		(0.0230)	(1.0237)		(0.0236)	(1.0237)		(0.0344)	(1.0238)	
	[0.0034]			[0.0103]			[0.0123]			[0.3621]		
Positive impact strategy	0.0634**	3.3751	29578	0.0626***	3.3289	20492	0.0631***	3.3289	20492	0.0506	3.3287	39813
	(0.0276)	(0.8863)		(0.0220)	(0.9082)		(0.0219)	(0.9082)		(0.0324)	(0.9083)	
	[0.0218]			[0.0045]			[0.0042]			[0.1184]		
Confident about bank's future	0.0381*	3.6382	29578	0.0378**	3.6115	20492	0.0363**	3.6115	20492	0.0422	3.6115	39813
	(0.0223)	(0.9088)		(0.0163)	(0.9302)		(0.0165)	(0.9302)		(0.0325)	(0.9304)	
	[0.0877]			[0.0210]			[0.0284]			[0.1938]		
Trust senior leadership	0.0768**	3.4205	29578	0.0567***	3.3741	20169	0.0562***	3.3741	20169	0.0647***	3.3737	39085
	(0.0328)	(1.0012)		(0.0201)	(1.0330)		(0.0204)	(1.0330)		(0.0246)	(1.0329)	
	[0.0195]			[0.0050]			[0.0061]			[0.0086]		
Trust direct manager	0.0582***	4.0505	29578	0.0530***	4.0313	20169	0.0527***	4.0313	20169	0.0465**	4.0311	39085
	(0.0180)	(0.9746)		(0.0184)	(0.9782)		(0.0176)	(0.9782)		(0.0211)	(0.9783)	
	[0.0013]			[0.0040]			[0.0028]			[0.0280]		
Bank encourages speak-up	0.0892***	3.6736	29578	0.0773***	3.6533	20319	0.0757***	3.6533	20319	0.0573***	3.6530	39420
	(0.0254)	(0.9296)		(0.0225)	(0.9454)		(0.0229)	(0.9454)		(0.0185)	(0.9454)	
	[0.0005]			[0.0006]			[0.0010]			[0.0020]		
Can speak up on concerns	0.0726***	3.7885	29578	0.0538***	3.7770	20319	0.0513***	3.7770	20319	0.0166	3.7766	39420
	(0.0177)	(0.9098)		(0.0189)	(0.9223)		(0.0193)	(0.9223)		(0.0172)	(0.9222)	
	[0.0000]			[0.0045]			[0.0080]			[0.3362]		
Opinions are taken seriously	0.1052***	3.4661	29578	0.0844***	3.4547	20319	0.0810***	3.4547	20319	0.0674***	3.4543	39419
	(0.0239)	(1.0396)		(0.0222)	(1.0469)		(0.0225)	(1.0469)		(0.0187)	(1.0469)	
	[0.0000]			[0.0002]			[0.0003]			[0.0003]		
Recognised for good work	0.0822***	3.6315	29578	0.0604***	3.6114	18021	0.0596***	3.6114	18021	0.0259	3.6112	31559
	(0.0185)	(0.9491)		(0.0160)	(0.9702)		(0.0158)	(0.9702)		(0.0231)	(0.9701)	
	[0.0000]			[0.0002]			[0.0002]			[0.2632]		
Recognised for living values	0.0794***	3.7157	29578	0.0501***	3.6983	18021	0.0502***	3.6983	18021	0.0146	3.6980	31559
	(0.0176)	(0.9478)		(0.0172)	(0.9639)		(0.0165)	(0.9639)		(0.0249)	(0.9639)	
	[0.0000]			[0.0037]			[0.0025]			[0.5576]		
Bank treats customers fairly	0.0721***	3.9731	29578	0.0649***	3.9584	20061	0.0664***	3.9584	20061	0.0478**	3.9582	38847
	(0.0245)	(0.7191)		(0.0215)	(0.7282)		(0.0217)	(0.7282)		(0.0239)	(0.7282)	
	[0.0034]			[0.0026]			[0.0023]			[0.0456]		
Bank meets customers' needs	0.0793***	3.8478	29578	0.0735***	3.8272	20061	0.0733***	3.8272	20061	0.0578***	3.8275	38847
	(0.0263)	(0.7413)		(0.0214)	(0.7555)		(0.0217)	(0.7555)		(0.0178)	(0.7555)	
	[0.0026]			[0.0006]			[0.0008]			[0.0012]		
Can challenge customer policies	0.1212***	3.6174	29578	0.0813***	3.6220	20061	0.0816***	3.6220	20061	0.0398**	3.6221	38847
	(0.0318)	(0.8848)		(0.0234)	(0.8973)		(0.0237)	(0.8973)		(0.0189)	(0.8970)	
	[0.0001]			[0.0005]			[0.0006]			[0.0354]		
Proud to work for bank	0.0612**	3.7490	29578	0.0500***	3.6988	19968	0.0489**	3.6988	19968	0.0470*	3.6986	38607
	(0.0274)	(0.8867)		(0.0188)	(0.9044)		(0.0193)	(0.9044)		(0.0276)	(0.9044)	
	[0.0259]			[0.0079]			[0.0115]			[0.0890]		
Feel valued by bank	0.0862***	3.4113	29578	0.0703***	3.3770	19968	0.0701***	3.3770	19968	0.0583	3.3766	38607
	(0.0282)	(1.0296)		(0.0258)	(1.0509)		(0.0264)	(1.0509)		(0.0411)	(1.0508)	
	[0.0023]			[0.0065]			[0.0080]			[0.1565]		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. Specifications (1)–(3) use post-treatment observations only (timeID $\geq$ 3) and include semester and strata fixed effects. Specifications (2) and (3) additionally control for the corresponding baseline outcome (measured in semester 2; if unavailable, semester 1) interacted with semester indicators. Specification (3) further applies stabilized inverse probability weights (IPW), constructed from a logit model of post-treatment survey response on pre-treatment HR and survey covariates (including indicators for missing values) and strata, and stabilized within strata. The baseline covariates include a big city indicator, full-time equivalent status, performance and values, grade and function dummies, tenure and age category dummies, the manager indicator, team and office size, and strata fixed effects. These weights re-balance the post-treatment sample to resemble the baseline population conditional on observables. Specification (4) is a difference-in-differences regression that restricts the sample to employees observed at least once before and once after treatment and includes individual and semester fixed effects; it reports the coefficient on the interaction between the treatment indicator and a post-treatment dummy. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

Robustness Appendix Table B21 reports two additional ITT specifications. The first re-estimates equation (15) restricting the sample to employees who are never observed in an office outside the initial sample. The second is a PDS LASSO specification following Belloni et al. (2014), which includes a broad set of pre-treatment covariates.[28] In both cases, the ITT effects are very similar to those in Table 6 and, if anything, slightly larger and more precisely estimated. For the standardized culture index, the estimated ITT rises from about 0.07 in the benchmark specification (column (2) of Table 6) to about 0.08 in the restricted sample and with PDS LASSO (all $p < 0.01$).

Additionally, Appendix Table B22 re-estimates equation (15) adding office and team fixed effects. These fixed effects slightly attenuate the ITT estimates. For the culture index, the coefficient falls from about 0.07 in the benchmark specification to about 0.06 with office fixed effects and about 0.05 with team fixed effects ($p< 0.01$). A similar pattern of mild attenuation appears for the individual survey questions, but the sign, magnitude, and statistical significance of the coefficients are largely unchanged.

Finally, Appendix Table B23 replicates columns (2) and (4) of Table 6 using List-adjusted p-values that correct for multiple hypothesis testing across the full set of cultural outcomes. The difference-in-differences estimates remain robust to this correction, with the main index measures retaining statistical significance at conventional levels. For the FE + baseline specification, the adjustment modestly attenuates precision: the standardized culture index yields a List-adjusted p-value of 10.7%, while the PCA index remains significant at the 5% level. For individual survey items, some coefficients lose significance after correction, as expected given the number of outcomes considered, but adjusted p-values remain below 15% in all cases.

7.2 Office Level Cultural Change

Consistent with the model in Section 2, we discuss whether the intervention shifted the distribution of cultural fit at the office level. The model predicts that a credibility shock can move offices toward a higher cultural equilibrium and, along the adjustment path, towards lower dispersion of beliefs within offices.

7.2.1 Descriptive Evidence

Figure 10 plots kernel densities of the office-level Culture Fit Index before and after the intervention. The post-treatment distribution shifts clearly to the right. A two-sample Kolmogorov–Smirnov test strongly rejects equality of the pre- and post-treatment distributions ($D = 0.361$, $p < 0.001$).

[28] LASSO variables include function, managerial status, age and seniority categories, full-time employment status, baseline performance and values ratings, semester fixed effects, and strata fixed effects, together with their interactions.

Figure 10: Office-Level Culture Fit Index (Pre vs. Post)

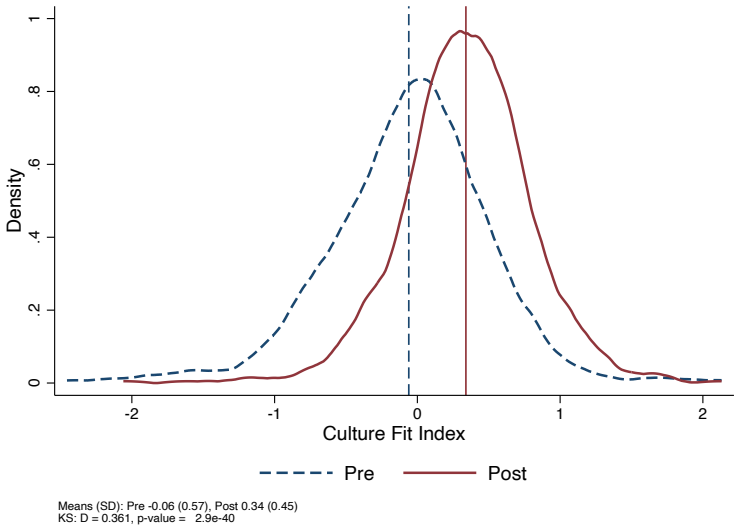


Figure 11 reports the distribution of the office-level share of employees who "strongly agree" or "agree" that "Where I work, people have a strong sense of responsibility to society," our social responsibility measure. The mean increases from 0.76 to 0.78, and the distribution shifts modestly to the right. Although the shift is more limited than for the overall index, the Kolmogorov–Smirnov test again rejects equality of distributions ($D = 0.099$, $p = 0.002$), indicating a statistically significant change in perceived social responsibility across offices.

Figure 11: Office-Level Social Responsibility (Pre vs. Post)

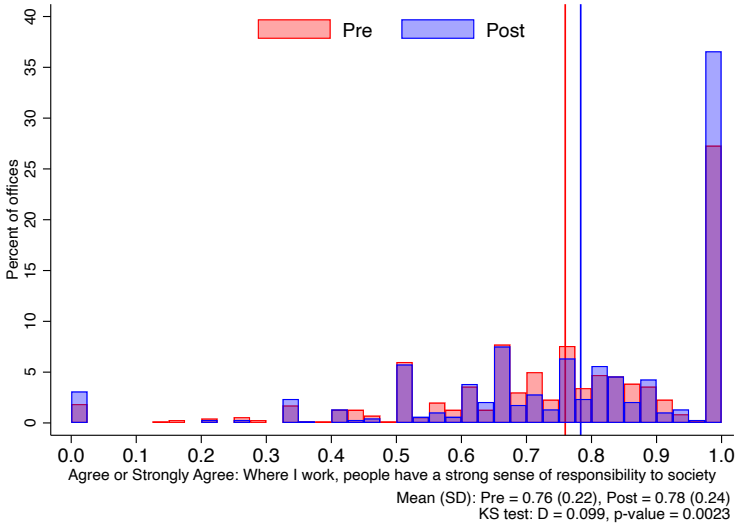


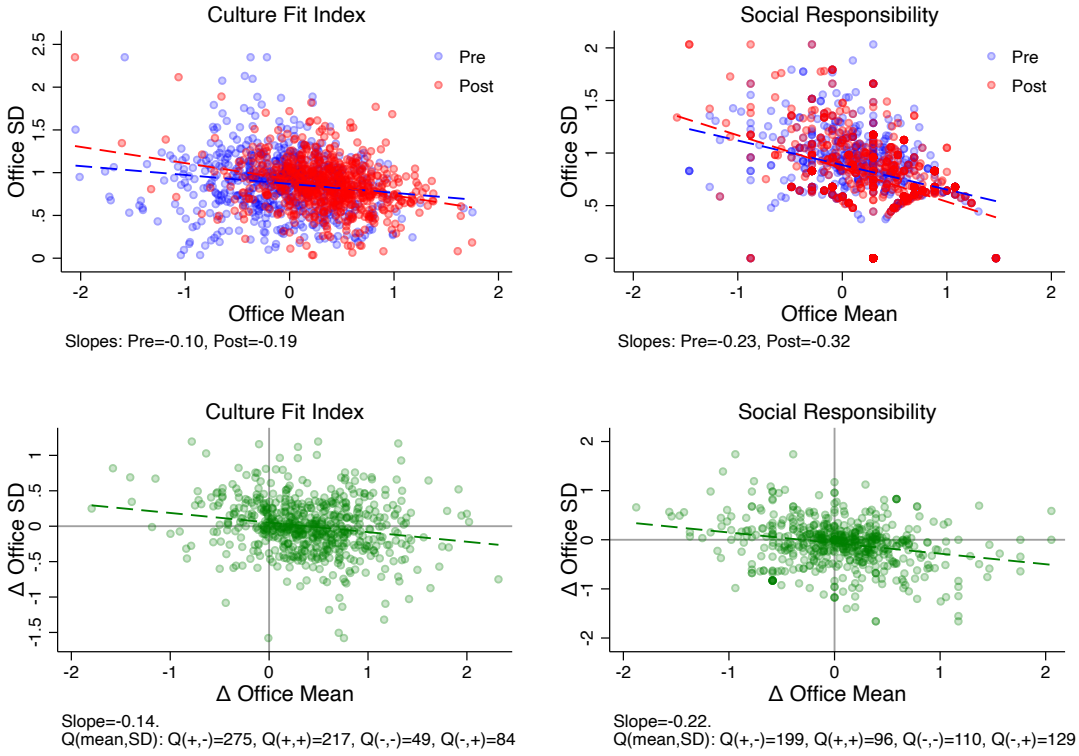
Figure 12 provides visual evidence on the joint evolution of the mean and dispersion of cultural beliefs across offices, directly linking to the convergence mechanism described in Section 2. In the framework, the volunteering initiative operates as a credibility shock that shifts the cultural updating function upward. When conformist forces are sufficiently strong, offices transition toward a higher steady-state share of prosocial types. Along this path, the model predicts not only an increase in the mean level of cultural alignment, but also a compression of within-office heterogeneity.

The top row of Figure 12 plots, for each office, the within-office standard deviation against the office mean, separately for the pre- and post-treatment periods. For both the Culture Fit Index (left) and Social Responsibility (right), the post-treatment cloud shifts to the right, reflecting the upward movement in average beliefs documented above. At the same time, the relationship between the mean and dispersion becomes more negative (Index: from -0.10, $p < 1\%$ pre to -0.19, $p < 1\%$ post; SR: from -0.23, $p < 1\%$ to -0.32, $p < 1\%$). Thus, in the post period, offices with higher cultural alignment tend to display lower within-office dispersion. This pattern is consistent with movement toward a high-culture equilibrium in which beliefs are both stronger and more homogeneous.

The bottom row reinforces this interpretation by plotting changes in dispersion against changes in the mean. For the Culture Fit Index (bottom left), the slope is negative (-0.14, $p < 1\%$), indicating that offices with larger increases in average cultural fit also experience larger reductions in dispersion. The quadrant breakdown underscores this result: 275 offices lie in the upper-left quadrant (Q(+,-)), exhibiting both rising means and declining dispersion, compared to 217 in Q(+,+) and far fewer in the quadrants associated with declining means. Cultural improvement is therefore frequently accompanied by convergence rather than polarization. A similar pattern holds for Social Responsibility (bottom right), where the slope is even steeper (-0.22, $p < 1\%$), and 199 offices combine rising means with falling dispersion (Q(+,-)).

Taken together, the figure aligns closely with the model's predictions. The intervention shifts offices toward higher average cultural alignment and, in many cases, toward a tighter distribution of beliefs. This joint movement in means and dispersion is precisely what one would expect if the policy functions as a credibility shock that moves offices along a convergence path toward a higher steady-state cultural equilibrium.

Figure 12: Convergence: Office-Level Scatter Plot



Notes. *The top panels plot, separately for the pre- and post-treatment periods, the relationship between the office-level mean of beliefs and the within-office standard deviation. The bottom panels plot the change in dispersion (post minus pre) against the change in the office-level mean. Each dot represents an office, and dashed lines denote linear fits.*

In the bottom panels, offices are partitioned into four quadrants according to the sign of changes in the mean and dispersion. ***Q1:*** $\Delta Mean > 0$, $\Delta SD < 0$ *— convergence toward a higher cultural alignment;* ***Q2:*** $\Delta Mean > 0$, $\Delta SD > 0$ *— higher mean with increasing dispersion;* ***Q3:*** $\Delta Mean < 0$, $\Delta SD < 0$ *— convergence toward a lower mean;* ***Q4:*** $\Delta Mean < 0$, $\Delta SD > 0$ *— declining mean with increasing dispersion.*

7.2.2 Causal Estimates

A remark on sample attrition in survey responses When analyzing culture at the office level, measurement precision varies across offices. Our outcomes are computed from survey respondents rather than from the full population of employees. Because response rates differ across offices and across years, some offices—especially smaller ones with low response rates—provide noisier estimates of underlying culture. In addition, the set of respondents changes over time, so post-treatment office means are based on different effective sample sizes. To address this concern, we proceed in two steps. First, we collapse the post-treatment period so that, for the Culture Fit Index, we average across the two post waves (for Social Responsibility we observe only one post-treatment measure). Second, we report both unweighted and weighted OLS specifications.

The weights are constructed using baseline survey data only and are described in detail in Section B.6, which also documents their orthogonality to treatment intensity, ruling out mechanical inflation of treatment effects. Concretely, we weight offices by (i) the baseline number of respondents, (ii) the baseline number of respondents multiplied by one minus the baseline non-response rate, and (iii) squared versions of these measures. These weights give greater influence to offices whose baseline culture is measured using more information and with lower non-response. Conceptually, this serves two purposes. First, it approximates feasible GLS: under standard sampling logic, the variance of an office mean decreases with the number of respondents, so offices with larger baseline samples provide more precise estimates. Second, it limits the influence of small offices with noisy means, reducing the risk that idiosyncratic composition changes in small samples drive the results.

Office-Level Mean Shift in Cultural Beliefs In Table 7, we estimate the following office–level specification separately for the Culture Fit Index and for Social Responsibility:

$$y_{o,s}^{\text{post}} = \alpha + \beta T_o + \gamma y_o^{\text{pre}} + \kappa \, \text{miss}_o^{\text{pre}} + \lambda_s + \varepsilon_{o,s}, \tag{17}$$

where $y_{o,s}^{\text{post}}$ denotes the post-treatment office-level outcome for office o in stratum s, T_o is the office treatment share, y_o^{pre} is the baseline office-level outcome (averaged over semesters 1–2), $\text{miss}_o^{\text{pre}}$ is the baseline share of missing survey responses in office o, and λ_s are randomization-strata fixed effects. The coefficient β in equation (17) captures whether greater treatment exposure increases the office-level mean of cultural beliefs, indicating, when positive, a shift toward a higher steady-state level in terms of perceived cultural fit or social responsibility. Columns (1)–(2) report unweighted OLS estimates of equation (17); Columns (3)–(6) report weighted least squares estimates.

Across all specifications, the row labeled *Office treatment share* reports the estimate of $\hat{\beta}$. For the Culture Fit Index, the coefficient ranges from roughly 0.14 standard deviations in the unweighted regressions to about 0.19 standard deviations when using squared weights, and is statistically significant at conventional levels throughout. The magnitude is economically meaningful and remains stable across weighting schemes.

Table 7: Office-Level Treatment Effects on Cultural Measures

	(1)	(2)	(3) W=N	(4) W=N#Sh.	(5) W=N^2	(6) W=(N#Sh.)^2	(7) Mean (SD)	(8) N
Panel A: Culture Fit Index								
Office treatment share	0.1438**	0.1426**	0.1410**	0.1247*	0.1845***	0.1901***		
	(0.0704)	(0.0710)	(0.0688)	(0.0645)	(0.0507)	(0.0479)		
	[0.0416]	[0.0452]	[0.0409]	[0.0536]	[0.0003]	[0.0001]		
Baseline	0.3058***	0.3060***	0.4960***	0.5555***	0.6033***	0.6013***	0.0178	664
	(0.0396)	(0.0399)	(0.0611)	(0.0598)	(0.0466)	(0.0465)	(0.5652)	
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]		
Baseline share missing		-0.0243	-0.4510***	-0.4864***	-1.3033***	-1.2872***		
		(0.1092)	(0.1487)	(0.1531)	(0.0706)	(0.0722)		
		[0.8237]	[0.0025]	[0.0016]	[0.0000]	[0.0000]		
Panel B: Social Responsibility								
Office treatment share	0.1176	0.1005	0.0957**	0.1001***	0.0885***	0.1006***		
	(0.0744)	(0.0738)	(0.0391)	(0.0376)	(0.0309)	(0.0344)		
	[0.1146]	[0.1739]	[0.0147]	[0.0080]	[0.0043]	[0.0036]		
Baseline	0.3331***	0.3312***	0.5450***	0.5968***	0.8516***	0.8617***	3.9703	645
	(0.0471)	(0.0467)	(0.0519)	(0.0573)	(0.0769)	(0.0735)	(0.4375)	
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]		
Baseline share missing		-0.2681**	-0.4309***	-0.4292***	-0.5780***	-0.5783***		
		(0.1103)	(0.1011)	(0.1039)	(0.0427)	(0.0385)		
		[0.0154]	[0.0000]	[0.0000]	[0.0000]	[0.0000]		

Notes: *Each column reports estimates from office-level regressions of the indicated cultural outcome on the office treatment share. Columns (1)–(2) report unweighted OLS specifications. Columns (3)–(6) report weighted least squares specifications: Column (3) weights offices by the baseline number of survey respondents; Column (4) weights by the baseline respondent count adjusted by the baseline response share; Columns (5) and (6) use squared versions of Columns (3) and (4), respectively. All weights are constructed using baseline survey data only. All regressions include randomization-strata fixed effects. Column (7) reports the baseline mean and standard deviation of the dependent variable, and Column (8) reports the number of office-level observations. Robust standard errors are reported in parentheses and p-values in brackets.* $*p < 0.10, **p < 0.05, ***p < 0.01$.

For Social Responsibility, the point estimates are positive and broadly similar in magnitude but less precisely estimated. In the unweighted specifications, the coefficients are 0.12 and 0.10 standard deviations (with $p = 0.11$ and $p = 0.17$, respectively). The weighted regressions yield estimates of comparable size, though they remain statistically insignificant. The lower precision is consistent with two features of this outcome: its categorical nature, which reduces variation and therefore statistical power, and the fact that it is observed in only one post-treatment wave, increasing measurement noise due to non-response.

Taken together, the results indicate that the intervention shifted perceived cultural fit upward at the office level. For the Culture Fit Index, the effects are larger than those estimated at the individual level in Table 6, consistent with the office-level amplification predicted by our theoretical framework.

Office-Level Change in Within-Office Dispersion In Table 8, we examine whether the intervention affects cultural dispersion within offices. Specifically, we estimate:

$$\Delta \text{sd}_{o,s} = \alpha + \beta T_o + \gamma y_o^{\text{pre}} + \kappa \, \text{miss}_o^{\text{pre}} + \lambda_s + \varepsilon_{o,s}, \tag{18}$$

where $\Delta \text{sd}_{o,s} = \text{sd}_o^{\text{post}} - \text{sd}_o^{\text{pre}}$, and $\text{sd}_o^{\text{post}}$ and sd_o^{pre} denote the post-treatment and baseline within-office standard deviations of the outcome, respectively. y_o^{pre} is the baseline office-level mean, and the remaining terms are defined as in equation (17). As before, Columns (1)–(3) report unweighted OLS estimates, while Columns (4)–(7) report weighted least squares estimates using baseline weights.

The coefficient β captures whether greater treatment exposure is associated with systematic changes in within-office dispersion. A negative β implies convergence at the office level: treated offices experience larger declines (or smaller increases) in the standard deviation of beliefs relative to less-treated offices, indicating compression of the within-office distribution.

We find suggestive evidence of a reduction in within-office dispersion, particularly for the Culture Fit Index. In the unweighted specifications, the coefficients on office treatment share are negative but imprecisely estimated. However, once we weight offices to account for differences in baseline measurement precision, the pattern becomes consistent and statistically meaningful. In the weighted regressions the estimated reduction in the within-office standard deviation ranges between 0.10 and 0.14. These magnitudes are economically meaningful relative to the overall variation in dispersion and indicate compression of the cultural distribution in more-treated offices. For Social Responsibility, the pattern is qualitatively similar but the coefficients are smaller and imprecisely estimated.

To complement Table 8, we examine whether convergence reflects differential adjustment at the individual level in Appendix Table B26. For both the Culture Fit Index and Social Responsibility, we estimate:

$$y_{i,\tau,s}^{\text{post}} = \alpha + \beta y_i^{\text{pre}} + \theta T_{o(i)} + \pi \big(T_{o(i)} \times y_i^{\text{pre}} \big) + \lambda_s + \delta_\tau + \varepsilon_{i,\tau,s}, \qquad \tau \geq 3, \tag{19}$$

where y_i^{pre} denotes the individual's baseline belief (averaged over semesters 1–2), and $T_{o(i)}$ is the treatment share in individual i's office. The specification includes strata fixed effects λ_s and semester fixed effects δ_τ, with standard errors clustered at the office level. The interaction coefficient π captures whether treatment effects vary with baseline beliefs, with a negative π implying convergence.[29]

Table B26 reports the estimates of equation (19). The interaction between office treatment share and baseline beliefs is negative for both outcomes, consistent with stronger

[29] In alternative specifications, we replace the continuous interaction $T_{o(i)} \times y_i^{\text{pre}}$ with interactions between treatment exposure and baseline belief groups. For the Culture Fit Index, individuals are partitioned into terciles based on baseline quartiles (Low, Medium, High). For Social Responsibility, baseline categories follow the discrete response scale (Low = 1–3, Medium = 4, High = 5), with the medium group as the reference category.

Table 8: Office-Level Treatment Effects on Changes in Within-Office Dispersion (SD)

	(1)	(2)	(3)	(4) W=N	(5) W=N#Sh.	(6) W=N^2	(7) W=(N#Sh.)^2	Mean (SD)	N
Panel A: Δ Culture Fit Index									
Office treatment share	-0.0494	-0.0486	-0.0252	-0.0994**	-0.0977**	-0.1304***	-0.1380***	0.079	654
	(0.0665)	(0.0666)	(0.0648)	(0.0459)	(0.0461)	(0.0347)	(0.0341)	(0.390)	
	[0.4579]	[0.4654]	[0.6978]	[0.0306]	[0.0343]	[0.0002]	[0.0001]		
Baseline mean		0.0130	0.0085	0.0717**	0.0781**	0.1589***	0.1584***		
		(0.0388)	(0.0365)	(0.0301)	(0.0325)	(0.0388)	(0.0373)		
		[0.7383]	[0.8167]	[0.0173]	[0.0167]	[0.0000]	[0.0000]		
Baseline share missing			0.4538***	0.2572***	0.2206***	0.3663***	0.3623***		
			(0.0926)	(0.0646)	(0.0678)	(0.0559)	(0.0536)		
			[0.0000]	[0.0001]	[0.0012]	[0.0000]	[0.0000]		
Panel B: Δ Social Responsibility									
Office treatment share	0.0705	0.0528	0.0647	-0.0260	-0.0353	-0.0805	-0.0903**	0.007	598
	(0.0751)	(0.0743)	(0.0741)	(0.0507)	(0.0508)	(0.0489)	(0.0429)	(0.429)	
	[0.3483]	[0.4779]	[0.3829]	[0.6081]	[0.4866]	[0.1000]	[0.0359]		
Baseline mean		0.1793***	0.1844***	0.1365***	0.1309***	0.1597***	0.1565***		
		(0.0517)	(0.0500)	(0.0455)	(0.0484)	(0.0537)	(0.0446)		
		[0.0006]	[0.0002]	[0.0028]	[0.0070]	[0.0030]	[0.0005]		
Baseline share missing			0.3347***	0.1368*	0.1098	0.2119***	0.2159***		
			(0.1122)	(0.0734)	(0.0839)	(0.0658)	(0.0550)		
			[0.0030]	[0.0630]	[0.1909]	[0.0014]	[0.0001]		

Notes: *Each column reports estimates from office-level regressions of the change in within-office dispersion on the office treatment share. All regressions include randomization-strata fixed effects. Columns (1)–(2) report unweighted OLS specifications. Columns (3)–(6) replicate Column (2) using baseline-derived weights. Specifically, Column (3) weights offices by the number of baseline respondents; Column (4) adjusts this weight by the baseline response share; Columns (5) and (6) use squared versions of Columns (3) and (4), respectively. All weights are constructed using baseline survey data only. Column (7) reports the baseline mean and standard deviation of the dependent variable, and Column (8) reports the number of office-level observations. Robust standard errors are reported in parentheses and p-values in brackets.* $*p < 0.10, **p < 0.05, ***p < 0.01$.

adjustment among initially lower-belief individuals ($p = 0.078$ and $p = 0.10$, respectively).[30]

Taken together, Table 7 and Table 8 (and Table B26) suggest that the intervention not only shifts the mean level of cultural alignment upward but also reduces dispersion within offices, particularly for the Culture Fit Index. More-treated offices exhibit stronger convergence toward a higher level of cultural fit. The effects on dispersion are smaller and less precisely estimated than those on the mean, which may reflect the fact that convergence in variance is a slower-moving margin.

[30] In the categorical specifications, using the medium baseline category as the reference group, the interaction between treatment and the low-baseline group is positive and statistically significant at the 10% level for Social Responsibility, while the interaction with the high-baseline group is close to zero.

Culture and Performance Finally, consistent with the model, we examine whether convergence in cultural beliefs translates into improvements in performance. Table 9 reports estimates of the following office-level specification:

$$\Delta \text{perf}_o = \alpha + \beta_1 \Delta y_o + \beta_2 \Delta \text{sd}_o + \beta_3 \left(\Delta y_o \times \Delta \text{sd}_o\right) + \theta T_o + \lambda_s + \varepsilon_o, \tag{20}$$

where Δperf_o denotes the change in average performance in office o, Δy_o the change in the mean belief measure (Culture Fit Index or Social Responsibility), and Δsd_o the change in within-office dispersion. T_o is the office-level treatment share and λ_s are randomization-strata fixed effects.

Across specifications, changes in the mean level of cultural alignment are strongly and positively associated with improvements in performance. In Column (1), a one–standard–deviation increase in the Culture Fit Index is associated with an increase in performance of approximately 0.07 ($p < 0.01$), with a very similar magnitude for Social Responsibility. Importantly, these coefficients remain stable when controlling for changes in dispersion and when allowing for their interaction, indicating that mean shifts in beliefs are the primary driver of performance gains.

Moreover, once changes in beliefs are included, the office-level treatment share T_o is small and statistically insignificant across all specifications. This suggests that the direct effect of treatment at the office level operates through its impact on cultural beliefs. In other words, performance improvements are mainly explained by changes in culture rather than by treatment intensity per se.

By contrast, changes in dispersion are not independently related to performance. The coefficient on Δsd_o is small and statistically insignificant, and the interaction term is imprecisely estimated. However, when splitting the sample by the sign of dispersion changes, the effect of Δy_o is larger in offices where dispersion declines ($\Delta \text{sd}_o < 0$) than where it increases. While this difference is not statistically significant, it is consistent with the model's prediction that convergence toward a shared high-culture equilibrium reinforces the performance gains associated with higher average cultural alignment.

Table 9: Performance Effects of Belief Convergence

	(1)	(2)	(3)	(4)	(5) $\Delta SD < 0$	(6) $\Delta SD \geq 0$
Panel A: Culture Fit Index						
Δ Belief	0.0682***		0.0668***	0.0654***	0.0855***	0.0581**
	(0.0186)		(0.0189)	(0.0192)	(0.0282)	(0.0252)
	[0.0003]		[0.0004]	[0.0007]	[0.0027]	[0.0220]
Δ SD		-0.0278	-0.0092	-0.0185		
		(0.0276)	(0.0280)	(0.0334)		
		[0.3129]	[0.7434]	[0.5804]		
Δ Mean × Δ SD				0.0207		
				(0.0417)		
				[0.6200]		
Treatment	0.0324	0.0349	0.0327	0.0316	0.0617	0.0013
	(0.0360)	(0.0360)	(0.0360)	(0.0363)	(0.0488)	(0.0544)
	[0.3690]	[0.3327]	[0.3640]	[0.3846]	[0.2067]	[0.9814]
Mean DV	0.0684	0.0684	0.0684	0.0684	0.0703	0.0663
	(0.2326)	(0.2326)	(0.2326)	(0.2326)	(0.2361)	(0.2291)
Observations	625	625	625	625	324	301
Panel B: Social Responsibility						
Δ SR	0.0596***		0.0668***	0.0662***	0.0897***	0.0418
	(0.0191)		(0.0193)	(0.0193)	(0.0284)	(0.0283)
	[0.0019]		[0.0006]	[0.0007]	[0.0017]	[0.1408]
Δ SD		0.0129	0.0323	0.0325		
		(0.0238)	(0.0241)	(0.0240)		
		[0.5886]	[0.1798]	[0.1769]		
Δ Mean × Δ SD				-0.0157		
				(0.0401)		
				[0.6966]		
Treatment	0.0264	0.0201	0.0259	0.0260	0.0847*	-0.0701
	(0.0373)	(0.0375)	(0.0372)	(0.0372)	(0.0492)	(0.0597)
	[0.4797]	[0.5920]	[0.4858]	[0.4852]	[0.0861]	[0.2417]
Mean DV	0.0689	0.0689	0.0689	0.0689	0.0621	0.0772
	(0.2248)	(0.2248)	(0.2248)	(0.2248)	(0.2309)	(0.2174)
Observations	576	576	576	576	316	260

Notes. *Each column reports an office-level regression of the change in average performance (post minus pre). Panel A relates performance changes to changes in the Culture Fit Index; Panel B relates performance changes to changes in Social Responsibility. Column (1) includes only the change in the mean index; Column (2) includes only the change in within-office dispersion; Column (3) includes both mean and dispersion changes additively; Column (4) adds their interaction term; Columns (5) and (6) restrict the sample to offices with declining ($\Delta SD < 0$) and increasing ($\Delta SD \geq 0$) dispersion, respectively, and include only the change in the mean index. All regressions include randomization-strata fixed effects. Robust standard errors are reported in parentheses and p-values in brackets.* $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

8 Leaders

So far, we have focused on the universe of employees, from bank tellers to senior executives. In the model, effort decisions reflect a tradeoff between the private cost of effort and the perceived cultural environment. In practice, this cost component varies systematically across the hierarchy. At the bottom of the organization, compensation relies almost entirely on fixed pay, whereas at the top it depends heavily on performance-based variable pay in the form of bonuses and stock options. While we do not observe individual pay figures for confidentiality reasons, internal discussions with the bank's HR team confirm a stark divergence in compensation structure within our partner institution.[31]

In the framework, greater exposure to variable pay increases the opportunity cost of time, effectively raising the cost parameter in the effort decision. This suggests that senior employees may face higher private costs of participating in volunteering activities, potentially dampening take-up. We therefore ask whether differences in compensation structure — and thus opportunity costs — translate into differential engagement with the policy. Section 8.1 examines this channel.

Beyond opportunity costs, the model emphasizes the role of credibility in shifting equilibrium culture. In particular, the intervention's impact depends on whether employees perceive it as a genuine change in organizational priorities. Leaders play a central role in shaping this perception. If senior employees visibly participate, they reinforce the credibility of the prosocial shift; if they do not, the signal may weaken. This generates a second question: to what extent does cultural change depend on leaders' own engagement? Section 8.2 studies how leadership participation interacts with the intervention to shape cultural alignment within offices.

Identifying leaders We adopt two complementary definitions:

- **"Global Seniority"**: the bank's own classification of employees as junior, middle, or senior.
- **"Relative Seniority"**: within each office, we classify as "senior" the most senior employees in that office.[32]

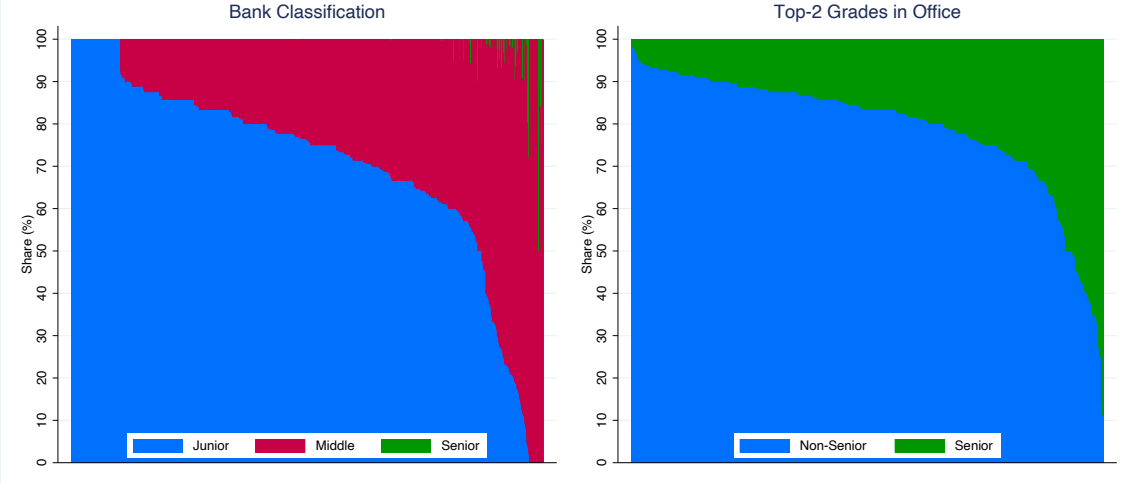
Appendix B.5 describes in detail how we construct these two measures. Table B25 reports summary statistics. As expected, relative seniority is a much broader classification that encompasses about 34% of the sample, whereas under the bank's global classification only about 9% of employees are classified as senior.

[31] This pattern is consistent with the broader UK banking sector, which recently successfully lobbied for the removal of the regulatory cap limiting bonuses to 100% of fixed salary. This suggests that, at the top of the ladder, variable pay often exceeds fixed pay.

[32] Under this classification, we exclude offices in which we cannot distinguish juniors from seniors, either because they are too small or because everyone is in the same seniority band.

Figure 13 plots the office level composition by seniority under the two definitions: global seniority (left panel) and relative seniority (right panel). Each vertical bar is an office, and each color represents the share of employees in that office belonging to a given seniority group. Under the bank's global classification (left), most offices contain only junior and middle-level staff, and only a small number of (large, urban) offices have any employees in the "senior" category. By contrast, under the relative classification (right), we explicitly restrict attention to offices in which we can observe both juniors and seniors, ensuring within-office variation in seniority across all offices.

Figure 13: Area Plot: Office Share by Seniority



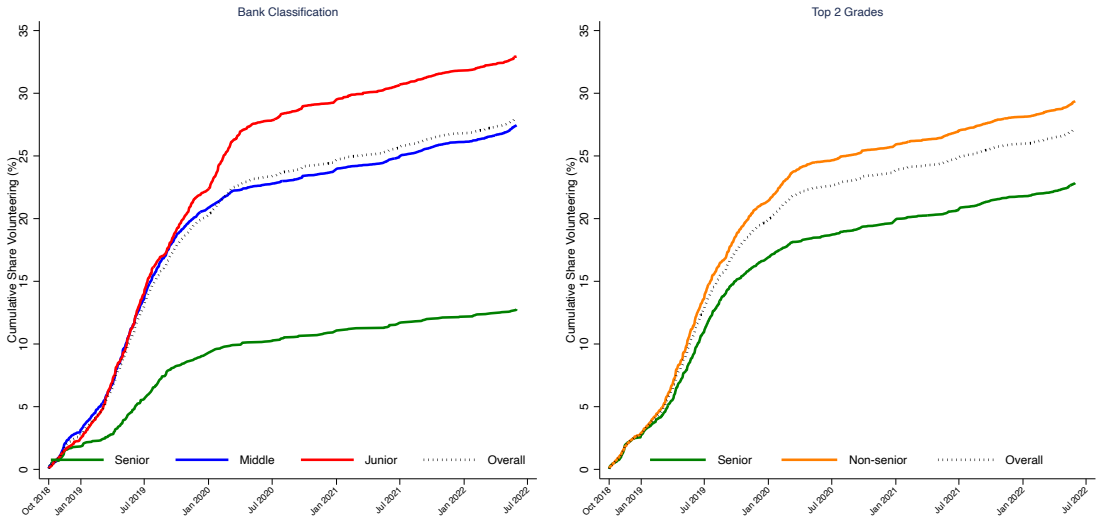
8.1 Take-up by Seniority

Figure 14 shows that leaders are consistently less likely to volunteer, regardless of how we define seniority. Under the bank's global classification, by the end of the study period seniors volunteer at a rate that is roughly one third of that of juniors (left panel). Under the relative seniority definition, cumulative take-up among seniors is about 10 percentage points below that of non-seniors throughout the study period.[33]

Additionally, Appendix Figure C10 shows that leaders with higher baseline cultural fit volunteer substantially more than leaders with lower baseline cultural fit.

[33]The "overall" lines are not identical because, under the relative definition, we restrict to the subset of offices in which we can define at least one senior and at least one junior; see Appendix B.5.

Figure 14: Volunteering Take Up by Seniority



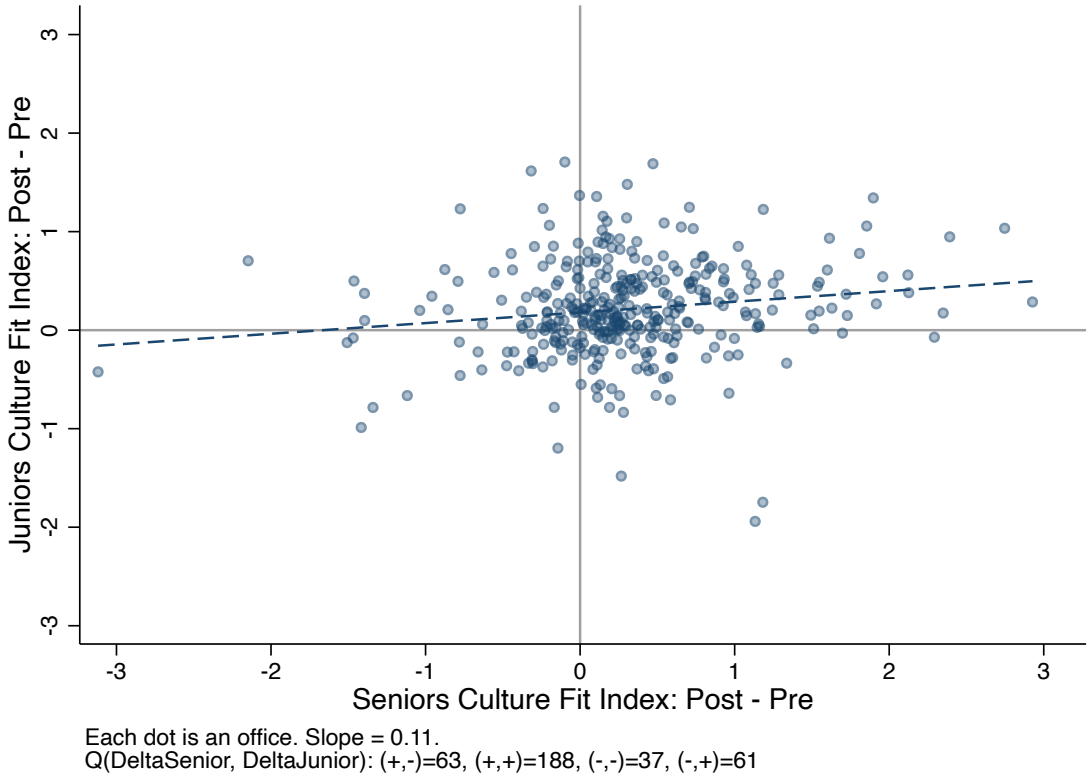
8.2 Seniors Influence on Cultural Change

Throughout this section, we study how important leaders' engagement with the policy is for cultural change.

Before turning to causal estimates, Figure 15 documents a positive association between cultural changes among seniors (defined as the top two grades within each office) and juniors. Each dot represents an office, plotting the change in the Culture Fit Index for seniors on the horizontal axis and for juniors on the vertical axis. The fitted slope is 0.11 ($p < 1\%$), indicating that offices in which seniors experience larger increases in cultural fit also tend to exhibit increases among juniors. The quadrant counts reinforce this pattern: the largest mass of offices lies in the $(+,+)$ quadrant (188 offices), where both seniors and juniors increase, followed by 63 offices in $(+,-)$, 61 in $(-,+)$, and 37 in $(-,-)$. This suggests that cultural change among seniors and juniors tends to move in the same direction within offices, consistent with the mechanism outlined in our model.[34]

[34]The sample is restricted to offices with at least two seniors and two juniors, since mean and dispersion changes cannot be computed otherwise.

Figure 15: Office-level Correlation Δ Seniors (Mean) Beliefs with Δ Juniors (Mean) Beliefs



Notes. *Each dot represents an office. The figure plots the change in average beliefs among junior employees against the change among senior employees between the pre and post periods. The dashed line shows the fitted linear regression of junior belief changes on senior belief changes; the reported slope corresponds to this regression coefficient.*

8.2.1 Volunteering Gap

For each office o and semester τ, we construct junior and senior volunteering shares (and analogously, treatment shares) as leave–one–out (LOO) means based on relative seniority at baseline. Individuals are classified as juniors or seniors within their office according to their position in the office-level seniority distribution.

The junior (senior) volunteering share equals the average volunteering rate among other juniors (seniors) in the same office and semester, excluding the individual from both numerator and denominator. The junior–senior volunteering gap is the difference between these two LOO shares. The junior–senior treatment gap is defined analogously using treatment

shares. Higher values indicate that juniors volunteer (or are treated) relatively more than seniors within the same office and semester.[35]

Table 10 reports estimates of the following individual level regressions:

- First stage:

$$\text{VolGap}_{o\tau} = \alpha + \pi\,\text{TreatGap}_{o\tau} + \lambda_s + \delta_\tau + \Gamma\big(X_i^{\text{pre}} \times \delta_\tau\big) + u_{o\tau}.$$

- Reduced form:

$$Y_{io\tau} = \alpha + \rho\,\text{TreatGap}_{o\tau} + \lambda_s + \delta_\tau + \Gamma\big(X_i^{\text{pre}} \times \delta_\tau\big) + \varepsilon_{io\tau}.$$

- 2SLS:

$$Y_{io\tau} = \alpha + \beta\,\widehat{\text{VolGap}}_{o\tau} + \lambda_s + \delta_\tau + \Gamma\big(X_i^{\text{pre}} \times \delta_\tau\big) + \varepsilon_{io\tau},$$

  where $\text{VolGap}_{o\tau}$ is instrumented by $\text{TreatGap}_{o\tau}$.

Here, $\text{VolGap}_{o\tau}$ and $\text{TreatGap}_{o\tau}$ denote the junior–senior volunteering and treatment gaps in office o and semester τ, and $Y_{io\tau}$ is the individual-level outcome. λ_s are randomization-strata fixed effects and δ_τ are semester fixed effects. X_i^{pre} denotes baseline beliefs, interacted with semester fixed effects. Standard errors are clustered at the office level. Panel A controls for office-level treatment shares; Panel B additionally controls for individual treatment assignment.

The first stage shows that the junior–senior treatment gap strongly predicts the junior–senior volunteering gap. Offices in which juniors are relatively more likely to be treated are also those in which juniors end up volunteering relatively more than seniors (Cluster robust F-statistic ≈ 15; Kleibergen–Paap F-statistic ≈ 9.9 in both panels). The experimental variation therefore shifts the relative composition of volunteering across hierarchical groups. The reduced form indicates that a larger junior–senior treatment gap is associated with lower cultural fit.

Crucially, the 2SLS estimates deliver a negative and economically meaningful LATE. Instrumenting the office–semester junior–senior volunteering gap with the office–semester junior–senior treatment gap, we find that, for compliers — i.e. individuals in offices where the randomized variation induces a change in the volunteering gap — an increase in juniors' volunteering relative to seniors causally reduces individual perceived cultural fit ($p \approx 0.08$ in both panels). Put differently: among offices where the experiment causes juniors (but not seniors) to volunteer more, treated individuals report weaker cultural alignment.

This pattern is consistent with the credibility channel in the model. The initiative affects culture not only through its introduction, but through whether it is visibly reinforced by leaders. When juniors engage but seniors do not, the intervention lacks validation from

[35] Gaps are constructed using relative seniority at baseline. Under the bank's global seniority classification, seniors are present in too few offices.

Table 10: Culture: Treatment Gaps & Volunteering Gaps

	(1) First stage	(2) Reduced form	(3) 2SLS	(4) Cluster-F	(5) KP-F	(6) Mean (SD)	(7) N
Panel A: Controls = Office T.Sh.							
Volunteering Gap	0.1887***	-0.1588**	-0.8881*	15.25	9.86	0.0865	18057
	(0.0483)	(0.0731)	(0.5039)			(0.9863)	
	[0.0001]	[0.0305]	[0.0780]				
Panel B: Controls = Office T.Sh. + T.Assignment							
Volunteering Gap	0.1885***	-0.1592**	-0.8907*	15.24	9.87	0.0865	18057
	(0.0483)	(0.0730)	(0.5019)			(0.9863)	
	[0.0001]	[0.0297]	[0.0760]				

Notes: *All specifications include semester and strata fixed effects and control for baseline beliefs interacted with time. Seniority (relative) at baseline: individuals are classified as seniors if they belong to the highest (or top two) career grades in their office, and juniors otherwise. For each office and semester, we construct leave–one–out (LOO) shares separately for juniors and seniors. The junior (senior) volunteering share is computed as the mean volunteering rate among other juniors (seniors) in the same office and semester, excluding the individual. The junior–senior volunteering gap equals the LOO junior share minus the LOO senior share. Analogously, the junior–senior treatment gap is defined as the LOO difference in treatment shares between juniors and seniors within office and semester. Column (1) reports the first-stage regression of the junior–senior volunteering gap on the junior–senior treatment gap. Column (2) reports the reduced-form effect of the treatment gap on beliefs. Column (3) reports the IV coefficient on the volunteering gap instrumented by the treatment gap. Column (4) reports the cluster-robust Wald F-statistic testing the null that the excluded instrument has zero effect in the first stage. Column (5) reports the Kleibergen–Paap rk Wald F-statistic for weak identification. Panel A controls for office treatment share. Panel B adds individual treatment assignment as a control. Standard errors clustered at the office level are reported in parentheses; p-values in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

high-status actors. In that case, the policy does not move the office toward a higher cultural equilibrium. The negative 2SLS effect is therefore consistent with a setting in which volunteering concentrated among juniors fails to strengthen — and may even undermine — perceived prosocial alignment within the organization.

Robustness Appendix Table B27 replicates Table 10 after winsorizing the junior–senior treatment and volunteering gaps at the 1st and 99th percentiles. The 2SLS estimates become larger in absolute value and remain statistically significant. This indicates that the negative effect of a volunteering gap in favour of juniors is not driven by extreme offices. If

anything, trimming outliers strengthens the estimated impact.[36]

Finally, Appendix Table B30 examines the volunteering gap at the office level, defined as the difference between junior and senior post-period volunteering shares within each office (and analogously for treatment). In the unweighted specifications, the IV coefficient is negative but imprecisely estimated, reflecting a relatively weak first stage (Cluster-F ≈ 6.7; KP-F ≈ 6.7). Once we weight offices by baseline respondent count, or by respondent count adjusted for response rates, the first stage strengthens substantially (Cluster-F ≈ 13.1 and 10.5, respectively; KP-F ≈ 11.6 and 9.2), and the 2SLS coefficient becomes large and negative ($p = 0.0554$ and $p = 0.0689$, respectively).

8.2.2 Peer Effects

To directly estimate peer effects, we first construct leave–one–out (LOO) baseline senior shares for each office (and team) for both seniority definitions. Then, we partition seniors by their baseline Culture Fit Index, defining "low culture fit" seniors as those whose baseline index falls below a given percentile of the given senior distribution (median, 33rd, 25th, or 10th). Third, we interact the LOO senior share with the "low culture fit" indicator to obtain, for each office (and team) at baseline, the share of seniors with low culture.

In Table 11 (and Appendix Tables B31, B32, B34), we estimate

$$\begin{aligned} y_{is\tau} = \alpha \;+\; \beta\, T_i \;+\; \theta^{\text{off}} S_i^{\text{off}} \;+\; \theta^{\text{team}} S_i^{\text{team}} \\ +\; \pi^{\text{off}}\big(T_i \times S_i^{\text{off}}\big) \;+\; \pi^{\text{team}}\big(T_i \times S_i^{\text{team}}\big) \\ +\; \phi\, y_i^{\text{pre}} \;+\; \sum_{\tau \in \{4,6\}} \kappa_\tau\big(y_i^{\text{pre}} \cdot \mathbf{1}\{\tau\}\big) \;+\; \lambda_s \;+\; \delta_\tau \;+\; \varepsilon_{is\tau}. \end{aligned} \tag{21}$$

where $y_{is\tau}$ denotes the Culture Fit Index for individual i in stratum s and semester τ, and T_i is the individual level treatment indicator. The variables S_i^{off} and S_i^{team} are the baseline LOO senior shares (or "low culture fit" senior shares) in individual i's office and team. The term y_i^{pre} is the baseline Culture Fit Index, included both in levels and interacted with semester dummies. We include stratum fixed effects λ_s and semester fixed effects δ_τ, and cluster standard errors at the office level.[37]

Table 11, our main table in this section, uses the relative seniority definition to construct seniors shares. Across all columns, the treatment effect on the Culture Fit Index remains positive and stable, with $\hat{\beta}$ ranging between 0.1209 and 0.1301 (all $p < 0.01$). In Columns (1) and (2), S^{off} and S^{team} are the baseline LOO senior shares. The interaction between treatment and the office senior share, $\hat{\pi}^{\text{off}}$, is negative and statistically significant in both

[36] We replicate both specifications using Social Responsibility as the outcome (Appendix Tables B28 and B29). The 2SLS coefficients have the same sign as in the baseline results, but are not statistically significant. This is consistent with a weaker first-stage, likely reflecting two features of the SR measure: it is categorical (reducing variation), and it is observed in only one post-treatment wave, which limits precision.

[37] In columns (1), (3), and (5) we only use office exposure, so S_i^{team} and the associated coefficients θ^{team} and π^{team} are set to zero.

cases ($\hat{\pi}^{\text{off}} = -0.1437$, $p = 0.020$ in Column (1); $\hat{\pi}^{\text{off}} = -0.2015$, $p = 0.029$ in Column (2)), showing that culture gains from treatment are smaller in offices with more seniors. In Columns (3) and (4), S^{off} and S^{team} are defined as the LOO shares of seniors whose baseline Culture Fit Index is below the median; the estimated treatment–office share interactions, $\hat{\pi}^{\text{off}}$, are over three times larger in absolute value, although only marginally significant at the 10% level ($\hat{\pi}^{\text{off}} = -0.4839$, $p = 0.091$; $\hat{\pi}^{\text{off}} = -0.6288$, $p = 0.076$). Finally, in Columns (5) and (6), we restrict S^{off} and S^{team} to seniors in the bottom quartile of the baseline culture distribution; the corresponding interaction estimates are over six times as large and precisely estimated ($\hat{\pi}^{\text{off}} = -1.1622$, $p = 0.016$; $\hat{\pi}^{\text{off}} = -1.2205$, $p = 0.039$). Notably, across all specifications, the coefficients on the team senior share and its interaction with treatment, $\hat{\theta}^{\text{team}}$ and $\hat{\pi}^{\text{team}}$, are small and not statistically different from zero, confirming that peer effects are occurring at the office level rather than at the team level.

Focusing on Columns (2), (4), and (6) as our preferred specifications, we now interpret the interactions in relation to the treatment effect. In Column (2), the baseline treatment effect when the office senior share is zero is $\hat{\beta} = 0.1220$. A ten percentage point increase in the office senior share ($\Delta S^{\text{off}} = 0.10$) reduces the treatment effect by $0.10 \times \hat{\pi}^{\text{off}} = 0.10 \times (-0.2015) \approx -0.020$, a decline of about 17 percent. In Column (4), where S^{off} is the share of seniors with baseline Culture Fit Index below the median, the same ten percentage point increase lowers the treatment effect by $0.10 \times (-0.6288) \approx -0.063$, about 49 percent. In Column (6), where S^{off} is the share of seniors in the bottom quartile of the culture distribution, a ten percentage point increase reduces the treatment effect by $0.10 \times (-1.2205) \approx -0.122$, from 0.1301 to about 0.008, a 94 percent drop. This implies that even modest increases in the share of low-culture-fit-seniors can lead to substantial erosion of the effectiveness of the policy.

As in the previous section, the peer effect evidence shows that the intervention shifts culture only when it is reinforced by senior peers. When a larger share of seniors hold low baseline prosocial beliefs, the treatment effect on culture is substantially attenuated. This is consistent with the model's prediction that leadership behavior determines the credibility of the shock and therefore whether offices move toward a higher cultural equilibrium.

Robustness Appendix Table B31 shows that these patterns are robust to adding rich baseline office level controls (Culture Fit Index, performance, values, office size, function mix, and turnover). In particular, when including these baseline covariates, the interaction coefficients are more precisely estimated across all regressions. In Columns (1) and (2) the magnitudes of the interactions are very similar to those in Table 11. By contrast, in Columns (3) to (6) the treatment–office share interactions for low culture seniors show larger estimates relative to Table 11, at -0.5831 ($p = 0.049$) and -0.7485 ($p = 0.034$), -1.3369 ($p < 0.01$) and -1.3928 ($p = 0.018$), respectively. These results confirm that the heterogeneous effects we document are not simply driven by where seniors were initially placed across offices.

Appendix Table B32 splits seniors according to whether their baseline Culture Fit Index

lies below the median, the 33rd percentile, the 25th percentile, or the 10th percentile. We find a striking monotonic pattern: as we 'tighten' the threshold, the treatment–office share interaction for low culture seniors becomes steadily more negative and more precisely estimated. With the median split, the treatment–office share interaction is -0.4839 ($p = 0.091$) and -0.6288 ($p = 0.076$); at the 33rd percentile it is -0.7554 ($p = 0.054$) and -0.9953 ($p = 0.039$); at the 25th percentile -1.1622 ($p = 0.016$) and -1.2205 ($p = 0.039$); and at the 10th percentile -2.8553 ($p < 0.01$) and -3.0413 ($p < 0.01$).[38]

Finally, Appendix Table B34 replicates Table 11 using the bank's "global seniority" classification. Across all columns, the treatment effect on the Culture Fit Index remains around 0.10 ($p < 0.01$). In Columns (1) and (2), S^{off} and S^{team} are the baseline LOO shares of global seniors. The interaction between treatment and the office senior share, $\hat{\pi}^{\text{off}}$, is negative and precisely estimated ($\hat{\pi}^{\text{off}} = -0.3094$, $p = 0.002$; $\hat{\pi}^{\text{off}} = -0.3359$, $p = 0.037$), implying that offices with more global seniors see smaller culture gains from treatment. In Columns (3) and (4), we restrict S^{off} and S^{team} to seniors whose baseline Culture Fit Index lies below the median; the interaction estimates become larger in absolute value, $\hat{\pi}^{\text{off}} = -1.0354$ ($p = 0.045$) and $\hat{\pi}^{\text{off}} = -1.3620$ ($p = 0.039$). In Columns (5) and (6), where S^{off} and S^{team} capture seniors in the bottom quartile of the baseline culture distribution, the treatment–office share interactions are even more negative, at $\hat{\pi}^{\text{off}} = -2.2220$ ($p = 0.036$) and $\hat{\pi}^{\text{off}} = -2.8411$ ($p = 0.023$). As in Table 11, the team share coefficients $\hat{\theta}^{\text{team}}$ and $\hat{\pi}^{\text{team}}$ are small and not statistically different from zero. On the other hand, a notable difference relative to Table 11 is that the coefficient on the office senior share in levels is itself large, negative, and precisely estimated. This indicates that offices with a higher fraction of global seniors (typically, large urban offices) have lower culture outcomes to start with, even after conditioning on baseline culture, and that the treatment effect is weaker there, as these seniors act as a drag on cultural change.

Finally, Tables B35 and B36 replicate Tables 11 and B34 using Social Responsibility as the outcome. Because this variable is measured on a 1–5 scale, we classify "low" Social Responsibility as baseline responses of 3 or below. Under the relative seniority definition (Table B35), the interaction between treatment and the share of low–SR seniors is negative, mirroring the pattern found for the Culture Fit Index, but imprecisely estimated. By contrast, under the bank's global seniority classification (Table B36), the interaction is large, negative, and statistically significant. Offices with a higher share of seniors holding low baseline SR experience substantially weaker cultural gains from the intervention. This pattern is consistent with the results for the Culture Fit Index and reinforces the interpretation that leadership alignment is central to the credibility of the prosocial shift.

[38] As a sanity check, Appendix Table B33 replicates Table B32 as a placebo, using only semesters 1 and 2 and replacing the treatment indicator with an indicator for semester 2; we do not find significant interactions in any of these placebo regressions ($p > 0.36$).

Table 11: Peer Effects: "Relative" Seniors

Culture Fit Index	(1) Seniors b/se/p	(2) Seniors b/se/p	(3) Median b/se/p	(4) Median b/se/p	(5) P25 b/se/p	(6) P25 b/se/p
Treat	0.1209***	0.1220***	0.1219***	0.1286***	0.1281***	0.1301***
	(0.0334)	(0.0346)	(0.0353)	(0.0357)	(0.0306)	(0.0313)
	[0.0003]	[0.0005]	[0.0006]	[0.0004]	[0.0000]	[0.0000]
OffSh. Seniors	-0.1645*	-0.1270				
	(0.0926)	(0.1074)				
	[0.0763]	[0.2378]				
Treat × OffSh. Seniors	-0.1437**	-0.2015**				
	(0.0617)	(0.0919)				
	[0.0201]	[0.0287]				
TeamSh. Seniors		-0.0320				
		(0.0371)				
		[0.3887]				
Treat × TeamSh. Seniors		0.0529				
		(0.0605)				
		[0.3825]				
OffSh. Seniors (Index<Med)			-0.6439	-0.6010		
			(0.4308)	(0.4683)		
			[0.1357]	[0.2001]		
Treat × OffSh. Seniors (Index<Med)			-0.4839*	-0.6288*		
			(0.2856)	(0.3534)		
			[0.0909]	[0.0759]		
TeamSh. Seniors (Index<Med)				-0.0367		
				(0.1085)		
				[0.7354]		
Treat × TeamSh. Seniors (Index<Med)				0.0694		
				(0.1438)		
				[0.6294]		
OffSh. Seniors (Index<P25)					-1.1002	-1.1033
					(0.7053)	(0.7186)
					[0.1195]	[0.1254]
Treat × OffSh. Seniors (Index<P25)					-1.1622**	-1.2205**
					(0.4821)	(0.5887)
					[0.0163]	[0.0388]
TeamSh. Seniors (Index<P25)						-0.0006
						(0.1931)
						[0.9974]
Treat × TeamSh. Seniors (Index<P25)						0.0172
						(0.2387)
						[0.9426]
N	19659	18733	15077	14346	15077	14346
N Offices	547	543	444	437	444	437
R2	0.3935	0.3923	0.4291	0.4273	0.4300	0.4280
Mean DV	0.1834	0.1796	0.1875	0.1819	0.1875	0.1819

Notes: *Each column reports estimates from individual-level regressions of the Culture Fit Index on treatment and baseline senior shares, including interactions between treatment and the relevant office- and/or team-level senior share. Column (1) includes only the treatment indicator. Columns (2)–(3) use the baseline leave–one–out (LOO) share of seniors within the office (and team, in Column (3)), based on the relative top-two grade definition. Columns (4)–(5) replace the senior share with the share of seniors whose baseline Culture Fit Index lies below the median of the senior distribution within the corresponding category. Columns (6)–(7) use instead the share of seniors below the 25th percentile of the baseline Culture Fit Index. All specifications control for the baseline Culture Fit Index (measured in semester 2; if unavailable, semester 1) interacted with semester indicators, and include semester and randomization-strata fixed effects. Office- and team-level shares are constructed as leave–one–out baseline shares. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

8.3 Vertical and Horizontal Offices

We next examine whether the effectiveness of the intervention depends on the degree of hierarchy within the office. Cultural change may be more difficult in vertically structured offices, where communication and influence operate primarily within career bands, potentially limiting the diffusion of new norms.

Table 12 reports estimates of the following office-level specification:

$$y_o = \beta_0 + \beta_1 T_o + \beta_2 \big(T_o \times V_o\big) + \gamma y_o^{\text{pre}} + \theta \log(\text{size}_o) + \kappa \text{miss}_o^{\text{pre}} + \lambda_s + \varepsilon_o.$$

Here, y_o denotes the office-level outcome (either the Culture Fit Index or Social Responsibility), T_o is the office treatment share, and y_o^{pre} is the baseline value of the outcome. The term λ_s captures randomization-strata fixed effects. We also control for log office size and the baseline share of missing survey responses. The verticality measure V_o is constructed in three alternative ways. In Column (1), verticality is defined as the vertical span within the office—measured as the difference between the highest and lowest career band present—and offices are grouped into terciles based on this raw span. In Column (2), we use the same vertical span measure but scale it by the log of office size to account for the fact that larger offices mechanically tend to include more career bands; the resulting scaled measure is entered as a continuous variable. Finally, in Column (3), we divide this scaled verticality measure into terciles.

Across specifications, the interaction terms between treatment and verticality are negative, indicating that treatment effects weaken as offices become more hierarchical. In Column (1), the baseline treatment effect on the Culture Fit Index is 0.2370 ($p = 0.041$), while the interaction terms for medium and high verticality offices are negative but imprecisely estimated. A similar pattern holds for Social Responsibility. In Column (2), the attenuation becomes more pronounced. For the Culture Fit Index, the baseline treatment effect is 0.4794 ($p = 0.023$) and the interaction term is -0.3086 ($p = 0.096$). For Social Responsibility, the baseline effect is 0.6024 ($p = 0.011$) and the interaction is -0.4590 ($p = 0.026$), indicating statistically significant attenuation. Finally, in Column (3), which splits the scaled measure into terciles, the baseline treatment effect on the Culture Fit Index is 0.3011 ($p = 0.003$), while the interaction for the highest verticality tercile is -0.3284 ($p = 0.025$), implying a near-zero net effect in highly vertical offices ($0.3011 - 0.3284 \approx -0.03$). For Social Responsibility, the high-verticality interaction is -0.4017 ($p = 0.0096$), again indicating strong attenuation.

Overall, the intervention generates substantially larger cultural gains in flatter offices, while hierarchical structure markedly dampens its effectiveness. This pattern maps directly to the model. Cultural change operates through two forces: a credibility shock that shifts beliefs directly, and conformist transmission that amplifies this shift through horizontal peer diffusion. The size of the equilibrium response therefore depends not only on the credibility parameter, but also on the strength of peer-to-peer updating. Verticality affects precisely this second channel. In highly hierarchical offices, interaction and influence are more seg-

mented within career bands, limiting cross-band observation and imitation and thereby weakening conformist transmission. As a result, even if the credibility shock initially raises beliefs, the equilibrium multiplier is smaller when verticality is high, leading to attenuated cultural change.

Table 12: Treatment Effects by Office

	(1)	(2)	(3)	Mean (SD)	N
Panel A: Culture Fit Index					
Office treatment share	0.2370**	0.4794**	0.3011***	0.0173	662
	(0.1159)	(0.2110)	(0.1029)	(0.5656)	
	[0.0412]	[0.0234]	[0.0035]		
Verticality ratio (cont.) × Treat		-0.3086*			
		(0.1849)			
		[0.0955]			
Verticality (Medium) × Treat	-0.1286		-0.1734		
	(0.1741)		(0.1727)		
	[0.4604]		[0.3159]		
Verticality (High) × Treat	-0.2083		-0.3284**		
	(0.1593)		(0.1465)		
	[0.1915]		[0.0253]		
Panel B: Social Responsibility					
Office treatment share	0.1860	0.6024**	0.1702	3.9710	643
	(0.1183)	(0.2366)	(0.1111)	(0.4378)	
	[0.1165]	[0.0111]	[0.1258]		
Verticality ratio (cont.) × Treat		-0.4590**			
		(0.2062)			
		[0.0264]			
Verticality (Medium) × Treat	-0.0747		0.2056		
	(0.1905)		(0.1774)		
	[0.6952]		[0.2469]		
Verticality (High) × Treat	-0.2425		-0.4017***		
	(0.1543)		(0.1546)		
	[0.1165]		[0.0096]		

Notes. *All regressions include randomization-strata fixed effects. Column (1) defines verticality by counting the number of career bands in each office and splitting offices into terciles according to it. Columns (2) and (3) scale the measure of verticality in (1) by log(office size): column (2) uses the continuous version, column (3) splits it into terciles. All measures are computed using baseline data. Variables not shown in the table for conciseness are baseline culture, log office size, and share missing at baseline; for columns (1) and (3) the level indicators for verticality groups are included in the regressions but omitted in the table. Robust standard errors are reported in parentheses and p-values in brackets.* $* < 0.10, ** < 0.05, *** < 0.01$.

9 Conclusion

This paper studies whether a credible prosocial intervention can shift organizational culture and generate equilibrium effects on performance and retention in a high-powered incentive environment. We show that integrating a prosocial initiative into core organizational policy can increase cultural alignment among employees, strengthen perceptions of social responsibility, and correct the misperceptions documented in Ashraf et al. (2020). Importantly, these shifts do not come at the expense of economic outcomes. If anything, performance improves, turnover declines, and disengagement falls. The evidence indicates that prosocial initiatives can alter the cultural equilibrium of an organization rather than merely affecting individual behavior at the margin.

A central finding is that effects are amplified at higher levels of aggregation. Offices with greater exposure experience disproportionately larger improvements in performance and retention, consistent with strategic complementarities in effort. Cultural change operates through interactions: as more employees internalize prosocial norms, the marginal return to discretionary effort rises for others. The intervention therefore generates an equilibrium multiplier.

At the same time, the success of the intervention depends critically on credibility. Limited engagement by senior employees attenuates cultural shifts, underscoring the role of leadership in sustaining belief updating and coordination around shared norms. When leaders reinforce the signal, offices move toward a higher cultural equilibrium; when they do not, belief updating stalls and cultural gains dissipate.

Our setting—a large global bank with strong individual incentives—provides a demanding test of whether culture can be reshaped without altering pay schemes. While caution is warranted in extrapolating to all contexts, the scale of the experiment and the structural features of the organization suggest that similar mechanisms may operate in other large, performance-oriented firms. More broadly, the results highlight that culture is not a residual or purely symbolic feature of organizations. When credible, it can be deliberately influenced and can generate economically meaningful equilibrium effects. Understanding the conditions under which such cultural shifts are sustained over the long run remains an important direction for future research.

References

Akerlof, George A and Rachel E Kranton, "Economics and identity," *The quarterly journal of economics*, 2000, *115* (3), 715–753.

_ and _ , "Identity and the Economics of Organizations," *Journal of Economic perspectives*, 2005, *19* (1), 9–32.

Alan, Sule, Gozde Corekcioglu, and Matthias Sutter, "Improving workplace climate in large corporations: A clustered randomized intervention," *The Quarterly Journal of Economics*, 2023, *138* (1), 151–203.

Ashraf, N. and O. Bandiera, "Social Incentives in Organizations," *Annual Review of Economics*, 2018.

Ashraf, Nava and Oriana Bandiera, "Altruistic capital," *American Economic Review*, 2017, *107* (5), 70–75.

_ , _ , and Alexia Delfino, "The Distinctive Values of Bankers," *AEA Papers and Proceedings*, May 2020, *110*, 167–71.

_ , _ , and B Kelsey Jack, "No margin, no mission? A field experiment on incentives for public service delivery," *Journal of public economics*, 2014, *120*, 1–17.

_ , _ , Virginia Minni, and Luigi Zingales, "Meaning at work," Technical Report, National Bureau of Economic Research 2025.

Bandiera, Oriana, Andrea Prat, Stephen Hansen, and Raffaella Sadun, "CEO behavior and firm performance," *Journal of Political Economy*, 2020, *128* (4), 1325–1369.

_ , Iwan Barankay, and Imran Rasul, "Social preferences and the response to incentives: Evidence from personnel data," *The Quarterly Journal of Economics*, 2005, *120* (3), 917–962.

_ , _ , and _ , "Incentives for managers and inequality among workers: Evidence from a firm-level experiment," *The Quarterly Journal of Economics*, 2007, *122* (2), 729–773.

_ , _ , and _ , "Team incentives: Evidence from a firm level experiment," *Journal of the European Economic Association*, 2013, *11* (5), 1079–1114.

Basil, Debra Z, Mary S Runte, M Easwaramoorthy, and Cathy Barr, "Company support for employee volunteering: A national survey of companies in Canada," *Journal of Business Ethics*, 2009, *85*, 387–398.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "Inference on treatment effects after selection among high-dimensional controls," *Review of Economic Studies*, 2014, *81* (2), 608–650.

Bénabou, Roland and Jean Tirole, "Incentives and prosocial behavior," *American economic review*, 2006, *96* (5), 1652–1678.

Benevity Inc., "State of Corporate Purpose 2024," https://benevity.com/state-of-corporate-purpose-2024 2024. Accessed: November 14, 2025.

Bertrand, Marianne and Antoinette Schoar, "Managing with style: The effect of managers on firm policies," *The Quarterly journal of economics*, 2003, *118* (4), 1169–1208.

Beunza, Daniel, "Taking the floor: Models, morals, and management in a Wall Street trading room," 2019.

Bircan, Cagatay, Guido Friebel, and Tristan Stahl, "Knowledge Teams, Careers, and Gender," *Working paper*, 2022.

Bisin, Alberto and Thierry Verdier, "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic theory*, 2001, *97* (2), 298–319.

Blasco, Andrea, Olivia S Jung, Karim R Lakhani, and Michael Menietti, "Incentives for public goods inside organizations: Field experimental evidence," *Journal of Economic Behavior & Organization*, 2019, *160*, 214–229.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen, "The organization of firms across countries," *The quarterly journal of economics*, 2012, *127* (4), 1663–1705.

Bolton, Patrick, Markus K Brunnermeier, and Laura Veldkamp, "Leadership, coordination, and corporate culture," *Review of Economic Studies*, 2013, *80* (2), 512–537.

Borghans, Lex, Bas Ter Weel, and Bruce A Weinberg, "Interpersonal styles and labor market outcomes," *Journal of Human resources*, 2008, *43* (4), 815–858.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani, "The morale effects of pay inequality," *The Quarterly Journal of Economics*, 2018, *133* (2), 611–663.

Brudney, Jeffrey L and Beth Gazley, "Moving ahead or falling behind? Volunteer promotion and data collection," *Nonprofit Management and Leadership*, 2006, *16* (3), 259–276.

CECP, "Giving in numbers, 2024 edition," https://cecp.co/wp-content/uploads/2024/11/Giving-in-Numbers-2024.pdf 2024.

Cohn, Alain, Ernst Fehr, and Michel André Maréchal, "Business culture and dishonesty in the banking industry," *Nature*, Dec 2014, *516* (7529), 86–89.

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg, "Peer effects in the workplace," *American Economic Review*, 2017, *107* (2), 425–456.

Cullen, Zoe and Ricardo Perez-Truglia, "The old boys' club: Schmoozing and the gender gap," *American Economic Review*, 2023, *113* (7), 1703–1740.

Cycyota, Cynthia S, Claudia J Ferrante, and Jessica M Schroeder, "Corporate social responsibility and employee volunteerism: What do the best companies do?," *Business Horizons*, 2016, *59* (3), 321–329.

DellaVigna, Stefano and Devin Pope, "What motivates effort? Evidence and expert forecasts," *The Review of Economic Studies*, 2018, *85* (2), 1029–1069.

_ , John A List, Ulrike Malmendier, and Gautam Rao, "Estimating social preferences and gift exchange at work," *American Economic Review*, 2022, *112* (3), 1038–1074.

Deming, David J, "The growing importance of social skills in the labor market," *The Quarterly Journal of Economics*, 2017, *132* (4), 1593–1640.

den Steen, Eric Van, "Culture clash: The costs and benefits of homogeneity," *Management Science*, 2010, *56* (10), 1718–1738.

_ , "On the origin of shared beliefs (and corporate culture)," *The RAND Journal of Economics*, 2010, *41* (4), 617–648.

Dessein, Wouter and Andrea Prat, "Organizational capital, corporate leadership, and firm dynamics," *Journal of Political Economy*, 2022, *130* (6), 1477–1536.

Dohmen, Thomas and Armin Falk, "Performance pay and multidimensional sorting: Productivity, preferences, and gender," *American economic review*, 2011, *101* (2), 556–590.

Egan, Mark, Gregor Matvos, and Amit Seru, "The market for financial adviser misconduct," *Journal of Political Economy*, 2019, *127* (1), 233–295.

Falk, Armin and Michael Kosfeld, "The hidden costs of control," *American Economic Review*, 2006, *96* (5), 1611–1630.

Fehr, Ernst and Armin Falk, "Psychological foundations of incentives," *European economic review*, 2002, *46* (4-5), 687–724.

_ **and Klaus M Schmidt**, "Fairness, incentives, and contractual choices," *European Economic Review*, 2000, *44* (4-6), 1057–1068.

Friebel, Guido, Matthias Heinz, Miriam Krueger, and Nikolay Zubanov, "Team incentives and performance: Evidence from a retail chain," *American Economic Review*, 2017, *107* (8), 2168–2203.

Gartenberg, Claudine, Andrea Prat, and George Serafeim, "Corporate purpose and financial performance," *Organization Science*, 2019, *30* (1), 1–18.

Gibbons, Robert and Michael Waldman, "Careers in organizations: Theory and evidence," *Handbook of labor economics*, 1999, *3*, 2373–2437.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini, "Performance in competitive environments: Gender differences," *The quarterly journal of economics*, 2003, *118* (3), 1049–1074.

Gosnell, Greer K, John A List, and Robert D Metcalfe, "The impact of management practices on employee productivity: A field experiment with airline captains," *Journal of Political Economy*, 2020, *128* (4), 1195–1233.

Graham, John R, Jillian Grennan, Campbell R Harvey, and Shivaram Rajgopal, "Corporate culture: Evidence from the field," *Journal of financial economics*, 2022, *146* (2), 552–593.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, June 2006, *20* (2), 23–48.

_ , _ , and _ , "The value of corporate culture," *Journal of Financial Economics*, 2015, *117* (1), 60–76. NBER Conference on the Causes and Consequences of Corporate Culture.

Hoffman, Mitchell and Steven Tadelis, "People management skills, employee attrition, and manager rewards: An empirical analysis," *Journal of Political Economy*, 2021, *129* (1), 243–285.

Holmström, Bengt, "Pay for performance and beyond," *American Economic Review*, 2017, *107* (7), 1753–1777.

i Vidal, Jordi Blanes and Mareike Nossol, "Tournaments without prizes: Evidence from personnel records," *Management science*, 2011, *57* (10), 1721–1736.

Kandel, Eugene and Edward P Lazear, "Peer pressure and partnerships," *Journal of political Economy*, 1992, *100* (4), 801–817.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental analysis of neighborhood effects," *Econometrica*, 2007, *75* (1), 83–119.

Kosse, Fabian and Michela M Tincani, "Prosociality predicts labor market success around the world," *Nature communications*, 2020, *11* (1), 5298.

Larkin, Ian, "The cost of high-powered incentives: Employee gaming in enterprise software sales," *Journal of Labor Economics*, 2014, *32* (2), 199–227.

Lavy, Victor and Analia Schlosser, "Mechanisms and impacts of gender peer effects at school," *American Economic Journal: Applied Economics*, 2011, *3* (2), 1–33.

Lazear, E., "Compensation and Incentives in the Workplace," *Journal of Economic Perspectives*, 2018.

_ **and Robert L. Moore**, "Incentives, Productivity, and Labor Contracts," *Quarterly Journal of Economics*, 1984, *99*, 275–296.

Lazear, Edward P, "Performance pay and productivity," *American Economic Review*, 2000, *90* (5), 1346–1361.

_ , "Firm-specific human capital: A skill-weights approach," *Journal of political economy*, 2009, *117* (5), 914–940.

_ **, Kathryn L Shaw, and Christopher T Stanton**, "The Value of Bosses," *Journal of Labor Economics*, 2015, *33* (4), 823–861.

List, John A, Azeem M Shaikh, and Yang Xu, "Multiple hypothesis testing in experimental economics," *Experimental Economics*, 2019, *22*, 773–793.

Mas, Alexandre and Enrico Moretti, "Peers at work," *American Economic Review*, 2009, *99* (1), 112–145.

Perez, María Jesús, JA Poyatos, Daniela Bosioc, Gabriella Civico, Kashif Khan, and Silvia Loro, "Employee volunteering and employee volunteering in humanitarian AID in Europe," *European Volunteer Centre, European Commission: Brussels, Belgium*, 2014.

Prendergast, Canice, "The provision of incentives in firms," *Journal of economic literature*, 1999, *37* (1), 7–63.

Sandvik, Jason J, Richard E Saouma, Nathan T Seegert, and Christopher T Stanton, "Workplace Knowledge Flows," *The Quarterly Journal of Economics*, 2020.

Savitskiy, Viacheslav, Raphael Mu, and Joseph Henrich, "Cultural Evolution for Economists," 2025.

Thakor, Anjan V, "Corporate culture in banking," *SSRN Electronic Journal*, 2015.

A Model Proofs

Proof of Proposition 1. Fix $\delta' < \delta''$. By assumption, for all $p \in [0,1]$,

$$G(p;\delta') \leq G(p;\delta'').$$

Let $p^*(\delta')$ denote the largest fixed point of $G(\cdot;\delta')$, so that

$$p^*(\delta') = G\big(p^*(\delta');\delta'\big).$$

Then

$$G\big(p^*(\delta');\delta''\big) \geq G\big(p^*(\delta');\delta'\big) = p^*(\delta').$$

Because $G(\cdot;\delta'')$ is continuous and nondecreasing on the compact interval $[0,1]$, the sequence defined by

$$p_{n+1} = G(p_n;\delta''), \qquad p_0 = p^*(\delta'),$$

is nondecreasing and bounded above by 1. Hence it converges to some limit $\bar{p} \in [0,1]$. By continuity,

$$\bar{p} = G(\bar{p};\delta''),$$

so $\bar{p}$ is a fixed point of $G(\cdot;\delta'')$. Moreover, since the sequence is nondecreasing and starts at $p^*(\delta')$, we have

$$\bar{p} \geq p^*(\delta').$$

Since $p^*(\delta'')$ is defined as the largest fixed point of $G(\cdot;\delta'')$, it follows that

$$p^*(\delta'') \geq \bar{p} \geq p^*(\delta').$$

Thus $p^*(\delta)$ is weakly increasing in δ. □

Proof of Corollary 1.1. Parts (i) and (ii) follow directly from Assumption 1, which implies

$$\delta(1,x) \geq \delta(0,x) \quad \text{and} \quad \delta(T,1) \geq \delta(T,0),$$

together with Proposition 1, which establishes that $p^*(\delta)$ is weakly increasing in δ.

Part (iii) follows from Assumption 2, which implies

$$\delta(1,1) - \delta(0,1) \;\geq\; \delta(1,0) - \delta(0,0).$$

Since Proposition 1 establishes that $p^*(\delta)$ is weakly increasing in δ, a larger increase in credibility yields a weakly larger increase in equilibrium culture. Therefore,

$$p^*\big(\delta(1,1)\big) - p^*\big(\delta(0,1)\big) \;\geq\; p^*\big(\delta(1,0)\big) - p^*\big(\delta(0,0)\big),$$

as required. □

Proof of Proposition 2. Fix a worker i and her cultural type s_{it}. From Equation (1), the net payoff from exerting effort is

$$\Delta U_{it}(p_t) = b + a\mathbf{1}\{s_{it} = H\} + \gamma p_t - c.$$

Since $\gamma > 0$, $\Delta U_{it}(p_t)$ is weakly increasing in p_t for fixed s_{it}.

Effort is chosen if and only if $\Delta U_{it}(p_t) \geq 0$. Therefore, if $p_{t+1} \geq p_t$,

$$\Delta U_{it}(p_{t+1}) \geq \Delta U_{it}(p_t),$$

which implies $e_{i,t+1} \geq e_{it}$. □

Proof of Corollary 2.1. Fix a worker i and hold her cultural type fixed at s_{it}.

By Proposition 2, if $p_{t+1} \geq p_t$, then

$$e_{i,t+1} \geq e_{it}.$$

Since individual output satisfies

$$y_{it} = Ae_{it} \quad \text{with } A > 0,$$

it follows immediately that

$$y_{i,t+1} \geq y_{it}.$$

Turning to exit, for a fixed cultural type s_{it}, Equation (1) implies that contemporaneous utility is weakly increasing in p_t. By assumption, the continuation value $V_i(p_t)$ is weakly increasing in contemporaneous utility. Hence $V_i(p_t)$ is weakly increasing in p_t. Therefore, if $p_{t+1} \geq p_t$,

$$V_i(p_{t+1}) \geq V_i(p_t).$$

Since exit occurs if and only if $V_i(p_t) < U_0$, a weak increase in V_i weakly reduces the likelihood of exit. □

Proof of Corollary 2.2. Fix workers' cultural types. Recall that $e_{it} \in \{0, 1\}$ and $y_{it} = Ae_{it}$ (Equation (2)). Define the individual and aggregate changes

$$\Delta y_{it} \equiv y_{i,t+1} - y_{it}, \qquad \Delta Y \equiv Y_{t+1} - Y_t = \int_0^1 \Delta y_{it}\, di.$$

By Proposition 2, $e_{i,t+1} \geq e_{it}$ for every i, hence $\Delta y_{it} \geq 0$ for all i. By assumption there exists a measurable set $S \subseteq [0, 1]$ with positive Lebesgue measure $\mu(S) > 0$ such that for every $i \in S$,

$$e_{i,t+1} > e_{it}.$$

Since effort is binary and $A > 0$, a strict increase implies $e_{it} = 0$ and $e_{i,t+1} = 1$, so for every $i \in S$,

$$\Delta y_{it} = A\big(e_{i,t+1} - e_{it}\big) = A.$$

Decompose aggregate output as

$$\Delta Y = \int_0^1 \Delta y_{it}\, di = \int_S \Delta y_{it}\, di + \int_{[0,1]\setminus S} \Delta y_{it}\, di.$$

Because $\Delta y_{it} \geq 0$ everywhere and equals A on S, we obtain

$$\Delta Y \geq \int_S A\, di = A\,\mu(S).$$

Since $A > 0$ and $\mu(S) > 0$, it follows that

$$\Delta Y > 0,$$

proving the first claim.

Now define the turnover rate at culture level p as

$$\tau(p) \equiv \int_0^1 \mathbf{1}\{V_i(p) < U_0\}\, di.$$

Then

$$\tau(p_{t+1}) - \tau(p_t) = \int_0^1 \Big(\mathbf{1}\{V_i(p_{t+1}) < U_0\} - \mathbf{1}\{V_i(p_t) < U_0\}\Big)\, di.$$

By assumption, $V_i(p_{t+1}) > V_i(p_t)$ for every $i \in S$. Suppose further that there exists a measurable subset $S' \subseteq S$ with positive Lebesgue measure $\mu(S') > 0$ such that for all $i \in S'$,

$$V_i(p_t) < U_0 \quad \text{and} \quad V_i(p_{t+1}) \geq U_0.$$

For each $i \in S'$ the integrand equals -1, while elsewhere it is ≤ 0. Hence

$$\tau(p_{t+1}) - \tau(p_t) \leq -\int_{S'} 1\, di = -\mu(S') < 0,$$

so $\tau(p_{t+1}) < \tau(p_t)$, proving the second claim. □

Proof of Proposition 3. Suppose $p_0 > p^c(\delta)$. Then, for any $p_t \in (p^c, p_H^*)$, we have

$$G(p_t; \delta) > p_t.$$

Hence

$$p_{t+1} - p_t = \lambda\big(G(p_t; \delta) - p_t\big) > 0,$$

so the sequence (p_t) is strictly increasing as long as it remains below p_H^*.

Moreover, since $G(p_t; \delta) \leq p_H^*$ whenever $p_t \leq p_H^*$, we have

$$p_{t+1} = (1-\lambda)p_t + \lambda G(p_t; \delta) \leq (1-\lambda)p_t + \lambda p_H^* < p_H^*,$$

so the sequence is bounded above by p_H^*. Therefore (p_t) is monotone increasing and bounded, and hence converges to some limit $L \le p_H^*$.

Taking limits and using continuity of G, we obtain $G(L;\delta) = L$, so L is a fixed point. Since $L \ge p^c$ and p_H^* is the only fixed point in $[p^c, 1]$, it follows that $L = p_H^*$.

The argument for $p_0 < p^c(\delta)$ is symmetric: in this case $G(p_t;\delta) < p_t$ for $p_t \in (p_L^*, p^c)$, so (p_t) is strictly decreasing and bounded below by p_L^*, implying convergence to p_L^*. □

Proof of Corollary 3.1. Under $\eta = 1$, we have $G(p_t;\delta) = \phi(p_t)$, so

$$p_{t+1} = (1-\lambda)p_t + \lambda\phi(p_t), \qquad 0 < \lambda < 1.$$

Suppose $p_0 \in \left(\frac{1}{2}, p_H^*\right)$. If $p_t < p_H^*$, then $\phi(p_t) > p_t$, implying

$$p_{t+1} - p_t = \lambda(\phi(p_t) - p_t) > 0.$$

Thus the sequence (p_t) is increasing. Moreover, since $\phi(p_t) \le p_H^*$ whenever $p_t \le p_H^*$,

$$p_{t+1} \le (1-\lambda)p_t + \lambda p_H^* < p_H^*,$$

so the sequence is bounded above by p_H^*. Monotone convergence therefore implies $p_t \to L \le p_H^*$. Taking limits and using continuity yields $\phi(L) = L$, and uniqueness of the stable interior fixed point implies $L = p_H^*$. Hence $p_t \to p_H^*$. □

Proof of Corollary 3.2. The within-office variance is

$$\mathrm{Var}(s_{it}) = p_t(1-p_t),$$

which is strictly decreasing for all $p_t > \frac{1}{2}$.

From Corollary 3.1, the sequence (p_t) is strictly increasing and satisfies $p_t > \frac{1}{2}$ for all t, converging to $p_H^* \in \left(\frac{1}{2}, 1\right)$.

Since $p_{t+1} > p_t > \frac{1}{2}$ and the function $v(p) = p(1-p)$ is strictly decreasing on this region, it follows that

$$p_{t+1}(1-p_{t+1}) < p_t(1-p_t).$$

Hence, the within-office variance strictly decreases along the convergence path. □

B Further discussion of data and empirical design

B.1 Panel

In the HR data, we observe six measurement dates: January 2018, July 2018, January 2019, July 2019, January 2020, and July 2020. Each measurement is taken at the beginning of a six-month period and is intended to summarize outcomes from the preceding semester. Thus, January 2018 and July 2018 correspond to pre-treatment periods, while January 2019

through July 2020 correspond to post-treatment periods. For ease, we refer to semesters 1 and 2 as pre-treatment (January and July 2018 measures) and semesters 3–6 as post-treatment (January 2019 to July 2020 measures).[39]

Some important remarks. First, concerning the timing of treatment, the treatment is implemented in two waves, in early July and early October 2018, and we define an individual as treated if they receive the email in either wave. Semester 3 covers July–December 2018 and is therefore partially pre- and partially post-treatment for those who are only treated in October. Rather than dropping this data, we decide to classify semester 3 as post-treatment, for three main reasons: (i) Most of the period is post-treatment for all treated employees; (ii) It is likely that, when outcomes such as performance are evaluated in January 2019, managers/HR staff place more weight on recent months than on events four to six months earlier; (iii) any remaining contamination from pre-treatment months in semester 3 for October-treated employees implies that a "treated" outcome still reflects some pre-treatment behavior, at worse downward biasing the estimated treatment effect.

Second, volunteering data record activity at the daily level. On the platform, we observe (i) the date on which an employee applies and (ii) the date on which the opportunity is scheduled to take place. While the second date is the more 'natural' measure, partners at the bank cautioned us that it may suffer from measurement error. This is also visible in the data: in some cases, the scheduled date precedes the application date, a clear inconsistency. To limit this issue, we define the "date of volunteering" as the earliest between (i) and (ii). Given the higher frequency data relative to HR, we can precisely drop any volunteering before October 2018, when the second email was sent. We then map this continuous date variable into a categorical one that reflects the six-semester structure in the HR data. Notably, since platform data are available through the summer of 2022, we classify any volunteering that occurs after the end of semester 6 as taking place in semester 6.[40]

Finally, the employee survey is administered once a year in May, and we have access to four waves of data over the study period. To align these data with the six-semester HR panel, we map each wave to the closest semester: Wave 1 is treated as semester 1 (pre-January 2018), Wave 2 as semester 2 (January–June 2018), Wave 3 as semester 4 (January–June 2019), and Wave 4 as semester 6 (January–June 2020). Although Wave 1 slightly precedes semester 1, both are pre-treatment. Under this mapping, Waves 1–2 are baseline measures while Waves 3–4 are post-treatment.

[39] In detail:

- Semester 1: up to December 2017 (measured in January 2018)
- Semester 2: January–June 2018 (measured in July 2018)
- Semester 3: July–December 2018 (measured in January 2019)
- Semester 4: January–June 2019 (measured in July 2019)
- Semester 5: July–December 2019 (measured in January 2020)
- Semester 6: January–June 2020 (measured in July 2020)

[40] This has very limited implications for the analysis: our main volunteering measure is a dummy for ever volunteering over the study window, and there are very few employees in our sample who volunteer for the first time more than two years after the start of the program and during the pandemic period.

B.2 Identifiers

B.2.1 Office

Office identifiers are constructed from the postal codes reported in HR records and in our sample. The sample was created in May 2018, while HR postcodes are observed in January and July of each year from 2018 to 2020. To reconcile these two sources and define consistent office IDs over time, we proceed as follows.

Sample postcodes

1. We start from the set of unique sample postcodes. We correct obvious typographical errors (for example, transposed characters or missing digits) and drop records coded as "Unspecified."

2. We match each postcode to geographic coordinates using the UK national postcode directory (November 2018). To aggregate individual locations into offices, we truncate the last two characters, obtaining a set of "short postcodes," some of which still contain more than one office.

3. When multiple offices share the same short postcode, we compute the pairwise distance between them. Offices located within 300 meters ($\approx$ 1000 ft) of each other are treated as a single office; offices farther apart are kept distinct. In a small number of cases, we also correct clear coding errors (e.g., postcodes corresponding to delivery depots rather than branches) by merging them with the closest operational office. In the rare cases where several offices share the same short postcode (dense urban areas), we determine groupings manually using public information on branch locations (e.g., whether a postcode corresponds to a retail branch or to a head-office building).

HR postcodes The construction of the office panel requires merging postcodes in the sample with postcodes from HR. However, 115 postcodes observed in the HR records do not appear in the sample. This reflects differences in reporting across data sources, employees moving abroad or into newly opened or relocated offices, and the fact that UK postcodes can change over time even when the underlying location remains the same. We proceed as follows:

1. We first cross-check these 115 postcodes against the UK national postcode directories. Most are valid, but 14 cannot be found. These appear either to correspond to overseas locations or to be invalid codes, and we exclude them.

2. For the remaining 101 valid UK postcodes, we attempt to match them to sample postcodes using spatial proximity:

- If an HR postcode lies within 100 meters of a sample postcode, we treat them as the same office.
- If it lies within 300 meters and shares the same first 4 digits of the postcode as the nearest sample postcode, we also treat them as the same office.
- Twelve HR postcodes share the same short postcode as sample postcodes but are located more than 300 meters away. For these cases, we conduct a manual review using publicly available information on branch locations and aggregate ten of them with the closest sample postcode. These cases are typically due to minor coding discrepancies or large postcode coverage areas.

3. Applying these criteria, we match 72 HR postcodes to sample postcodes, leaving 29 unmatched. For the 29 unmatched HR postcodes, we assign office IDs using the same geocoding and aggregation procedure as for the sample postcodes, yielding 28 additional offices.

Our final dataset contains 759 unique offices, of which 714 have an assigned treatment share (based on the sample). Over the six semesters, the number of offices is fairly stable, ranging from 677 to 714.

B.2.2 Team Identifiers

Teams are defined as departments within an office. Because updating office identifiers to ensure consistency between sample and HR postcodes also affects team definitions, we adjust team identifiers as follows:

- For semester 2, we replace the HR department number with the sample department number.
- We then align team IDs across semesters to preserve temporal continuity. Starting with semester 5, we compare each team ID with its record in the following semester (semester 6). If an individual remains in the same office across semesters, we replace the earlier semester's team ID with that of the later one. We repeat this procedure iteratively for semesters 4, 3, 2, and 1.

Our final dataset contains 8362 unique teams, of which 5419 have an assigned treatment share (based on the sample). The number of teams varies much more over time (from 5,419 to 3,385), which is consistent with the fact that people change teams more frequently and that teams are often merged or split within the firm. As a result, the team ID is a much less stable measure than the office ID.

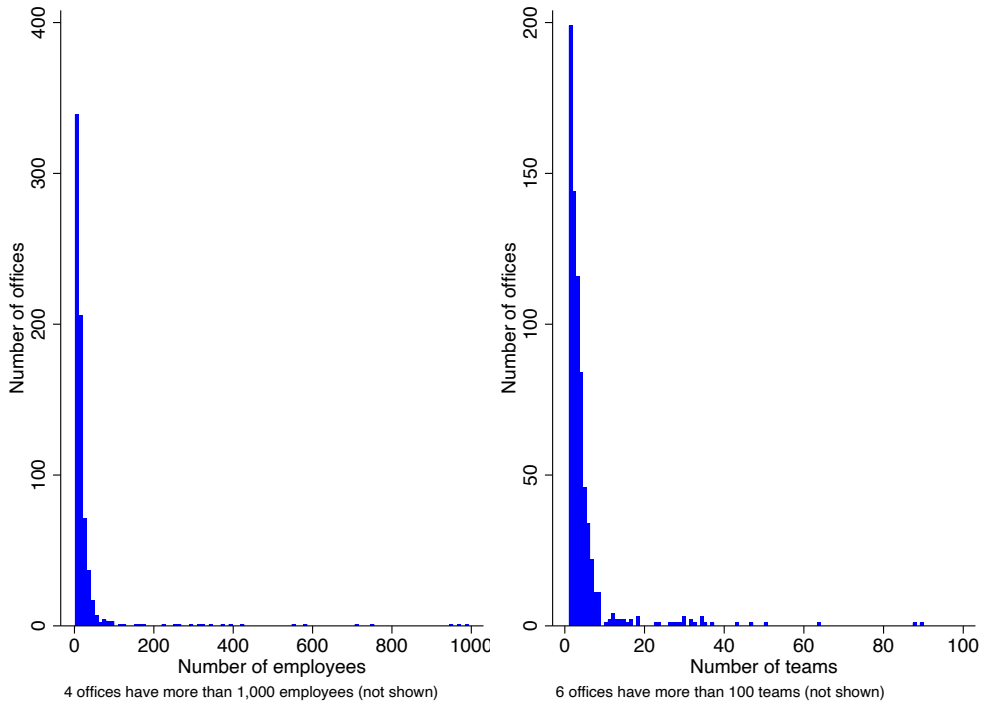
B.3 Survey Measures

Employees rate their agreement with the following statements on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree).

1. I am recognised when I do a good job.
2. I am recognised when I demonstrate [Bank Name]'s values.
3. [Bank Name] is genuine in its commitment to delivering fair outcomes for customers.
4. I feel confident that customers receive products and services that meet their needs.
5. I can challenge policies or procedures that I think will be bad for customers.
6. I am proud to say I work for [Bank Name].
7. I feel valued by [Bank Name].
8. People can state their opinions without fear of negative consequences.
9. My job conditions allow me to be as productive as I can be.
10. I feel able to achieve my career objectives at [Bank Name].
11. I see the positive impact of our strategy.
12. I feel confident about [Bank Name]'s future.
13. I trust senior leadership.
14. I trust my direct manager.
15. [Bank Name] is genuine in encouraging colleagues to speak up.
16. I feel able to speak up when I see behaviour I consider wrong.
17. *(Social responsibility)* Where I work, people have a strong sense of responsibility to society.

B.4 Stratification

The distribution of the number of teams per office, which we use for the July treatment assignment, is highly skewed.



To take this into account in our office level randomization, we created ten strata by counting the number of teams per office. In particular, stratum 1 includes offices with one team, stratum 2 with two teams, stratum 3 with three teams, and so on until stratum 6. Stratum 7 includes offices with 7 or 8 teams, stratum 8 includes offices with 9 or 10 teams, stratum 9 includes offices with between 11 and 18 teams, and stratum 10 includes offices with more than 18 teams.

For offices in strata 4 and above, we assigned them to one of five treatment groups with the following probabilities: 1/4 (25%) were assigned to the control group (where no teams received treatment), and the remaining 3/4 were evenly split among the four treatment groups (each with a probability of 3/16). For offices in the first three strata, we also assigned 1/4 (25%) to the control group. Among the remaining 3/4 (75%): in stratum 1, all offices were assigned to the 100% treated group (all teams treated); in stratum 2, all were assigned to the 50% treated group (half of the teams treated); and in stratum 3, half of the offices (3/8 of the stratum) were assigned to the 25% treated group and half (3/8) to the 75% treated group.

B.5 Seniority

The bank classifies employees into nine categories. Based on discussions with both HR and senior management, we group categories 0 to 3 as seniors, 4 to 6 as middle, and 7 to

8 as juniors. We define "Global Seniority" as an indicator equal to one for employees in categories 0 to 3.

0 = highest, 8 = lowest		
Grade	Freq.	Percent
0	41	0.12
1	71	0.21
2	709	2.09
3	2,396	7.06
4	6,467	19.07
5	6,467	19.07
6	5,720	16.87
7	6,844	20.18
8	5,199	15.33
Total	33,914	100.00

$\Longrightarrow$

3-categories		
Seniority	Freq.	Percent
Senior [0–3]	3,217	9.49
Middle [4–6]	18,654	55.00
Junior [7–8]	12,043	35.51
Total	33,914	100.00

On the other hand, to construct our "Relative Seniority" measure, we proceed as follows:

Step 1 We first collapse grades 0–3 into a single category, so that we are left with six grades in total.

Step 2 Within each office, we then identify the two most senior grades on this six-point scale and classify employees in these grades as "relative seniors"; all remaining employees in the office are treated as "relative juniors".

Step 3 To ensure that relative seniors are also reasonably senior in absolute terms (i.e. to avoid classifying junior employees as seniors simply because no true senior is present in the office), we reclassify as juniors any employees whose bank grade lies in the bottom three grades.

Step 4 Finally, we restrict attention to offices that contain at least one senior and at least one junior.

"Raw" Classification

Grade	Freq.	Percent
0	41	0.12
1	71	0.21
2	709	2.09
3	2,396	7.06
4	6,467	19.07
5	6,467	19.07
6	5,720	16.87
7	6,844	20.18
8	5,199	15.33
Total	33,914	100.00

⇓

Step 1: Collapse to 6-categories

Grade	Freq.	Percent
1	3,217	9.49
2	6,467	19.07
3	6,467	19.07
4	5,720	16.87
5	6,844	20.18
6	5,199	15.33
Total	33,914	100.00

⇓

Step 2: Seniors = top-2 grades in each office

Value	Freq.	Percent
0	21,474	63.32
1	12,440	36.68
Total	33,914	100.00

⇓

Step 3: Classify bottom 3 categories as junior

Value	Freq.	Percent
0	22,987	67.78
1	10,927	32.22
Total	33,914	100.00

⇓

Step 4: Keep Offices with ≥ 1 Senior & ≥ 1 Junior

Value	Freq.	Percent
0	21,293	66.14
1	10,902	33.86
Total	32,195	100.00

B.6 Culture Weights

All weights used in the paper are constructed exclusively from baseline survey data (semesters 1–2). Let $n_{o,\text{pre}}$ denote the number of baseline respondents in office o, and let $\text{Miss}_{o,\text{pre}}$ denote the baseline share of missing responses (i.e., non–respondents) in office o.

We consider four types of weights:

- Respondent count weight:
$$w_o^N = n_{o,\text{pre}}.$$
Offices with more baseline respondents receive greater weight.

- Effective respondent weight:
$$w_o^{\text{eff}} = n_{o,\text{pre}} \left(1 - \text{Miss}_{o,\text{pre}}\right).$$
This down–weights offices with high baseline non–response rates. An office with many respondents but a low response rate may provide less reliable information about the underlying office population.

- Squared respondent count weight:
$$w_o^{N^2} = n_{o,\text{pre}}^2.$$

- Squared effective respondent weight:
$$w_o^{\text{eff}^2} = \left[n_{o,\text{pre}} \left(1 - \text{Miss}_{o,\text{pre}}\right)\right]^2.$$
The squared weights give greater emphasis to offices with more precisely measured baseline culture, serving as a robustness check that more heavily down-weights noisier offices.

B.6.1 Testing Orthogonality to Treatment

Test 1: Raw Correlations. We first compute simple correlations between office-level treatment share and each baseline weight, separately for the Culture Fit Index and Social Responsibility.

Off. TSh.	N. Resp.	N.#Sh. Resp.	N. Resp.^{2}	N.#Sh. Resp.^{2}
Index	-0.0400	-0.0422	-0.0140	-0.0158
SR	-0.0399	-0.0421	-0.0140	-0.0160

Correlations are small and close to zero across all weight definitions. There is no indication that offices with higher treatment exposure have systematically larger baseline samples or response rates.

Test 2: Regression Balance. We formally test whether treatment intensity predicts baseline weights by estimating

$$w_o = \alpha + \beta\, o_treat_o + \lambda_s + \varepsilon_o, \tag{22}$$

where w_o denotes a baseline office-level weight, o_treat_o is the office treatment share, and λ_s are randomization-strata fixed effects. Each regression includes one weight at a time.

Table 13: Treatment Orthogonality to Baseline Belief Weights

	(1) N. Resp. b/se/p	(2) N.#Sh. Resp. b/se/p	(3) N. Resp.^{2} b/se/p	(4) N.#Sh. Resp.^{2} b/se/p
Off. TSh.	-6.9124	-3.9845	482.5053	-456.0226
	(13.0851)	(8.3192)	(3.62e+04)	(1.18e+04)
	[0.5975]	[0.6321]	[0.9894]	[0.9691]
N	714	714	714	714

Notes: Each column regresses a baseline office-level Culture Fit Index weight on office treatment share, controlling for randomization-strata fixed effects. Heteroskedasticity-robust standard errors are in parentheses and p-values in brackets. *, **, *** denote significance at the 10%, 5%, and 1% levels.

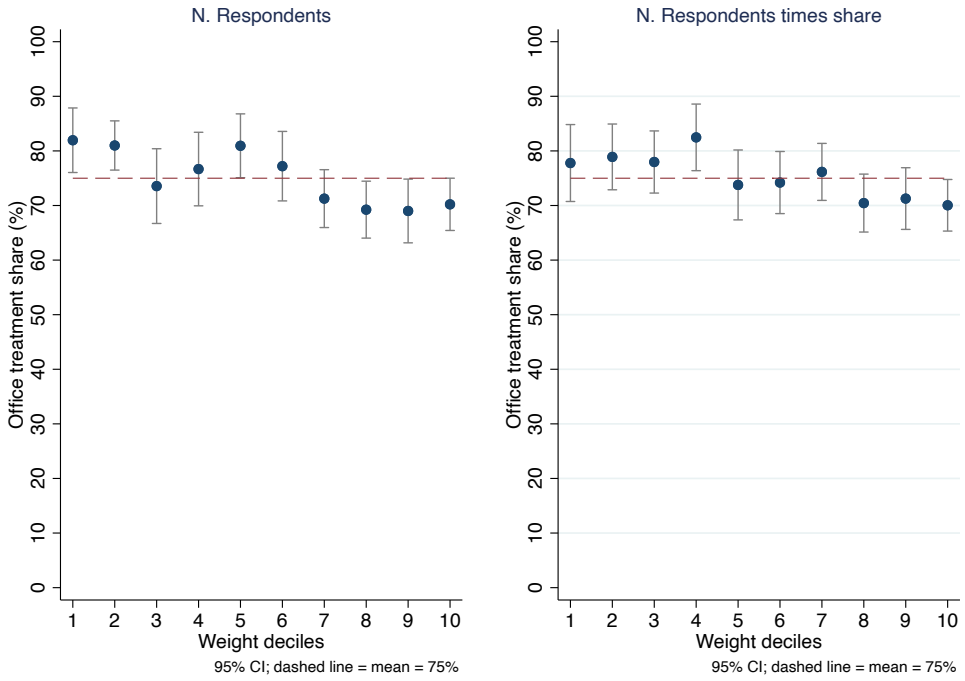
Table 14: Treatment Orthogonality to Baseline Social Responsibility Weights

	(1) N. Resp. b/se/p	(2) N.#Sh. Resp. b/se/p	(3) N. Resp.^{2} b/se/p	(4) N.#Sh. Resp.^{2} b/se/p
Off. TSh.	-6.5929	-3.7132	630.9546	-378.8638
	(12.9414)	(8.1503)	(3.47e+04)	(1.10e+04)
	[0.6106]	[0.6488]	[0.9855]	[0.9725]
N	714	714	714	714

Notes: Each column regresses a baseline office-level Social Responsibility weight on office treatment share, controlling for randomization-strata fixed effects. Heteroskedasticity-robust standard errors are in parentheses and p-values in brackets. *, **, *** denote significance at the 10%, 5%, and 1% levels.

Across specifications, treatment shares do not significantly predict baseline weights. Coefficients are small and statistically indistinguishable from zero, indicating that baseline survey coverage is orthogonal to treatment intensity.

Test 3: Binned Balance. We further group offices into deciles based on each weight definition (number of respondents and number-times-share of respondents) and plot mean treatment share with 95% confidence intervals.



Treatment shares remain close to the overall mean across deciles, with no systematic pattern. If anything, higher-weight offices exhibit slightly lower treatment exposure, implying that weighting would attenuate rather than mechanically inflate estimated treatment effects.

C Additional Tables

Table B1: Placebo Test (Main Outcomes)

	Individual			Team			Office		
	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N
Performance Rating [1-4]	0.0784	2.6749	31174	0.0355	2.7087	5226	0.1143	2.5124	714
	(0.0236)	(0.7999)		(0.0243)	(0.6013)		(0.0518)	(0.3712)	
	{ 0.1000}			{ 0.7990}			{ 0.5330}		
Low Performance Rating	-0.0037	0.0420	31174	-0.0095	0.0389	5226	0.0074	0.0848	714
	(0.0039)	(0.2006)		(0.0066)	(0.1436)		(0.0163)	(0.1075)	
	{ 0.9240}			{ 0.8480}			{ 1.0000}		
Values Rating [1-4]	0.0797**	3.0607	31058	0.0543***	3.1072	5220	0.1055	3.0144	714
	(0.0209)	(0.5484)		(0.0175)	(0.4100)		(0.0394)	(0.2376)	
	{ 0.0480}			{ 0.0010}			{ 0.2180}		
Low Values Rating	-0.0350	0.1114	31058	-0.0185	0.0881	5220	-0.0460	0.1463	714
	(0.0159)	(0.3146)		(0.0101)	(0.2093)		(0.0245)	(0.1438)	
	{ 0.4400}			{ 0.6020}			{ 0.6980}		
Exiter	-0.0151***	0.0682	34489	-0.0169	0.0654	5419	-0.0281	0.0682	715
	(0.0039)	(0.2520)		(0.0063)	(0.1809)		(0.0157)	(0.0920)	
	{ 0.0010}			{ 0.1380}			{ 0.8060}		
Beliefs Index [1-5]	0.0501	3.5789	19110	0.0592	3.5800	4446	0.0309	3.5288	692
	(0.0286)	(0.6828)		(0.0367)	(0.5380)		(0.0655)	(0.3908)	
	{ 0.6900}			{ 0.8250}			{ 1.0000}		
Social Responsibility [1-5]	0.0453	3.7717	18717	0.0980	3.7923	4404	0.1013	3.9509	688
	(0.0287)	(0.8521)		(0.0394)	(0.6608)		(0.0670)	(0.4536)	
	{ 0.8000}			{ 0.4230}			{ 0.9180}		

Notes: *All variables are measured at baseline (semester 2; semester 1 if missing). Each column reports placebo regressions of the indicated outcome on treatment. Individual and team regressions cluster standard errors at the office level; office regressions report heteroskedasticity-robust standard errors. Standard errors are shown in parentheses. List-adjusted p-values (reported in braces) are computed following List et al. (2019) using 1000 cluster wild bootstrap replications in a Romano–Wolf stepdown procedure. Significance stars are based on adjusted p-values: * p<0.10, ** p<0.05, *** p<0.01.*

Table B2: Placebo Test (Additional HR Measures)

	Individual			Team			Office		
	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N
Age [1-6]	0.1031	3.6282	25893	0.0435	3.7515	4947	0.2582	3.9111	709
	(0.0531)	(1.0869)		(0.0325)	(0.8017)		(0.0912)	(0.5873)	
	{ 0.6830}			{ 0.8540}			{ 0.1830}		
Commercial Banking	-0.0004	0.1619	34415	-0.0012	0.1763	5416	-0.0092	0.0947	715
	(0.0350)	(0.3684)		(0.0255)	(0.3790)		(0.0312)	(0.2311)	
	{ 0.9970}			{ 0.9680}			{ 1.0000}		
Retail Banking and Wealth Mgmt	0.0015	0.4352	34415	0.0591	0.4023	5416	0.0295	0.8654	715
	(0.0418)	(0.4958)		(0.0229)	(0.4888)		(0.0367)	(0.2835)	
	{ 1.0000}			{ 0.3890}			{ 0.9980}		
Global Banking and Markets	0.0349	0.1054	34415	0.0181	0.0806	5416	0.0067	0.0059	715
	(0.0291)	(0.3071)		(0.0360)	(0.2704)		(0.0068)	(0.0632)	
	{ 0.9130}			{ 0.9900}			{ 0.9930}		
Career Grade [0-8]	-0.2722	5.6252	34409	-0.0364	5.2249	5416	0.0560	6.8037	715
	(0.1756)	(1.6425)		(0.0818)	(1.3581)		(0.0816)	(0.6945)	
	{ 0.8060}			{ 0.9930}			{ 1.0000}		
Manager	0.0798	0.1726	34489	0.0598	0.1820	5419	0.0136	0.1124	715
	(0.0167)	(0.3779)		(0.0201)	(0.2761)		(0.0126)	(0.0815)	
	{ 0.4800}			{ 0.6070}			{ 0.9890}		
Tenure [1-6]	0.4082***	3.9964	34415	0.3850***	4.1440	5416	0.4145***	4.6430	715
	(0.0771)	(1.5626)		(0.0897)	(1.2268)		(0.1162)	(0.6989)	
	{ 0.0010}			{ 0.0010}			{ 0.0010}		
Office size	211.2370	2911.9860	34489	-243.9520	2509.1187	5419	-17.6355	48.2969	714
	(356.3960)	(3837.9724)		(408.1853)	(3726.0344)		(17.5090)	(372.1851)	
	{ 0.9850}			{ 0.9910}			{ 0.9870}		
Team size	-12.6360	35.0789	34489	-0.8804	6.2844	5419	-4.1670	10.1364	714
	(12.7279)	(71.6001)		(0.6843)	(13.5572)		(2.3322)	(14.3323)	
	{ 0.9530}			{ 0.9020}			{ 0.9220}		

Notes: *All variables are measured at baseline (semester 2; semester 1 if missing). Each column reports placebo regressions of the indicated outcome on treatment. Individual and team regressions cluster standard errors at the office level; office regressions report heteroskedasticity-robust standard errors. Standard errors are shown in parentheses. List-adjusted p-values (reported in braces) are computed following List et al. (2019) using 1000 cluster wild bootstrap replications in a Romano–Wolf stepdown procedure. Significance stars are based on adjusted p-values: * p<0.10, ** p<0.05, *** p<0.01.*

Table B3: Placebo Test (Additional Survey Measures)

	Individual			Team			Office		
	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N
Respondent	0.1031	3.6282	25893	0.0435	3.7515	4947	0.2582	3.9111	709
	(0.0531)	(1.0869)		(0.0325)	(0.8017)		(0.0912)	(0.5873)	
	{ 0.6830}			{ 0.8540}			{ 0.1830}		
Beliefs Index [1-5]	0.0501	3.5789	19110	0.0592	3.5800	4446	0.0309	3.5288	692
	(0.0286)	(0.6828)		(0.0367)	(0.5380)		(0.0655)	(0.3908)	
	{ 0.6900}			{ 0.8250}			{ 1.0000}		
Beliefs Index (PCA)	0.2321	-0.1763	16092	0.3462	-0.1864	4147	0.0345	-0.4513	685
	(0.1196)	(2.8553)		(0.1467)	(2.2992)		(0.2797)	(1.6748)	
	{ 0.6080}			{ 0.4750}			{ 1.0000}		
Can Achieve Career Goals [1-5]	0.0775	3.3694	18598	0.0978	3.3559	4397	0.0244	3.2600	687
	(0.0343)	(1.0565)		(0.0481)	(0.8276)		(0.0880)	(0.5660)	
	{ 0.4160}			{ 0.5530}			{ 1.0000}		
Can Challenge Policies [1-5]	0.0957	3.6010	18550	0.0571	3.6088	4391	0.0811	3.4881	686
	(0.0323)	(0.9807)		(0.0449)	(0.7629)		(0.0928)	(0.5704)	
	{ 0.3550}			{ 0.8790}			{ 0.9960}		
Confident about Bank's Future [1-5]	0.0321	3.7359	19110	0.0414	3.7392	4446	-0.0130	3.6142	692
	(0.0429)	(0.9085)		(0.0573)	(0.7075)		(0.0859)	(0.4978)	
	{ 0.9640}			{ 0.9860}			{ 1.0000}		
Fair Outcomes for Customers [1-5]	0.0490	3.9391	18550	0.0252	3.9575	4391	-0.0061	3.9264	686
	(0.0252)	(0.7737)		(0.0363)	(0.5847)		(0.0816)	(0.4418)	
	{ 0.7120}			{ 0.9890}			{ 0.9470}		
Customers' Needs are Met [1-5]	0.0365	3.8256	18550	0.0141	3.8517	4391	0.0529	3.9753	686
	(0.0288)	(0.8066)		(0.0377)	(0.6135)		(0.0845)	(0.4472)	
	{ 0.8850}			{ 0.9810}			{ 1.0000}		
See Strategy's Positive Impact [1-5]	0.0517	3.4423	19110	0.0487	3.4315	4446	-0.0091	3.4094	692
	(0.0486)	(0.8801)		(0.0548)	(0.6920)		(0.0809)	(0.4824)	
	{ 0.9230}			{ 0.9840}			{ 0.9910}		
Recognised for Values [1-5]	0.0439	3.2131	18432	0.0649	3.1944	4387	0.0178	2.9970	686
	(0.0411)	(1.1273)		(0.0576)	(0.8951)		(0.1123)	(0.6739)	
	{ 0.9180}			{ 0.8970}			{ 1.0000}		
Opinions Accepted [1-5]	0.0394	3.3065	18842	0.0507	3.3134	4415	0.0456	3.2534	689
	(0.0238)	(1.0758)		(0.0413)	(0.8355)		(0.1021)	(0.6256)	
	{ 0.6050}			{ 0.8950}			{ 1.0000}		
Able to be Productive [1-5]	0.0069	3.1111	18922	0.0279	3.0792	4424	-0.1195	3.0156	690
	(0.0355)	(1.1683)		(0.0551)	(0.9285)		(0.1055)	(0.6275)	
	{ 0.9990}			{ 0.9950}			{ 0.9900}		
Proud to Work Here [1-5]	0.0381	3.7614	18432	0.0415	3.7652	4387	-0.0296	3.6183	686
	(0.0373)	(0.9270)		(0.0539)	(0.7299)		(0.0855)	(0.5392)	
	{ 0.9320}			{ 0.9840}			{ 1.0000}		
Recognition: Good Job [1-5]	0.0583	3.4956	16098	0.1131	3.5033	4147	0.1113	3.5285	685
	(0.0313)	(1.0015)		(0.0413)	(0.7815)		(0.0918)	(0.5661)	
	{ 0.6350}			{ 0.2150}			{ 0.9860}		
Recognition: Values [1-5]	0.0756	3.5940	16098	0.1288***	3.5981	4147	0.1831	3.5332	685
	(0.0291)	(1.0117)		(0.0379)	(0.7944)		(0.0905)	(0.5785)	
	{ 0.2780}			{ 0.0010}			{ 0.6910}		
Bank Encourages Speaking Up [1-5]	0.0499	3.6780	18842	0.0818	3.6763	4415	-0.0911	3.6853	689
	(0.0281)	(0.9983)		(0.0357)	(0.7794)		(0.0959)	(0.5797)	
	{ 0.6430}			{ 0.3220}			{ 0.9930}		
Can Speak Up on Wrongdoing [1-5]	0.0439	3.9478	18842	0.0595	3.9649	4415	-0.0830	4.0223	689
	(0.0194)	(0.9329)		(0.0310)	(0.7149)		(0.0736)	(0.5000)	
	{ 0.3010}			{ 0.5880}			{ 0.9870}		
Trust: Line Mgmt [1-5]	0.0235	4.0178	18667	0.0778	3.9981	4401	0.1093	3.9852	687
	(0.0260)	(1.0072)		(0.0330)	(0.8131)		(0.0982)	(0.5762)	
	{ 0.9360}			{ 0.2250}			{ 0.9900}		
Trust: Senior Mgmt [1-5]	0.0495	3.3287	18667	0.0914	3.3223	4401	-0.0248	3.2036	687
	(0.0483)	(1.0703)		(0.0644)	(0.8550)		(0.0970)	(0.5852)	
	{ 0.9310}			{ 0.8930}			{ 1.0000}		
Minority	-0.0351	0.1479	16970	-0.0274	0.1424	4266	-0.0812	0.1129	683
	(0.0235)	(0.3550)		(0.0148)	(0.2771)		(0.0382)	(0.2093)	
	{ 0.8060}			{ 0.6890}			{ 0.6620}		
Female	0.0323	0.5173	25893	0.0402	0.4842	4947	0.0437	0.7148	709
	(0.0131)	(0.4997)		(0.0205)	(0.3897)		(0.0316)	(0.2142)	
	{ 0.3500}			{ 0.7480}			{ 0.9300}		
Customer-facing Role	-0.0180	0.4261	19216	0.0024	0.4299	4459	0.0675	0.8541	693
	(0.0354)	(0.4945)		(0.0339)	(0.4621)		(0.0348)	(0.2369)	
	{ 0.9810}			{ 0.9980}			{ 0.7020}		
How long plan to stay at bank [1-4]	0.0516	3.2757	14384	0.0610	3.2431	3992	0.0828	3.1915	660
	(0.0343)	(1.0203)		(0.0469)	(0.8221)		(0.1138)	(0.6374)	
	{ 0.8950}			{ 0.8890}			{ 1.0000}		

Notes: *All variables are measured at baseline (semester 2; semester 1 if missing). Each column reports placebo regressions of the indicated outcome on treatment. Individual and team regressions cluster standard errors at the office level; office regressions report heteroskedasticity-robust standard errors. Standard errors are shown in parentheses. List-adjusted p-values (reported in braces) are computed following List et al. (2019) using 1000 cluster wild bootstrap replications in a Romano–Wolf stepdown procedure. Significance stars are based on adjusted p-values: * p<0.10, ** p<0.05, *** p<0.01.*

Table B4: Placebo Test (Main Outcomes)

	Individual			Team			Office		
	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N
Performance Rating [1-4]	0.0784***	2.6749	31174	0.0355	2.7087	5226	0.1143**	2.5124	714
	(0.0236)	(0.7999)		(0.0243)	(0.6013)		(0.0518)	(0.3712)	
	{0.0009}			{0.1448}			{0.0277}		
Low Performance Rating	-0.0037	0.0420	31174	-0.0095	0.0389	5226	0.0074	0.0848	714
	(0.0039)	(0.2006)		(0.0066)	(0.1436)		(0.0163)	(0.1075)	
	{0.3370}			{0.1503}			{0.6513}		
Values Rating [1-4]	0.0797***	3.0607	31058	0.0543***	3.1072	5220	0.1055***	3.0144	714
	(0.0209)	(0.5484)		(0.0175)	(0.4100)		(0.0394)	(0.2376)	
	{0.0001}			{0.0020}			{0.0076}		
Low Values Rating	-0.0350**	0.1114	31058	-0.0185*	0.0881	5220	-0.0460*	0.1463	714
	(0.0159)	(0.3146)		(0.0101)	(0.2093)		(0.0245)	(0.1438)	
	{0.0275}			{0.0678}			{0.0608}		
Exiter	-0.0151***	0.0682	34489	-0.0169***	0.0654	5419	-0.0281*	0.0682	715
	(0.0039)	(0.2520)		(0.0063)	(0.1809)		(0.0157)	(0.0920)	
	{0.0001}			{0.0077}			{0.0743}		
Beliefs Index [1-5]	0.0501*	3.5789	19110	0.0592	3.5800	4446	0.0309	3.5288	692
	(0.0286)	(0.6828)		(0.0367)	(0.5380)		(0.0655)	(0.3908)	
	{0.0799}			{0.1065}			{0.6380}		
Social Responsibility [1-5]	0.0453	3.7717	18717	0.0980**	3.7923	4404	0.1013	3.9509	688
	(0.0287)	(0.8521)		(0.0394)	(0.6608)		(0.0670)	(0.4536)	
	{0.1151}			{0.0131}			{0.1311}		

Notes: *All variables are measured at baseline (semester 2; semester 1 if missing). Each column reports placebo regressions of the indicated outcome on treatment. Individual and team regressions cluster standard errors at the office level; office regressions report heteroskedasticity-robust standard errors. Standard errors are shown in parentheses. P-values (reported in braces) correspond to the regression estimates shown in each column. Significance stars are based on these p-values: * p<0.10, ** p<0.05, *** p<0.01.*

Table B5: Placebo Test (Additional HR Measures)

	Individual			Team			Office		
	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N
Age [1-6]	0.1031*	3.6282	25893	0.0435	3.7515	4947	0.2582***	3.9111	709
	(0.0531)	(1.0869)		(0.0325)	(0.8017)		(0.0912)	(0.5873)	
	{0.0524}			{0.1804}			{0.0048}		
Commercial Banking	-0.0004	0.1619	34415	-0.0012	0.1763	5416	-0.0092	0.0947	715
	(0.0350)	(0.3684)		(0.0255)	(0.3790)		(0.0312)	(0.2311)	
	{0.9899}			{0.9614}			{0.7669}		
Retail Banking and Wealth Mgmt	0.0015	0.4352	34415	0.0591**	0.4023	5416	0.0295	0.8654	715
	(0.0418)	(0.4958)		(0.0229)	(0.4888)		(0.0367)	(0.2835)	
	{0.9721}			{0.0101}			{0.4228}		
Global Banking and Markets	0.0349	0.1054	34415	0.0181	0.0806	5416	0.0067	0.0059	715
	(0.0291)	(0.3071)		(0.0360)	(0.2704)		(0.0068)	(0.0632)	
	{0.2301}			{0.6154}			{0.3292}		
Career Grade [0-8]	-0.2722	5.6252	34409	-0.0364	5.2249	5416	0.0560	6.8037	715
	(0.1756)	(1.6425)		(0.0818)	(1.3581)		(0.0816)	(0.6945)	
	{0.1214}			{0.6563}			{0.4927}		
Manager	0.0798***	0.1726	34489	0.0598***	0.1820	5419	0.0136	0.1124	715
	(0.0167)	(0.3779)		(0.0201)	(0.2761)		(0.0126)	(0.0815)	
	{0.0000}			{0.0030}			{0.2816}		
Tenure [1-6]	0.4082***	3.9964	34415	0.3850***	4.1440	5416	0.4145***	4.6430	715
	(0.0771)	(1.5626)		(0.0897)	(1.2268)		(0.1162)	(0.6989)	
	{0.0000}			{0.0000}			{0.0004}		
Office size	211.2370	2911.9860	34489	-243.9520	2509.1187	5419	-17.6355	48.2969	714
	(356.3960)	(3837.9724)		(408.1853)	(3726.0344)		(17.5090)	(372.1851)	
	{0.5536}			{0.5503}			{0.3142}		
Team size	-12.6360	35.0789	34489	-0.8804	6.2844	5419	-4.1670*	10.1364	714
	(12.7279)	(71.6001)		(0.6843)	(13.5572)		(2.3322)	(14.3323)	
	{0.3212}			{0.1987}			{0.0744}		

Notes: *All variables are measured at baseline (semester 2; semester 1 if missing). Each column reports placebo regressions of the indicated outcome on treatment. Individual and team regressions cluster standard errors at the office level; office regressions report heteroskedasticity-robust standard errors. Standard errors are shown in parentheses. P-values (reported in braces) correspond to the regression estimates shown in each column. Significance stars are based on these p-values: * p<0.10, ** p<0.05, *** p<0.01.*

Table B6: Placebo Test (Additional Survey Measures)

	Individual			Team			Office		
	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N	T (SE)	Mean (SD)	N
Respondent	0.0668***	0.5541	34489	0.0735***	0.5853	5419	0.0873***	0.4654	715
	(0.0109)	(0.4971)		(0.0148)	(0.3695)		(0.0320)	(0.2197)	
	{0.0000}			{0.0000}			{0.0066}		
Beliefs Index [1-5]	0.0501*	3.5789	19110	0.0592	3.5800	4446	0.0309	3.5288	692
	(0.0286)	(0.6828)		(0.0367)	(0.5380)		(0.0655)	(0.3908)	
	{0.0799}			{0.1065}			{0.6380}		
Beliefs Index (PCA)	0.2321*	-0.1763	16092	0.3462**	-0.1864	4147	0.0345	-0.4513	685
	(0.1196)	(2.8553)		(0.1467)	(2.2992)		(0.2797)	(1.6748)	
	{0.0528}			{0.0186}			{0.9018}		
Can Achieve Career Goals [1-5]	0.0775**	3.3694	18598	0.0978**	3.3559	4397	0.0244	3.2600	687
	(0.0343)	(1.0565)		(0.0481)	(0.8276)		(0.0880)	(0.5660)	
	{0.0243}			{0.0423}			{0.7820}		
Can Challenge Policies [1-5]	0.0957***	3.6010	18550	0.0571	3.6088	4391	0.0811	3.4881	686
	(0.0323)	(0.9807)		(0.0449)	(0.7629)		(0.0928)	(0.5704)	
	{0.0032}			{0.2035}			{0.3825}		
Confident about Bank's Future [1-5]	0.0321	3.7359	19110	0.0414	3.7392	4446	-0.0130	3.6142	692
	(0.0429)	(0.9085)		(0.0573)	(0.7075)		(0.0859)	(0.4978)	
	{0.4547}			{0.4705}			{0.8797}		
Fair Outcomes for Customers [1-5]	0.0490*	3.9391	18550	0.0252	3.9575	4391	-0.0061	3.9264	686
	(0.0252)	(0.7737)		(0.0363)	(0.5847)		(0.0816)	(0.4418)	
	{0.0519}			{0.4884}			{0.9405}		
Customers' Needs are Met [1-5]	0.0365	3.8256	18550	0.0141	3.8517	4391	0.0529	3.9753	686
	(0.0288)	(0.8066)		(0.0377)	(0.6135)		(0.0845)	(0.4472)	
	{0.2063}			{0.7094}			{0.5312}		
See Strategy's Positive Impact [1-5]	0.0517	3.4423	19110	0.0487	3.4315	4446	-0.0091	3.4094	692
	(0.0486)	(0.8801)		(0.0548)	(0.6920)		(0.0809)	(0.4824)	
	{0.2882}			{0.3740}			{0.9108}		
Recognised for Values [1-5]	0.0439	3.2131	18432	0.0649	3.1944	4387	0.0178	2.9970	686
	(0.0411)	(1.1273)		(0.0576)	(0.8951)		(0.1123)	(0.6739)	
	{0.2853}			{0.2602}			{0.8741}		
Opinions Accepted [1-5]	0.0394*	3.3065	18842	0.0507	3.3134	4415	0.0456	3.2534	689
	(0.0238)	(1.0758)		(0.0413)	(0.8355)		(0.1021)	(0.6256)	
	{0.0980}			{0.2198}			{0.6553}		
Able to be Productive [1-5]	0.0069	3.1111	18922	0.0279	3.0792	4424	-0.1195	3.0156	690
	(0.0355)	(1.1683)		(0.0551)	(0.9285)		(0.1055)	(0.6275)	
	{0.8460}			{0.6121}			{0.2577}		
Proud to Work Here [1-5]	0.0381	3.7614	18432	0.0415	3.7652	4387	-0.0296	3.6183	686
	(0.0373)	(0.9270)		(0.0539)	(0.7299)		(0.0855)	(0.5392)	
	{0.3080}			{0.4423}			{0.7296}		
Recognition: Good Job [1-5]	0.0583*	3.4956	16098	0.1131***	3.5033	4147	0.1113	3.5285	685
	(0.0313)	(1.0015)		(0.0413)	(0.7815)		(0.0918)	(0.5661)	
	{0.0634}			{0.0063}			{0.2256}		
Recognition: Values [1-5]	0.0756***	3.5940	16098	0.1288***	3.5981	4147	0.1831**	3.5332	685
	(0.0291)	(1.0117)		(0.0379)	(0.7944)		(0.0905)	(0.5785)	
	{0.0097}			{0.0007}			{0.0435}		
Bank Encourages Speaking Up [1-5]	0.0499*	3.6780	18842	0.0818**	3.6763	4415	-0.0911	3.6853	689
	(0.0281)	(0.9983)		(0.0357)	(0.7794)		(0.0959)	(0.5797)	
	{0.0763}			{0.0223}			{0.3427}		
Can Speak Up on Wrongdoing [1-5]	0.0439**	3.9478	18842	0.0595*	3.9649	4415	-0.0830	4.0223	689
	(0.0194)	(0.9329)		(0.0310)	(0.7149)		(0.0736)	(0.5000)	
	{0.0241}			{0.0549}			{0.2597}		
Trust: Line Mgmt [1-5]	0.0235	4.0178	18667	0.0778**	3.9981	4401	0.1093	3.9852	687
	(0.0260)	(1.0072)		(0.0330)	(0.8131)		(0.0982)	(0.5762)	
	{0.3678}			{0.0186}			{0.2658}		
Trust: Senior Mgmt [1-5]	0.0495	3.3287	18667	0.0914	3.3223	4401	-0.0248	3.2036	687
	(0.0483)	(1.0703)		(0.0644)	(0.8550)		(0.0970)	(0.5852)	
	{0.3050}			{0.1560}			{0.7986}		
Minority	-0.0351	0.1479	16970	-0.0274*	0.1424	4266	-0.0812**	0.1129	683
	(0.0235)	(0.3550)		(0.0148)	(0.2771)		(0.0382)	(0.2093)	
	{0.1358}			{0.0639}			{0.0339}		
Female	0.0323**	0.5173	25893	0.0402*	0.4842	4947	0.0437	0.7148	709
	(0.0131)	(0.4997)		(0.0205)	(0.3897)		(0.0316)	(0.2142)	
	{0.0136}			{0.0500}			{0.1668}		
Customer-facing Role	-0.0180	0.4261	19216	0.0024	0.4299	4459	0.0675*	0.8541	693
	(0.0354)	(0.4945)		(0.0339)	(0.4621)		(0.0348)	(0.2369)	
	{0.6113}			{0.9428}			{0.0529}		
How long plan to stay at bank [1-4]	0.0516	3.2757	14384	0.0610	3.2431	3992	0.0828	3.1915	660
	(0.0343)	(1.0203)		(0.0469)	(0.8221)		(0.1138)	(0.6374)	
	{0.1325}			{0.1932}			{0.4673}		

Notes: *All variables are measured at baseline (semester 2; semester 1 if missing). Each column reports placebo regressions of the indicated outcome on treatment. Individual and team regressions cluster standard errors at the office level; office regressions report heteroskedasticity-robust standard errors. Standard errors are shown in parentheses. P-values (reported in braces) correspond to the regression estimates shown in each column. Significance stars are based on these p-values: * p<0.10, ** p<0.05, *** p<0.01.*

Table B7: Summary Statistics: Main Variables (Semester Averages)

		Individual					Team					Office				
		Mean	Sd.	Min.	Max	N	Mean	Sd.	Min.	Max	N	Mean	Sd.	Min.	Max	N
Volunteering	Volunteering	0.07	0.11	0.00	1.00	33729	0.08	0.14	0.00	1.00	6824	0.12	0.07	0.00	0.43	714
Main Outcomes	Performance Rating	2.69	0.63	1.00	4.00	33691	2.75	0.53	1.00	4.00	8128	2.58	0.29	1.33	4.00	755
	Value Rating	3.07	0.43	1.00	4.00	33670	3.13	0.36	1.00	4.00	8120	3.06	0.19	2.00	4.00	755
	Low Performance	0.04	0.14	0.00	1.00	33691	0.03	0.11	0.00	1.00	8128	0.07	0.07	0.00	0.67	755
	Low Values	0.11	0.24	0.00	1.00	33670	0.07	0.17	0.00	1.00	8120	0.12	0.11	0.00	1.00	755
	Exit	0.07	0.18	0.00	1.00	34498	0.04	0.12	0.00	1.00	8350	0.03	0.05	0.00	1.00	759
Culture	Beliefs Index	0.10	0.90	-3.88	2.35	25830	0.12	0.76	-3.80	2.35	6767	0.17	0.45	-2.45	1.77	744
	Social Responsibility	3.77	0.76	1.00	5.00	23090	3.79	0.62	1.00	5.00	5888	3.93	0.40	1.00	5.00	733

Table B8: Summary Statistics: Survey Variables (Semester Averages)

		Individual					Team					Office				
		Mean	Sd.	Min.	Max	N	Mean	Sd.	Min.	Max	N	Mean	Sd.	Min.	Max	N
Culture	Index	3.61	0.62	1.00	5.00	25830	3.63	0.52	1.00	5.00	6767	3.65	0.31	1.88	4.75	744
	Standardized Index	0.10	0.90	-3.88	2.35	25830	0.12	0.76	-3.80	2.35	6767	0.17	0.45	-2.45	1.77	744
	PCA	-0.04	2.60	-11.17	5.66	24549	0.03	2.20	-11.17	5.66	6406	0.13	1.27	-7.41	4.62	739
	Social Responsibility	3.77	0.76	1.00	5.00	23090	3.79	0.62	1.00	5.00	5888	3.93	0.40	1.00	5.00	733
	I feel valued by [Bank Name]	3.34	0.96	1.00	5.00	25526	3.35	0.81	1.00	5.00	6726	3.32	0.48	1.00	5.00	744
	I trust senior lead	3.41	0.91	1.00	5.00	25615	3.40	0.80	1.00	5.00	6738	3.40	0.44	1.00	5.00	744
	I trust my direct manager	4.05	0.84	1.00	5.00	25615	4.05	0.74	1.00	5.00	6738	4.07	0.45	1.00	5.00	744
	I can speak up when I see wrong behaviour	3.86	0.80	1.00	5.00	25687	3.89	0.66	1.00	5.00	6746	3.95	0.39	2.00	5.00	744
	Bank encourages colleagues to speak up	3.68	0.84	1.00	5.00	25687	3.69	0.71	1.00	5.00	6746	3.73	0.41	1.00	5.00	744
	Recognised when I demonstrate values	3.67	0.87	1.00	5.00	24552	3.69	0.73	1.00	5.00	6406	3.73	0.39	1.00	5.00	739
	Recognised when I do a good job	3.60	0.86	1.00	5.00	24552	3.61	0.72	1.00	5.00	6406	3.72	0.39	1.00	5.00	739
	Proud to say I work here	3.77	0.82	1.00	5.00	25526	3.77	0.69	1.00	5.00	6726	3.76	0.39	2.00	5.00	744
	Conditions allow me to be productive	3.22	0.96	1.00	5.00	25728	3.23	0.83	1.00	5.00	6752	3.24	0.44	1.00	4.50	744
	People can state opinions without fear	3.42	0.93	1.00	5.00	25687	3.46	0.78	1.00	5.00	6746	3.49	0.48	1.00	5.00	744
	Seeing the positive impact of our strategy	3.41	0.76	1.00	5.00	25830	3.40	0.66	1.00	5.00	6767	3.47	0.34	2.00	5.00	744
	Customers needs are met	3.87	0.68	1.00	5.00	25575	3.88	0.56	1.00	5.00	6730	4.04	0.31	3.00	5.00	744
	Delivers fair outcomes for customers	3.97	0.65	1.00	5.00	25575	4.01	0.54	1.00	5.00	6730	4.03	0.31	2.00	5.00	744
	Confident about [Bank Name] future	3.65	0.79	1.00	5.00	25830	3.65	0.68	1.00	5.00	6767	3.64	0.37	2.00	5.00	744
	Able to challenge policies/procedures	3.63	0.82	1.00	5.00	25575	3.66	0.68	1.00	5.00	6730	3.60	0.40	1.00	5.00	744
	Able to achieve career objectives	3.37	0.91	1.00	5.00	25593	3.41	0.77	1.00	5.00	6736	3.38	0.41	1.00	5.00	744
Other	Ethnicity: white	0.85	0.35	0.00	1.00	16977	0.85	0.28	0.00	1.00	6580	0.88	0.21	0.00	1.00	736
	Ethnicity: black	0.02	0.13	0.00	1.00	16977	0.02	0.10	0.00	1.00	6580	0.02	0.08	0.00	1.00	736
	Ethnicity: asian	0.11	0.31	0.00	1.00	16977	0.11	0.25	0.00	1.00	6580	0.09	0.18	0.00	1.00	736
	Foreign Born	0.14	0.35	0.00	1.00	17339	0.14	0.28	0.00	1.00	6591	0.09	0.18	0.00	1.00	734
	Female	0.52	0.50	0.00	1.00	25900	0.48	0.39	0.00	1.00	7690	0.70	0.22	0.00	1.00	754
	Customer-facing role	0.45	0.48	0.00	1.00	25878	0.38	0.45	0.00	1.00	6773	0.84	0.24	0.00	1.00	745

Table B9: Peer-exposure heterogeneous effects on volunteering

DV: Volunteered	(1) Office	(2) Team
Treat	-0.0474	0.1052***
	(0.0765)	(0.0253)
	[0.5355]	[0.0000]
Treat share (LOO)	-0.1161	0.0230
	(0.0915)	(0.0368)
	[0.2051]	[0.5318]
Treat × Treat share (LOO)	0.1744	-0.0690
	(0.1157)	(0.0539)
	[0.1322]	[0.2009]
Observations	33711	31809

Notes: *Each column reports estimates from regressions of volunteering on individual treatment, the leave-one-out share of treated peers within the office (Column 1) or team (Column 2), and their interaction. Peer exposure is defined as the fraction of other members in the same unit assigned to treatment. All specifications include randomization strata fixed effects. Standard errors are clustered at the office level and reported in parentheses, p-values in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

Table B10: Outcome Variable: Registered

	T(SE)[P]	$\overline{Y}(SD)$	N
Individual	0.1079***	0.4840	34489
	(0.0162)	(0.4998)	
	[0.0000]		
Team	0.1084***	0.5919	5419
	(0.0246)	(0.3690)	
	[0.0000]		
Office	0.0823***	0.6209	714
	(0.0288)	(0.1894)	
	[0.0044]		

Notes: *SE clustered by office for individual and team level regressions; heteroskedasticity robust SE for office-level regressions. Strata FE. $\bar{Y}$ is the control mean (individual level) or the sample mean (team/office level, where treatment is a share). *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Table B11: On-the-job Outcomes (Robustness: Drop transfers to offices not in the sample)

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_{preT}}(SD)$	N	T(SE)[P]	$\overline{Y_{preT}}(SD)$	N
Performance	0.0178***	2.6877	95870	0.0343**	2.7092	13923	0.0867***	2.5319	2716
	(0.0069)	(0.7957)		(0.0170)	(0.5997)		(0.0322)	(0.3671)	
	[0.0098]			[0.0440]			[0.0073]		
Values	0.0245***	3.0438	95340	0.0247**	3.1056	13898	0.0569**	3.0376	2716
	(0.0072)	(0.5429)		(0.0114)	(0.4071)		(0.0236)	(0.2392)	
	[0.0007]			[0.0311]			[0.0160]		
Low Performance	-0.0045**	0.0378	95870	-0.0098*	0.0403	13923	-0.0203**	0.0792	2716
	(0.0019)	(0.1907)		(0.0054)	(0.1476)		(0.0101)	(0.1075)	
	[0.0178]			[0.0706]			[0.0435]		
Low Values	-0.0089	0.1154	95340	-0.0092	0.0890	13898	-0.0217	0.1338	2716
	(0.0056)	(0.3195)		(0.0083)	(0.2105)		(0.0155)	(0.1443)	
	[0.1103]			[0.2656]			[0.1615]		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. All specifications include semester fixed effects and randomization strata fixed effects and control for baseline values of the outcome (measured in semester 2; if unavailable, semester 1). Control group means and standard deviations at the individual level are computed over the post-treatment period (timeID $\geq$ 3). At the team and office levels, reported means and standard deviations correspond to pre-treatment averages. Low performance and low values are indicator variables. The estimation sample excludes individuals who at any point move to an office that did not exist in the baseline sample. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

Table B12: On-the-job Outcomes (Robustness: Keep only survey respondents)

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_{preT}}(SD)$	N	T(SE)[P]	$\overline{Y_{preT}}(SD)$	N
Performance	0.0187***	2.7179	84262	0.0460***	2.7207	13329	0.0847***	2.5319	2704
	(0.0067)	(0.7900)		(0.0160)	(0.5787)		(0.0316)	(0.3647)	
	[0.0051]			[0.0042]			[0.0076]		
Values	0.0259***	3.0680	83845	0.0280**	3.1128	13302	0.0625***	3.0371	2704
	(0.0068)	(0.5320)		(0.0119)	(0.3920)		(0.0230)	(0.2352)	
	[0.0001]			[0.0185]			[0.0066]		
Low Performance	-0.0039**	0.0323	84262	-0.0089	0.0356	13329	-0.0197*	0.0795	2704
	(0.0017)	(0.1767)		(0.0056)	(0.1303)		(0.0102)	(0.1076)	
	[0.0219]			[0.1117]			[0.0534]		
Low Values	-0.0086	0.1004	83845	-0.0117	0.0843	13302	-0.0284*	0.1335	2704
	(0.0055)	(0.3006)		(0.0095)	(0.1985)		(0.0150)	(0.1425)	
	[0.1160]			[0.2189]			[0.0582]		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. All specifications include semester fixed effects and randomization strata fixed effects and control for baseline values of the outcome (measured in semester 2; if unavailable, semester 1). Control group means and standard deviations at the individual level are computed over the post-treatment period (timeID $\geq$ 3). At the team and office levels, reported means and standard deviations correspond to pre-treatment averages. Low performance and low values are indicator variables. Excludes employees that never answered to the survey. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

Table B13: On-the-job Performance (Robustness: PDS Lasso)

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_{preT}}(SD)$	N	T(SE)[P]	$\overline{Y_{preT}}(SD)$	N
Performance	0.0076	2.7113	105357	0.0134	2.7140	14412	0.0749**	2.5334	2713
	(0.0066)	(0.7942)		(0.0152)	(0.6015)		(0.0303)	(0.3717)	
	[0.2498]			[0.3779]			[0.0134]		
Values	0.0148***	3.0643	105107	0.0147	3.1077	14402	0.0630***	3.0375	2713
	(0.0048)	(0.5272)		(0.0101)	(0.4077)		(0.0209)	(0.2414)	
	[0.0018]			[0.1464]			[0.0026]		
Low Performance	0.0005	0.0342	105357	-0.0071	0.0389	14412	-0.0198**	0.0787	2713
	(0.0015)	(0.1817)		(0.0053)	(0.1450)		(0.0097)	(0.1074)	
	[0.7530]			[0.1794]			[0.0407]		
Low Values	-0.0045	0.0994	105107	-0.0098	0.0872	14402	-0.0311**	0.1336	2713
	(0.0052)	(0.2992)		(0.0076)	(0.2097)		(0.0133)	(0.1460)	
	[0.3950]			[0.1955]			[0.0191]		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. All specifications include semester fixed effects and randomization strata fixed effects and control for baseline values of the outcome (measured in semester 2; if unavailable, semester 1). Control group means and standard deviations at the individual level are computed over the post-treatment period (timeID $\geq$ 3). At the team and office levels, reported means and standard deviations correspond to pre-treatment averages. Low performance and low values are indicator variables. LASSO covariates consist of baseline measures for function, being a manager according to HR, senior, junior, full-time employment status, performance, values, time FE and strata FE, along with their interactions. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

Table B14: On-the-job Performance, MHT (1)

	Individual			Team			Office		
	T(SE)[P]	Mean(SD)	N	T(SE)[P]	Mean(SD)	N	T(SE)[P]	Mean(SD)	N
Performance	0.0170***	2.7031	105733	0.0307**	2.7140	14358	0.0866***	2.5334	2716
	(0.0063)	(0.7919)		(0.0168)	(0.6015)		(0.0322)	(0.3717)	
	{0.0010}			{0.0310}			{0.0010}		
Values	0.0229***	3.0550	105165	0.0216*	3.1077	14331	0.0572***	3.0375	2716
	(0.0064)	(0.5318)		(0.0111)	(0.4077)		(0.0236)	(0.2414)	
	{0.0010}			{0.0850}			{0.0010}		
Low Performance	-0.0038**	0.0344	105733	-0.0090	0.0389	14358	-0.0205	0.0787	2716
	(0.0016)	(0.1822)		(0.0052)	(0.1450)		(0.0100)	(0.1074)	
	{0.0270}			{0.1530}			{0.1020}		
Low Values	-0.0070	0.1055	105165	-0.0088	0.0872	14331	-0.0219*	0.1336	2716
	(0.0049)	(0.3072)		(0.0087)	(0.2097)		(0.0155)	(0.1460)	
	{0.1130}			{0.3650}			{0.0690}		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. All specifications include semester fixed effects and randomization strata fixed effects and control for baseline values of the outcome (measured in semester 2; if unavailable, semester 1). Control group means and standard deviations at the individual level are computed over the post-treatment period (timeID $\geq$ 3). At the team and office levels, reported means and standard deviations correspond to pre-treatment averages. Low performance and low values are indicator variables. Standard errors are clustered at the office level and reported in parentheses. List-adjusted p-values (reported in braces) are computed separately within each level of aggregation across the four outcome variables using 1000 cluster wild bootstrap replications in a Romano–Wolf stepdown procedure following List et al. (2019). Significance stars are based on List-adjusted p-values. * p<0.10, ** p<0.05, *** p<0.01.*

Table B15: On-the-job Performance, MHT (2)

	Individual			Team			Office		
	T(SE)[P]	Mean(SD)	N	T(SE)[P]	Mean(SD)	N	T(SE)[P]	Mean(SD)	N
Performance	0.0170***	2.7031	105733	0.0307**	2.7140	14358	0.0866**	2.5334	2716
	(0.0063)	(0.7919)		(0.0168)	(0.6015)		(0.0322)	(0.3717)	
	{0.0010}			{0.0210}			{0.0170}		
Values	0.0229***	3.0550	105165	0.0216*	3.1077	14331	0.0572***	3.0375	2716
	(0.0064)	(0.5318)		(0.0111)	(0.4077)		(0.0236)	(0.2414)	
	{0.0010}			{0.0850}			{0.0010}		
Low Performance	-0.0038**	0.0344	105733	-0.0090	0.0389	14358	-0.0205*	0.0787	2716
	(0.0016)	(0.1822)		(0.0052)	(0.1450)		(0.0100)	(0.1074)	
	{0.0410}			{0.2080}			{0.0840}		
Low Values	-0.0070	0.1055	105165	-0.0088	0.0872	14331	-0.0219*	0.1336	2716
	(0.0049)	(0.3072)		(0.0087)	(0.2097)		(0.0155)	(0.1460)	
	{0.1130}			{0.3540}			{0.0520}		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. All specifications include semester fixed effects and randomization strata fixed effects and control for baseline values of the outcome (measured in semester 2; if unavailable, semester 1). Control group means and standard deviations at the individual level are computed over the post-treatment period (timeID $\geq$ 3). At the team and office levels, reported means and standard deviations correspond to pre-treatment averages. Low performance and low values are indicator variables. Standard errors are clustered at the office level and reported in parentheses. List-adjusted p-values (reported in braces) are computed computed jointly across all 12 hypotheses (4 outcomes $\times$ 3 aggregation levels) using 1000 cluster wild bootstrap replications in a Romano–Wolf stepdown procedure following List et al. (2019). Significance stars are based on List-adjusted p-values. * p<0.10, ** p<0.05, *** p<0.01.*

Table B16: Turnover and Disengagement (Robustness: Drop transfers to offices not in sample)

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N
Exiter	-0.0268**	0.1590	29817	-0.0141	0.0365	5063	-0.0622***	0.1282	713
	(0.0115)	(0.3657)		(0.0152)	(0.0957)		(0.0214)	(0.1233)	
	[0.0201]			[0.3508]			[0.0038]		
Disengagement: Quiet Quitting									
Performance	-0.0260***	0.4699	28619	-0.0305**	2.6985	4985	-0.0770***	0.5206	713
	(0.0058)	(0.4991)		(0.0127)	(0.6096)		(0.0294)	(0.2049)	
	[0.0000]			[0.0169]			[0.0090]		
Values	-0.0240***	0.2004	28509	-0.0422***	3.1025	4976	-0.0543**	0.2042	713
	(0.0063)	(0.4003)		(0.0163)	(0.4170)		(0.0235)	(0.1537)	
	[0.0002]			[0.0098]			[0.0209]		
Performance or Values	-0.0193***	0.4743	28496	-0.0261**	5.8006	4973	-0.0717**	0.5237	713
	(0.0061)	(0.4994)		(0.0131)	(0.9146)		(0.0289)	(0.2051)	
	[0.0017]			[0.0465]			[0.0133]		

Notes: *SE clustered by office for individual and team level regressions; heteroskedasticity robust SE for office-level regressions. Strata FE included. $\bar{Y}$ is the control mean (individual level) or the pre-treatment mean (team/office level, where treatment is a share). Excludes individuals that at any point moved to an office that does not belong to the sample. Controls for baseline values in qq-values, baseline performance in qq-performance and baseline values + performance in qq-performance-values. Control for baseline turnover in team and office level regressions. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Table B17: Turnover and Disengagement (Robustness: Keep only survey respondents)

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N
Exiter	-0.0105	0.1124	24765	-0.0014	0.0226	4868	-0.0312*	0.0858	707
	(0.0082)	(0.3158)		(0.0099)	(0.0757)		(0.0188)	(0.1121)	
	[0.1997]			[0.8915]			[0.0982]		
Disengagement: Quiet Quitting									
Performance	-0.0188***	0.4167	23828	-0.0380***	2.7379	4793	-0.0709**	0.4652	707
	(0.0063)	(0.4931)		(0.0118)	(0.6064)		(0.0344)	(0.2272)	
	[0.0028]			[0.0013]			[0.0398]		
Values	-0.0091	0.1380	23754	-0.0134	3.1265	4787	-0.0239	0.1303	707
	(0.0079)	(0.3449)		(0.0092)	(0.4146)		(0.0234)	(0.1453)	
	[0.2482]			[0.1475]			[0.3090]		
Performance or Values	-0.0116*	0.4206	23746	-0.0345***	5.8647	4783	-0.0717**	0.4684	707
	(0.0063)	(0.4937)		(0.0114)	(0.9054)		(0.0337)	(0.2276)	
	[0.0682]			[0.0025]			[0.0336]		

Notes: *SE clustered by office for individual and team level regressions; heteroskedasticity robust SE for office-level regressions. Strata FE included. $\bar{Y}$ is the control mean (individual level) or the pre-treatment mean (team/office level, where treatment is a share). Excludes individuals with all waves of beliefs missing. Controls for baseline values in qq-values, baseline performance in qq-performance and baseline values + performance in qq-performance-values. Control for baseline turnover in team and office level regressions. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Table B18: Turnover and Disengagement (Robustness: Keep only survey respondents)

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N
Exiter	-0.0105	0.1124	24765	-0.0014	0.0226	4868	-0.0312*	0.0858	707
	(0.0082)	(0.3158)		(0.0099)	(0.0757)		(0.0188)	(0.1121)	
	[0.1997]			[0.8915]			[0.0982]		
Disengagement: Quiet Quitting									
Performance	-0.0188***	0.4167	23828	-0.0380***	2.7379	4793	-0.0709**	0.4652	707
	(0.0063)	(0.4931)		(0.0118)	(0.6064)		(0.0344)	(0.2272)	
	[0.0028]			[0.0013]			[0.0398]		
Values	-0.0091	0.1380	23754	-0.0134	3.1265	4787	-0.0239	0.1303	707
	(0.0079)	(0.3449)		(0.0092)	(0.4146)		(0.0234)	(0.1453)	
	[0.2482]			[0.1475]			[0.3090]		
Performance or Values	-0.0116*	0.4206	23746	-0.0345***	5.8647	4783	-0.0717**	0.4684	707
	(0.0063)	(0.4937)		(0.0114)	(0.9054)		(0.0337)	(0.2276)	
	[0.0682]			[0.0025]			[0.0336]		

Notes: *SE clustered by office for individual and team level regressions; heteroskedasticity robust SE for office-level regressions. Strata FE included. $\bar{Y}$ is the control mean (individual level) or the pre-treatment mean (team/office level, where treatment is a share). Excludes individuals with all waves of beliefs missing. Controls for baseline values in qq-values, baseline performance in qq-performance and baseline values + performance in qq-performance-values. Control for baseline turnover in team and office level regressions. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Table B19: Turnover and Disengagement, MHT (1)

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N
Exiter	-0.0195	0.1466	32456	-0.0007	0.0348	5317	-0.0607***	0.1261	713
	(0.0109)	(0.3537)		(0.0102)	(0.0922)		(0.0207)	(0.1211)	
	{ 0.1500}			{ 0.9600}			{ 0.0050}		
Disengagement: Quiet Quitting									
Performance	-0.0207***	0.4527	31174	-0.0257**	2.7087	5226	-0.0857***	0.5190	713
	(0.0051)	(0.4978)		(0.0110)	(0.6013)		(0.0274)	(0.2027)	
	{ 0.0010}			{ 0.0400}			{ 0.0010}		
Values	-0.0152*	0.1809	31058	-0.0276**	3.1072	5220	-0.0516**	0.2014	713
	(0.0066)	(0.3850)		(0.0100)	(0.4100)		(0.0232)	(0.1526)	
	{ 0.0600}			{ 0.0390}			{ 0.0240}		
Performance or Values	-0.0139**	0.4567	31043	-0.0206*	5.8162	5217	-0.0800***	0.5220	713
	(0.0052)	(0.4981)		(0.0106)	(0.8989)		(0.0272)	(0.2029)	
	{ 0.0250}			{ 0.0820}			{ 0.0070}		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. Individual- and team-level regressions control for the baseline value of the outcome and include randomization strata fixed effects; office-level regressions include strata fixed effects and control for baseline office averages where applicable. Standard errors are clustered at the office level for individual and team regressions and are heteroskedasticity-robust for office-level regressions. Control group means and standard deviations at the individual level are computed over the post-treatment period; at the team and office levels they correspond to pre-treatment averages. List-adjusted p-values (reported in braces) are computed separately within each aggregation level across the four outcome variables using 1000 cluster wild bootstrap replications in a Romano–Wolf stepdown procedure following List et al. (2019). Significance stars are based on these list-adjusted p-values: * p<0.10, ** p<0.05, *** p<0.01.*

Table B20: Turnover and Disengagement, MHT (2)

	Individual			Team			Office		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N	T(SE)[P]	$\overline{Y}(SD)$	N
Exiter	-0.0195	0.1466	32456	-0.0007	0.0348	5317	-0.0607***	0.1261	713
	(0.0109)	(0.3537)		(0.0102)	(0.0922)		(0.0207)	(0.1211)	
	{ 0.1460}			{ 0.9560}			{ 0.0040}		
Disengagement: Quiet Quitting									
Performance	-0.0207***	0.4527	31174	-0.0257**	2.7087	5226	-0.0857***	0.5190	713
	(0.0051)	(0.4978)		(0.0110)	(0.6013)		(0.0274)	(0.2027)	
	{ 0.0010}			{ 0.0190}			{ 0.0050}		
Values	-0.0152*	0.1809	31058	-0.0276**	3.1072	5220	-0.0516**	0.2014	713
	(0.0066)	(0.3850)		(0.0100)	(0.4100)		(0.0232)	(0.1526)	
	{ 0.0760}			{ 0.0260}			{ 0.0180}		
Performance or Values	-0.0139**	0.4567	31043	-0.0206*	5.8162	5217	-0.0800***	0.5220	713
	(0.0052)	(0.4981)		(0.0106)	(0.8989)		(0.0272)	(0.2029)	
	{ 0.0220}			{ 0.0550}			{ 0.0040}		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. Individual- and team-level regressions control for the baseline value of the outcome and include randomization strata fixed effects; office-level regressions include strata fixed effects and control for baseline office averages where applicable. Standard errors are clustered at the office level for individual and team regressions and are heteroskedasticity-robust for office-level regressions. Control group means and standard deviations at the individual level are computed over the post-treatment period; at the team and office levels they correspond to pre-treatment averages. List-adjusted p-values (reported in braces) are computed jointly across all 12 hypotheses (four outcomes × three aggregation levels) using 1000 cluster wild bootstrap replications in a Romano–Wolf stepdown procedure following List et al. (2019). Significance stars are based on these list-adjusted p-values: * p<0.10, ** p<0.05, *** p<0.01.*

Table B21: Culture Fit (Robustness: Sample restriction and PDS LASSO)

	(1) FE + Baseline			(2) = (1), restricted sample			(3) PDS LASSO		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_C}(SD)$	N
Culture index (std.)	0.0704***	0.0777	20492	0.0821***	0.1049	18272	0.0842***	0.1036	27278
	(0.0200)	(0.9846)		(0.0201)	(0.9870)		(0.0273)	(0.9673)	
	[0.0005]			[0.0001]			[0.0020]		
Culture index (PCA)	0.2180***	-0.1921	18016	0.2560***	-0.1154	16209	0.2415***	-0.1250	27278
	(0.0570)	(2.8488)		(0.0569)	(2.8609)		(0.0781)	(2.7704)	
	[0.0001]			[0.0000]			[0.0020]		
Social responsibility	0.1108***	-0.0228	9980	0.1142***	-0.0028	8889	0.0979***	-0.0140	13143
	(0.0382)	(0.9631)		(0.0401)	(0.9719)		(0.0347)	(0.9465)	
	[0.0038]			[0.0045]			[0.0048]		
Can be productive in my job	0.0646***	0.1700	20375	0.0715***	0.1788	18163	0.0478*	0.2031	27278
	(0.0242)	(0.9284)		(0.0247)	(0.9324)		(0.0250)	(0.9119)	
	[0.0077]			[0.0040]			[0.0559]		
Can achieve career goals	0.0561**	-0.0147	20104	0.0655***	-0.0179	17918	0.0571***	-0.0066	27278
	(0.0218)	(0.9714)		(0.0228)	(0.9787)		(0.0196)	(0.9654)	
	[0.0103]			[0.0042]			[0.0036]		
Positive impact strategy	0.0687***	-0.0450	20492	0.0744***	-0.0073	18272	0.0590**	-0.0049	27278
	(0.0241)	(0.9966)		(0.0246)	(0.9852)		(0.0266)	(0.9724)	
	[0.0045]			[0.0025]			[0.0270]		
Confident about bank's future	0.0395**	-0.0240	20492	0.0415**	0.0095	18272	0.0440**	-0.0109	27278
	(0.0171)	(0.9731)		(0.0192)	(0.9573)		(0.0202)	(0.9465)	
	[0.0210]			[0.0308]			[0.0290]		
Trust senior leadership	0.0523***	0.0787	20169	0.0472**	0.1250	17975	0.0733***	0.1002	27278
	(0.0186)	(0.9531)		(0.0192)	(0.9395)		(0.0260)	(0.9263)	
	[0.0050]			[0.0144]			[0.0048]		
Trust direct manager	0.0531***	0.0122	20169	0.0616***	0.0156	17975	0.0496***	0.0309	27278
	(0.0184)	(0.9806)		(0.0198)	(0.9852)		(0.0165)	(0.9728)	
	[0.0040]			[0.0020]			[0.0026]		
Bank encourages speak-up	0.0775***	0.0240	20319	0.0821***	0.0420	18110	0.0799***	0.0296	27278
	(0.0225)	(0.9481)		(0.0244)	(0.9497)		(0.0197)	(0.9362)	
	[0.0006]			[0.0008]			[0.0001]		
Can speak up on concerns	0.0577***	-0.1403	20319	0.0595***	-0.1335	18110	0.0577***	-0.1304	27278
	(0.0202)	(0.9892)		(0.0209)	(1.0060)		(0.0152)	(0.9755)	
	[0.0045]			[0.0046]			[0.0001]		
Opinions are taken seriously	0.0784***	0.1323	20319	0.0818***	0.1337	18110	0.0705***	0.1332	27278
	(0.0206)	(0.9716)		(0.0232)	(0.9810)		(0.0186)	(0.9675)	
	[0.0002]			[0.0005]			[0.0001]		
Recognised for good work	0.0610***	0.1336	18021	0.0690***	0.1477	16213	0.0614***	0.1570	27278
	(0.0161)	(0.9799)		(0.0172)	(0.9886)		(0.0183)	(0.9552)	
	[0.0002]			[0.0001]			[0.0008]		
Recognised for living values	0.0491***	0.1619	18021	0.0570***	0.1748	16213	0.0756***	0.1763	27278
	(0.0169)	(0.9463)		(0.0174)	(0.9496)		(0.0147)	(0.9303)	
	[0.0037]			[0.0011]			[0.0000]		
Bank treats customers fairly	0.0827***	0.0437	20061	0.0824***	0.0732	17876	0.0520**	0.0575	27278
	(0.0273)	(0.9279)		(0.0286)	(0.9236)		(0.0229)	(0.9203)	
	[0.0026]			[0.0041]			[0.0234]		
Bank meets customers' needs	0.0906***	0.0125	20061	0.0862***	0.0549	17876	0.0660**	0.0413	27278
	(0.0263)	(0.9317)		(0.0232)	(0.9256)		(0.0258)	(0.9166)	
	[0.0006]			[0.0002]			[0.0106]		
Can challenge customers' policies	0.0825***	0.0474	20061	0.0880***	0.0571	17876	0.0825***	0.0388	27278
	(0.0237)	(0.9105)		(0.0234)	(0.9164)		(0.0266)	(0.8973)	
	[0.0005]			[0.0002]			[0.0019]		
Proud to work for bank	0.0532***	-0.0132	19968	0.0657***	0.0131	17792	0.0560**	0.0165	27278
	(0.0200)	(0.9625)		(0.0205)	(0.9503)		(0.0252)	(0.9438)	
	[0.0079]			[0.0015]			[0.0266]		
Feel valued by bank	0.0627***	0.1847	19968	0.0740***	0.2029	17792	0.0601***	0.2019	27278
	(0.0230)	(0.9362)		(0.0221)	(0.9286)		(0.0212)	(0.9160)	
	[0.0065]			[0.0008]			[0.0045]		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. All regressions are estimated at the individual level and use post-treatment observations only (timeID $\geq$ 3). All specifications include semester and strata fixed effects. Specification (1) additionally controls for the corresponding baseline outcome (measured in semester 2; if unavailable, semester 1) interacted with semester indicators. Specification (2) is identical to (1) but restricts the sample by excluding employees who are observed at any point in offices outside the initial sample. Specification (3) reports treatment effects estimated using PDS LASSO. The set of potential controls includes baseline measures of function, managerial status, age, seniority, full-time employment status, performance, values, semester fixed effects, strata fixed effects, and their interactions. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

Table B22: Culture Fit (Robustness: Office and Team Fixed Effects)

	(1) FE + Baseline			(2) = (1) + Office FE			(3) = (1) + Team FE		
	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_C}(SD)$	N	T(SE)[P]	$\overline{Y_C}(SD)$	N
Culture index (std.)	0.0704***	0.0777	20492	0.0624***	0.0771	20458	0.0456***	0.0756	19246
	(0.0200)	(0.9846)		(0.0165)	(0.9846)		(0.0153)	(0.9888)	
	[0.0005]			[0.0002]			[0.0030]		
Culture index (PCA)	0.2180***	-0.1921	18016	0.2059***	-0.1943	17980	0.1528***	-0.2036	16855
	(0.0570)	(2.8488)		(0.0468)	(2.8489)		(0.0498)	(2.8620)	
	[0.0001]			[0.0000]			[0.0022]		
Social responsibility	0.1108***	-0.0228	9980	0.0919***	-0.0232	9858	0.0765*	-0.0279	8371
	(0.0382)	(0.9631)		(0.0320)	(0.9644)		(0.0433)	(0.9663)	
	[0.0038]			[0.0043]			[0.0778]		
Can be productive in my job	0.0646***	0.1700	20375	0.0449***	0.1692	20340	0.0301*	0.1671	19136
	(0.0242)	(0.9284)		(0.0147)	(0.9285)		(0.0161)	(0.9286)	
	[0.0077]			[0.0023]			[0.0622]		
Can achieve career goals	0.0561**	-0.0147	20104	0.0473***	-0.0150	20070	0.0395**	-0.0181	18872
	(0.0218)	(0.9714)		(0.0181)	(0.9716)		(0.0194)	(0.9734)	
	[0.0103]			[0.0091]			[0.0421]		
Positive impact strategy	0.0687***	-0.0450	20492	0.0539***	-0.0456	20458	0.0526***	-0.0450	19246
	(0.0241)	(0.9966)		(0.0156)	(0.9967)		(0.0183)	(0.9965)	
	[0.0045]			[0.0006]			[0.0043]		
Confident about bank's future	0.0395**	-0.0240	20492	0.0337**	-0.0242	20458	0.0361**	-0.0260	19246
	(0.0171)	(0.9731)		(0.0136)	(0.9731)		(0.0152)	(0.9753)	
	[0.0210]			[0.0137]			[0.0177]		
Trust senior leadership	0.0523***	0.0787	20169	0.0447***	0.0782	20134	0.0292**	0.0792	18938
	(0.0186)	(0.9531)		(0.0152)	(0.9533)		(0.0141)	(0.9527)	
	[0.0050]			[0.0033]			[0.0382]		
Trust direct manager	0.0531***	0.0122	20169	0.0561***	0.0117	20134	0.0337*	0.0148	18938
	(0.0184)	(0.9806)		(0.0139)	(0.9809)		(0.0179)	(0.9815)	
	[0.0040]			[0.0001]			[0.0605]		
Bank encourages speak-up	0.0775***	0.0240	20319	0.0549***	0.0235	20285	0.0514***	0.0199	19083
	(0.0225)	(0.9481)		(0.0166)	(0.9481)		(0.0140)	(0.9510)	
	[0.0006]			[0.0010]			[0.0003]		
Can speak up on concerns	0.0577***	-0.1403	20319	0.0563***	-0.1405	20285	0.0291	-0.1426	19083
	(0.0202)	(0.9892)		(0.0203)	(0.9894)		(0.0190)	(0.9908)	
	[0.0045]			[0.0056]			[0.1262]		
Opinions are taken seriously	0.0784***	0.1323	20319	0.0780***	0.1319	20285	0.0545***	0.1342	19083
	(0.0206)	(0.9716)		(0.0185)	(0.9718)		(0.0159)	(0.9720)	
	[0.0002]			[0.0000]			[0.0007]		
Recognised for good work	0.0610***	0.1336	18021	0.0554***	0.1332	17985	0.0418**	0.1335	16859
	(0.0161)	(0.9799)		(0.0127)	(0.9801)		(0.0177)	(0.9789)	
	[0.0002]			[0.0000]			[0.0189]		
Recognised for living values	0.0491***	0.1619	18021	0.0510***	0.1611	17985	0.0291*	0.1603	16859
	(0.0169)	(0.9463)		(0.0141)	(0.9465)		(0.0176)	(0.9496)	
	[0.0037]			[0.0003]			[0.0985]		
Bank treats customers fairly	0.0827***	0.0437	20061	0.0668***	0.0432	20028	0.0586***	0.0388	18828
	(0.0273)	(0.9279)		(0.0257)	(0.9280)		(0.0212)	(0.9320)	
	[0.0026]			[0.0095]			[0.0059]		
Bank meets customers' needs	0.0906***	0.0125	20061	0.0701***	0.0121	20028	0.0608***	0.0097	18828
	(0.0263)	(0.9317)		(0.0201)	(0.9320)		(0.0232)	(0.9346)	
	[0.0006]			[0.0005]			[0.0091]		
Can challenge customers' policies	0.0825***	0.0474	20061	0.0655***	0.0470	20028	0.0521***	0.0451	18828
	(0.0237)	(0.9105)		(0.0183)	(0.9107)		(0.0171)	(0.9142)	
	[0.0005]			[0.0004]			[0.0024]		
Proud to work for bank	0.0532***	-0.0132	19968	0.0391**	-0.0136	19933	0.0306	-0.0176	18739
	(0.0200)	(0.9625)		(0.0165)	(0.9626)		(0.0187)	(0.9633)	
	[0.0079]			[0.0176]			[0.1024]		
Feel valued by bank	0.0627***	0.1847	19968	0.0550***	0.1841	19933	0.0437**	0.1800	18739
	(0.0230)	(0.9362)		(0.0199)	(0.9364)		(0.0175)	(0.9397)	
	[0.0065]			[0.0059]			[0.0129]		

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. All regressions are estimated at the individual level using post-treatment observations only (timeID $\geq$ 3) and include semester and strata fixed effects. Specification (1) additionally controls for the corresponding baseline outcome (measured in semester 2; if unavailable, semester 1) interacted with semester indicators. Specification (2) further includes office fixed effects. Specification (3) further includes team fixed effects. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets. * p<0.10, ** p<0.05, *** p<0.01.*

Table B23: Culture Fit (MHT)

	(1) FE + Baseline	(2) DiD
Beliefs Index (standardized)	0.0704	0.0682***
	(0.0200)	(0.0271)
	{ 0.1070}	{ 0.0010}
Beliefs Index (PCA)	0.2180**	0.1943***
	(0.0570)	(0.0771)
	{ 0.0470}	{ 0.0010}
People have a strong sense of responsibility to society	0.0944	0.0588***
	(0.0325)	(0.0260)
	{ 0.1250}	{ 0.0010}
Conditions in my job allow me to be about as productive as I can be	0.0754*	0.0877***
	(0.0282)	(0.0359)
	{ 0.0690}	{ 0.0010}
I feel able to achieve my career objectives at this company	0.0592*	0.0314***
	(0.0230)	(0.0344)
	{ 0.0980}	{ 0.0010}
I am seeing the positive impact of our strategy	0.0626	0.0506***
	(0.0220)	(0.0324)
	{ 0.1290}	{ 0.0010}
I feel confident about BankName's future	0.0378	0.0422***
	(0.0163)	(0.0325)
	{ 0.1420}	{ 0.0010}
I trust senior lead	0.0567	0.0647***
	(0.0201)	(0.0246)
	{ 0.1150}	{ 0.0010}
I trust my direct manager	0.0530	0.0465***
	(0.0184)	(0.0211)
	{ 0.1000}	{ 0.0010}
BankName is genuine in its commitment to encourage colleagues to speak up	0.0773	0.0573***
	(0.0225)	(0.0185)
	{ 0.1000}	{ 0.0010}
I feel able to speak up when I see behaviour which I consider to be wrong	0.0538	0.0166***
	(0.0189)	(0.0172)
	{ 0.1230}	{ 0.0010}
People can state their opinion without the fear of negative consequences	0.0844**	0.0674***
	(0.0222)	(0.0187)
	{ 0.0350}	{ 0.0010}
I am recognised when I do a good job	0.0604**	0.0259***
	(0.0160)	(0.0231)
	{ 0.0140}	{ 0.0010}
I am recognised when I demonstrate BankName's values	0.0501**	0.0146***
	(0.0172)	(0.0249)
	{ 0.0340}	{ 0.0010}
BankName is genuine in its commitment to delivering fair outcomes for its custom	0.0649	0.0478***
	(0.0215)	(0.0239)
	{ 0.1200}	{ 0.0010}
I feel confident customers receive products and services that meet their needs	0.0735*	0.0578***
	(0.0214)	(0.0178)
	{ 0.0990}	{ 0.0010}
I feel able to challenge policies/procedures I think will be bad for customers	0.0813	0.0398***
	(0.0234)	(0.0189)
	{ 0.1160}	{ 0.0010}
I am proud to say I work for this company	0.0500	0.0470***
	(0.0188)	(0.0276)
	{ 0.1550}	{ 0.0010}
I feel valued by BankName	0.0703	0.0583***
	(0.0258)	(0.0411)
	{ 0.1040}	{ 0.0010}

Notes: *Each column reports estimates from regressions of the indicated outcome on treatment. Column (1) uses post-treatment observations only and includes semester and strata fixed effects, controlling for the baseline value of the outcome interacted with semester indicators. Column (2) is a difference-in-differences regression that includes individual and semester fixed effects and restricts the sample to employees observed at least once before and once after treatment. Standard errors are clustered at the office level and reported in parentheses. List-adjusted p-values (reported in braces) are computed using a Romano–Wolf stepdown procedure with 1000 cluster wild bootstrap replications. * p<0.10, ** p<0.05, *** p<0.01.*

Table B24: Heterogeneity Customer Facing

	Placebo	Unweighted		Weighted	
Beliefs Index	DV = CF	CF = 0	CF = 1	CF = 0	CF = 1
Treat	-0.0180	0.0171	0.0342***	0.0154	0.0312***
	(0.0354)	(0.0109)	(0.0104)	(0.0104)	(0.0097)
N	19216	27869	20232	27867	20232
R^2 (within)	0.0004	0.6553	0.6184	0.6567	0.6214
Mean DV		0.0352	0.1418	0.0222	0.1217
SD DV		0.9893	1.0099	0.9904	1.0071

Notes: *Standard errors clustered at officeID. * p<0.10, ** p<0.05, *** p<0.01. All regressions are estimated at the individual level. Column (1) reports a baseline balance regression of the indicator for being in a customer-facing role on the treatment dummy (DV = CF), using baseline data from semester 2 and controlling for strata fixed effects. Columns (2)–(5) report panel regressions of the Beliefs Index on the treatment dummy and the baseline Beliefs Index (measured in semester 2; if missing, semester 1) interacted with semester dummies, using all survey waves and including semester and strata fixed effects. Columns (2)–(3) are unweighted and estimate separate specifications for employees who are not in a customer-facing role at baseline (CF = 0) and those who are (CF = 1). Columns (4)–(5) repeat these two specifications but apply stabilized inverse-probability weights (IPW) based on a logit of post-treatment survey response on baseline HR and survey covariates (and indicators for missing values) and strata, then stabilized within strata. Baseline Beliefs Index and its interactions with semester dummies are included in all regressions but not reported.*

Table B25: Summary Statistics: Seniority Measures (Individual Level, Semester Averages)

	Mean	Sd.	Min.	Max	N
Panel A: Global Seniority (Bank Classification)					
Senior	0.09	0.29	0.00	1.00	34498
Share (LOO) Senior in Office	0.09	0.11	0.00	0.50	34491
Share (LOO) Senior in Team	0.09	0.17	0.00	1.00	33183
Senior, Beliefs sample	0.10	0.30	0.00	1.00	14133
Share (LOO) Senior in Office, Beliefs sample	0.04	0.05	0.00	0.50	14130
Share (LOO) Senior in Team, Beliefs sample	0.04	0.10	0.00	1.00	13503
Share (LOO) Senior in Office: Beliefs $<$ Median	0.02	0.03	0.00	0.50	14130
Share (LOO) Senior in Office: Beliefs $<$ P25	0.01	0.01	0.00	0.10	14130
Share (LOO) Senior in Team: Beliefs $<$ Median	0.02	0.06	0.00	1.00	13503
Share (LOO) Senior in Team: Beliefs $<$ P25	0.01	0.04	0.00	0.58	13503
Panel B: Relative Seniority (Top-2 in Office)					
Senior	0.34	0.47	0.00	1.00	32195
Share (LOO) Senior in Office	0.33	0.23	0.00	0.78	32195
Share (LOO) Senior in Team	0.31	0.32	0.00	1.00	30980
Senior, Beliefs sample	0.40	0.49	0.00	1.00	13110
Share (LOO) Senior in Office, Beliefs sample	0.17	0.10	0.00	0.60	13110
Share (LOO) Senior in Team, Beliefs sample	0.18	0.21	0.00	1.00	12552
Share (LOO) Senior in Office: Beliefs $<$ Median	0.08	0.05	0.00	0.44	13110
Share (LOO) Senior in Office: Beliefs $<$ P25	0.04	0.03	0.00	0.37	13110
Share (LOO) Senior in Team: Beliefs $<$ Median	0.08	0.12	0.00	1.00	12552
Share (LOO) Senior in Team: Beliefs $<$ P25	0.04	0.08	0.00	1.00	12552

Notes: *LOO = Leave-One-Out Share, that is, we exclude the individual from the share computation. Median and P25 splits are defined using beliefs among seniors under the corresponding seniority definition.*

Table B26: Individual-Level Convergence in Cultural Measures

DepVar:	(1) Index	(2) Index	(3) SR	(4) SR
Baseline	0.6644***	0.6112***	0.4778***	0.3883***
	(0.0285)	(0.0157)	(0.0377)	(0.0258)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Office treat × baseline	-0.0662*		-0.0848	
	(0.0375)		(0.0521)	
	[0.0781]		[0.1038]	
Office treat × Low		0.1018		0.1797*
		(0.1049)		(0.0952)
		[0.3321]		[0.0597]
Office treat × High		-0.0329		-0.0114
		(0.0654)		(0.0762)
		[0.6153]		[0.8816]
Baseline mean (SD)	0.130	0.130	3.802	3.802
	(0.978)	(0.978)	(0.845)	(0.845)
Observations	19276	19276	9697	9697

Notes. *All specifications include office, time, and randomization-strata fixed effects. For Social Responsibility, baseline groups are defined as Low = 1–3 and High = 5, with 4 (the median) as the reference category. For the Index (standardized), Low* < -1 *and High* > 1*, with Medium* $(-1, 1)$ *as the reference category. Estimates for group indicators in levels are omitted for conciseness. Standard errors clustered at the office level are reported in parentheses and p-values in brackets.* $*p < 0.10, **p < 0.05, ***p < 0.01$.

Table B27: Culture: Treatment Gaps & Volunteering Gaps (Robustness: Winsorize 1;99)

	(1) First stage	(2) Reduced form	(3) 2SLS	(4) Cluster-F	(5) KP-F	(6) Mean (SD)	(7) N
Panel A: Controls = Office T.Sh.							
Volunteering Gap	0.2010***	-0.2090**	-1.1184*	14.02	10.10	0.0841	17444
	(0.0537)	(0.1027)	(0.6598)			(0.9868)	
	[0.0002]	[0.0425]	[0.0901]				
Panel B: Controls = Office T.Sh. + T.Assignment							
Volunteering Gap	0.2006***	-0.2135**	-1.1387*	14.01	10.08	0.0841	17444
	(0.0536)	(0.1014)	(0.6572)			(0.9868)	
	[0.0002]	[0.0359]	[0.0832]				

Notes: *All specifications include semester and strata fixed effects and control for baseline beliefs interacted with time. Seniority is defined relative within office at baseline: individuals are classified as seniors if they belong to the highest (or top two) career grades in their office, and juniors otherwise. For each office and semester, leave–one–out (LOO) shares are constructed separately for juniors and seniors. The junior (senior) volunteering share equals the mean volunteering rate among other juniors (seniors) in the same office and semester. The junior–senior volunteering gap equals the LOO junior share minus the LOO senior share. The junior–senior treatment gap is defined analogously using LOO treatment shares. Volunteering and treatment gaps are winsorized at the 1st and 99th percentiles. Column (4) reports the cluster-robust Wald F-statistic for the excluded instrument from the first stage. Column (5) reports the Kleibergen–Paap rk Wald F-statistic for weak identification. Standard errors clustered at the office level are reported in parentheses; p-values in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B28: Social Responsibility: Treatment Gaps & Volunteering Gaps

	(1) First stage	(2) Reduced form	(3) 2SLS	(4) Cluster-F	(5) KP-F	(6) Mean (SD)	(7) N
Panel A: Controls = Office T.Sh.							
Volunteering Gap	0.1914***	-0.0499	-0.3176	15.88	4.72	3.7255	9285
	(0.0480)	(0.0683)	(0.4762)			(0.8250)	
	[0.0001]	[0.4650]	[0.5048]				
Panel B: Controls = Office T.Sh. + T.Assignment							
Volunteering Gap	0.1912***	-0.0501	-0.3188	15.87	4.73	3.7255	9285
	(0.0480)	(0.0684)	(0.4752)			(0.8250)	
	[0.0001]	[0.4636]	[0.5023]				

Notes: *All specifications include semester and strata fixed effects and control for baseline social responsibility interacted with time. Seniority (relative) at baseline: individuals are classified as seniors if they belong to the highest (or top two) career grades in their office, and juniors otherwise. For each office and semester, we construct leave–one–out (LOO) shares separately for juniors and seniors. The junior (senior) volunteering share is computed as the mean volunteering rate among other juniors (seniors) in the same office and semester, excluding the individual. The junior–senior volunteering gap equals the LOO junior share minus the LOO senior share. Analogously, the junior–senior treatment gap is defined as the LOO difference in treatment shares between juniors and seniors within office and semester. Column (1) reports the first-stage regression of the junior–senior volunteering gap on the junior–senior treatment gap. Column (2) reports the reduced-form effect of the treatment gap on beliefs. Column (3) reports the IV coefficient on the volunteering gap instrumented by the treatment gap. Column (4) reports the cluster-robust Wald F-statistic testing the null that the excluded instrument has zero effect in the first stage. Column (5) reports the Kleibergen–Paap rk Wald F-statistic for weak identification. Panel A controls for office treatment share. Panel B adds individual treatment assignment as a control. Standard errors clustered at the office level are reported in parentheses; p-values in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B29: Social Responsibility: Treatment Gaps & Volunteering Gaps (Robustness: Winsorize 1;99)

	(1) First stage	(2) Reduced form	(3) 2SLS	(4) Cluster-F	(5) KP-F	(6) Mean (SD)	(7) N
Panel A: Controls = Office T.Sh.							
Volunteering Gap	0.2018*** (0.0535) [0.0002]	-0.1164 (0.0874) [0.1830]	-0.7462 (0.6832) [0.2748]	14.23	3.84	3.7225 (0.8292)	8965
Panel B: Controls = Office T.Sh. + T.Assignment							
Volunteering Gap	0.2018*** (0.0535) [0.0002]	-0.1164 (0.0874) [0.1830]	-0.7726 (0.6853) [0.2596]	14.21	3.83	3.7225 (0.8292)	8965

Notes: *All specifications include semester and strata fixed effects and control for baseline social responsibility interacted with time. Seniority (relative) at baseline: individuals are classified as seniors if they belong to the highest (or top two) career grades in their office, and juniors otherwise. For each office and semester, we construct leave–one–out (LOO) shares separately for juniors and seniors. The junior (senior) volunteering share is computed as the mean volunteering rate among other juniors (seniors) in the same office and semester, excluding the individual. The junior–senior volunteering gap equals the LOO junior share minus the LOO senior share. Analogously, the junior–senior treatment gap is defined as the LOO difference in treatment shares between juniors and seniors within office and semester. Column (1) reports the first-stage regression of the junior–senior volunteering gap on the junior–senior treatment gap. Column (2) reports the reduced-form effect of the treatment gap on beliefs. Volunteering and treatment gaps are winsorized at the 1st and 99th percentiles. Column (3) reports the IV coefficient on the volunteering gap instrumented by the treatment gap. Column (4) reports the cluster-robust Wald F-statistic testing the null that the excluded instrument has zero effect in the first stage. Column (5) reports the Kleibergen–Paap rk Wald F-statistic for weak identification. Panel A controls for office treatment share. Panel B adds individual treatment assignment as a control. Standard errors clustered at the office level are reported in parentheses; p-values in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B30: Culture (Office Level): Treatment Gaps & Volunteering Gaps

	(1) First stage	(2) Reduced form	(3) 2SLS	(4) Cluster-F	(5) KP-F	(6) Mean (SD)	(7) N
Panel A: Unweighted							
Volunteering Gap	0.1645***	-0.0369	-0.2241	6.72	6.69	0.3402	452
	(0.0635)	(0.0495)	(0.3089)			(0.4523)	
	[0.0095]	[0.4557]	[0.4682]				
Panel B: Weighted by Respondent Count							
Volunteering Gap	0.1963***	-0.1876**	-1.0268*	13.06	11.55	0.3402	452
	(0.0543)	(0.0773)	(0.5359)			(0.4523)	
	[0.0003]	[0.0152]	[0.0554]				
Panel C: Weighted by Respondent Count × Response Rate							
Volunteering Gap	0.1880***	-0.2134**	-1.2173*	10.49	9.20	0.3402	452
	(0.0580)	(0.0841)	(0.6691)			(0.4523)	
	[0.0012]	[0.0111]	[0.0689]				

Notes: *All specifications control for baseline beliefs and randomization strata fixed effects. Seniority is defined relative within office at baseline. Post-period junior and senior volunteering and treatment shares are computed at the office level; the junior–senior gaps equal the difference between junior and senior shares. Column (4) reports the cluster-robust Wald F-statistic for the excluded instrument from the first stage. Column (5) reports the Kleibergen–Paap rk Wald F-statistic for weak identification. Panel A reports unweighted estimates. Panels B and C use office-level weights equal to (i) baseline respondent count and (ii) baseline respondent count adjusted for nonresponse rate, respectively. Standard errors clustered at the office level are in parentheses; p-values in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B31: Peer Effects: "Relative" Seniors, with Office Baseline Controls

Culture Fit Index	(1) Seniors b/se/p	(2) Seniors b/se/p	(3) Median b/se/p	(4) Median b/se/p	(5) P25 b/se/p	(6) P25 b/se/p
Treat	0.1196***	0.1204***	0.1278***	0.1342***	0.1345***	0.1362***
	(0.0310)	(0.0314)	(0.0340)	(0.0333)	(0.0292)	(0.0289)
	[0.0001]	[0.0001]	[0.0002]	[0.0001]	[0.0000]	[0.0000]
OffSh. Seniors	0.0696	0.0895				
	(0.1055)	(0.1139)				
	[0.5098]	[0.4321]				
Treat × OffSh. Seniors	-0.1531***	-0.1900**				
	(0.0589)	(0.0864)				
	[0.0096]	[0.0284]				
TeamSh. Seniors		-0.0140				
		(0.0357)				
		[0.6946]				
Treat × TeamSh. Seniors		0.0329				
		(0.0579)				
		[0.5701]				
OffSh. Seniors (Index<Med)			0.0674	0.0912		
			(0.3643)	(0.4186)		
			[0.8532]	[0.8277]		
Treat × OffSh. Seniors (Index<Med)			-0.5831**	-0.7485**		
			(0.2956)	(0.3527)		
			[0.0492]	[0.0344]		
TeamSh. Seniors (Index<Med)				-0.0527		
				(0.1055)		
				[0.6178]		
Treat × TeamSh. Seniors (Index<Med)				0.0946		
				(0.1404)		
				[0.5009]		
OffSh. Seniors (Index<P25)					0.0105	0.0252
					(0.5522)	(0.6269)
					[0.9848]	[0.9680]
Treat × OffSh. Seniors (Index<P25)					-1.3369***	-1.3928**
					(0.4894)	(0.5852)
					[0.0066]	[0.0177]
TeamSh. Seniors (Index<P25)						-0.0150
						(0.1907)
						[0.9373]
Treat × TeamSh. Seniors (Index<P25)						0.0270
						(0.2425)
						[0.9114]
N	19657	18731	15077	14346	15077	14346
N Offices	546	542	444	437	444	437
R2	0.3980	0.3969	0.4342	0.4325	0.4347	0.4329
Mean DV	0.1833	0.1795	0.1875	0.1819	0.1875	0.1819

Notes: *All regressions are estimated at the individual level with the Culture Fit Index as the dependent variable. Seniors are defined using the relative seniority classification (top two grades within each office). Columns (1)–(2) use the office- and team-level leave–one–out (LOO) shares of seniors; Columns (3)–(4) define low-belief seniors using the median of the baseline Culture Fit Index among seniors; Columns (5)–(6) use the 25th percentile. Odd-numbered columns include only the office-level LOO share; even-numbered columns additionally include the corresponding team-level LOO share. All specifications control for the individual baseline Culture Fit Index (measured in semester 2; if missing, semester 1) interacted with semester indicators, include semester and randomization-strata fixed effects, and add baseline office-level controls for mean Culture Fit Index, performance, values, office size, function composition (Global Banking, Commercial Banking, Retail Banking and Wealth Management), and baseline turnover. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B32: Peer Effects: "Relative" Seniors (Different Thresholds)

Culture Fit Index	(1) P50 b/se/p	(2) P50 b/se/p	(3) P33 b/se/p	(4) P33 b/se/p	(5) P25 b/se/p	(6) P25 b/se/p	(7) P10 b/se/p	(8) P10 b/se/p
Treat	0.1219***	0.1286***	0.1206***	0.1254***	0.1281***	0.1301***	0.1223***	0.1278***
	(0.0353)	(0.0357)	(0.0320)	(0.0320)	(0.0306)	(0.0313)	(0.0247)	(0.0253)
	[0.0006]	[0.0004]	[0.0002]	[0.0001]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
OffSh. seniors (Index<Median)	-0.6439	-0.6010						
	(0.4308)	(0.4683)						
	[0.1357]	[0.2001]						
Treat × OffSh. seniors (Index<Median)	-0.4839*	-0.6288*						
	(0.2856)	(0.3534)						
	[0.0909]	[0.0759]						
TeamSh. seniors (Index<Median)		-0.0367						
		(0.1085)						
		[0.7354]						
Treat × TeamSh. seniors (Index<Median)		0.0694						
		(0.1438)						
		[0.6294]						
OffSh. seniors (Index<P33)			-0.9024	-0.8037				
			(0.6035)	(0.6577)				
			[0.1354]	[0.2223]				
Treat × OffSh. seniors (Index<P33)			-0.7554*	-0.9953**				
			(0.3919)	(0.4817)				
			[0.0544]	[0.0393]				
TeamSh. seniors (Index<P33)				-0.0812				
				(0.1609)				
				[0.6142]				
Treat × TeamSh. seniors (Index<P33)				0.1559				
				(0.2005)				
				[0.4371]				
OffSh. seniors (Index<P25)					-1.1002	-1.1033		
					(0.7053)	(0.7186)		
					[0.1195]	[0.1254]		
Treat × OffSh. seniors (Index<P25)					-1.1622**	-1.2205**		
					(0.4821)	(0.5887)		
					[0.0163]	[0.0388]		
TeamSh. seniors (Index<P25)						-0.0006		
						(0.1931)		
						[0.9974]		
Treat × TeamSh. seniors (Index<P25)						0.0172		
						(0.2387)		
						[0.9426]		
OffSh. seniors (Index<P10)							-0.9673	-0.8230
							(1.3169)	(1.5744)
							[0.4630]	[0.6014]
Treat × OffSh. seniors (Index<P10)							-2.8553***	-3.0413***
							(0.9046)	(1.0392)
							[0.0017]	[0.0036]
TeamSh. seniors (Index<P10)								0.0023
								(0.2924)
								[0.9937]
Treat × TeamSh. seniors (Index<P10)								-0.1632
								(0.3162)
								[0.6061]
N	15077	14346	15515	14745	15077	14346	15515	14745
N Offices	444	437	514	508	444	437	514	508
R2	0.4291	0.4273	0.4264	0.4247	0.4300	0.4280	0.4260	0.4243
Mean DV	0.1875	0.1819	0.1881	0.1831	0.1875	0.1819	0.1881	0.1831

Notes: *All regressions are estimated at the individual level with the Culture Fit Index as the dependent variable. Seniors are defined using the relative seniority classification (top two grades within each office). Columns (1)–(2) define low-belief seniors using the median (P50) of the baseline Culture Fit Index among seniors; Columns (3)–(4) use the 33rd percentile (P33); Columns (5)–(6) use the 25th percentile (P25); and Columns (7)–(8) use the 10th percentile (P10). Odd-numbered columns include only the office-level leave–one–out (LOO) share of low-belief seniors; even-numbered columns additionally include the corresponding team-level LOO share. All specifications control for the individual baseline Culture Fit Index (measured in semester 2; if missing, semester 1) interacted with semester indicators, and include semester and randomization-strata fixed effects. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B33: Peer Effects: "Relative" Seniors (Placebo)

Culture Fit Index	(1) P50 b/se/p	(2) P50 b/se/p	(3) P33 b/se/p	(4) P33 b/se/p	(5) P25 b/se/p	(6) P25 b/se/p	(7) P10 b/se/p	(8) P10 b/se/p
Post (Sem. 2 vs 1)	0.2756***	0.2729***	0.2640***	0.2657***	0.2689***	0.2694***	0.2578***	0.2596***
	(0.0490)	(0.0515)	(0.0413)	(0.0438)	(0.0379)	(0.0397)	(0.0301)	(0.0311)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
OffSh. seniors (Index<Median)	0.3419	0.7915						
	(0.7955)	(0.7904)						
	[0.6676]	[0.3173]						
Post × OffSh. seniors (Index<Median)	-0.4083	-0.4905						
	(0.6935)	(0.5598)						
	[0.5564]	[0.3814]						
TeamSh. seniors (Index<Median)		-0.4671***						
		(0.1274)						
		[0.0003]						
Post × TeamSh. seniors (Index<Median)		0.1049						
		(0.2379)						
		[0.6594]						
OffSh. seniors (Index<P33)			0.0779	0.7017				
			(1.0970)	(1.1659)				
			[0.9434]	[0.5476]				
Post × OffSh. seniors (Index<P33)			-0.5032	-0.5998				
			(0.9149)	(0.8067)				
			[0.5825]	[0.4575]				
TeamSh. seniors (Index<P33)				-0.6224***				
				(0.1281)				
				[0.0000]				
Post × TeamSh. seniors (Index<P33)				0.0513				
				(0.2456)				
				[0.8346]				
OffSh. seniors (Index<P25)					-0.3719	0.3733		
					(1.2578)	(1.3373)		
					[0.7677]	[0.7803]		
Post × OffSh. seniors (Index<P25)					-0.7065	-0.7682		
					(1.0267)	(0.9479)		
					[0.4918]	[0.4182]		
TeamSh. seniors (Index<P25)						-0.6721***		
						(0.1610)		
						[0.0000]		
Post × TeamSh. seniors (Index<P25)						0.0296		
						(0.2547)		
						[0.9074]		
OffSh. seniors (Index<P10)							-1.8647	-1.0824
							(2.2726)	(2.5324)
							[0.4123]	[0.6693]
Post × OffSh. seniors (Index<P10)							-1.4125	-1.5233
							(1.7247)	(1.6834)
							[0.4132]	[0.3660]
TeamSh. seniors (Index<P10)								-0.7232**
								(0.3008)
								[0.0166]
Post × TeamSh. seniors (Index<P10)								0.0080
								(0.2647)
								[0.9760]
N	21059	20128	21609	20624	21059	20128	21609	20624
N Offices	383	381	481	475	383	381	481	475
R2	0.0138	0.0159	0.0134	0.0163	0.0144	0.0168	0.0152	0.0161
Mean DV	0.0321	0.0298	0.0262	0.0248	0.0321	0.0298	0.0262	0.0248

Notes: *All regressions are estimated at the individual level using only pre-treatment survey waves (semesters 1 and 2), with the Culture Fit Index as the dependent variable. The indicator Post (Sem. 2 vs 1) equals one in semester 2 and zero in semester 1. Columns (1)–(2) define low-belief seniors using the median (P50) of the baseline Culture Fit Index among seniors; Columns (3)–(4) use the 33rd percentile (P33); Columns (5)–(6) use the 25th percentile (P25); and Columns (7)–(8) use the 10th percentile (P10). Odd-numbered columns include only the office-level leave–one–out (LOO) share of low-belief seniors; even-numbered columns additionally include the corresponding team-level LOO share. All specifications include randomization-strata fixed effects. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets. The absence of statistically significant interaction terms indicates no differential pre-trends.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B34: Peer Effects: "Global" Seniors

Culture Fit Index	(1) BankC. b/se/p	(2) BankC. b/se/p	(3) Median b/se/p	(4) Median b/se/p	(5) P25 b/se/p	(6) P25 b/se/p
Treat	0.1018***	0.1021***	0.1019***	0.1055***	0.1026***	0.1065***
	(0.0225)	(0.0235)	(0.0232)	(0.0239)	(0.0225)	(0.0231)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
OffSh. Seniors	-0.4336***	-0.3551*				
	(0.1555)	(0.1934)				
	[0.0055]	[0.0668]				
Treat × OffSh. Seniors	-0.3094***	-0.3359**				
	(0.1008)	(0.1607)				
	[0.0022]	[0.0370]				
TeamSh. Seniors		-0.0586				
		(0.0556)				
		[0.2923]				
Treat × TeamSh. Seniors		0.0208				
		(0.0916)				
		[0.8209]				
OffSh. Seniors (Index<Med)			-2.1375***	-1.9489**		
			(0.7162)	(0.7835)		
			[0.0029]	[0.0131]		
Treat × OffSh. Seniors (Index<Med)			-1.0354**	-1.3620**		
			(0.5145)	(0.6570)		
			[0.0446]	[0.0386]		
TeamSh. Seniors (Index<Med)				-0.0545		
				(0.1399)		
				[0.6968]		
Treat × TeamSh. Seniors (Index<Med)				0.1956		
				(0.1913)		
				[0.3071]		
OffSh. Seniors (Index<P25)					-3.9356***	-3.6117**
					(1.4455)	(1.4660)
					[0.0067]	[0.0140]
Treat × OffSh. Seniors (Index<P25)					-2.2220**	-2.8411**
					(1.0597)	(1.2423)
					[0.0364]	[0.0225]
TeamSh. Seniors (Index<P25)						-0.0881
						(0.1367)
						[0.5194]
Treat × TeamSh. Seniors (Index<P25)						0.3326*
						(0.1905)
						[0.0813]
N	20486	19512	16121	15314	16121	15314
N Offices	662	660	627	623	627	623
R2	0.3932	0.3917	0.4280	0.4263	0.4281	0.4264
Mean DV	0.1883	0.1848	0.1921	0.1876	0.1921	0.1876

Notes: *All regressions are estimated at the individual level with the Culture Fit Index as the dependent variable. Seniors are defined using the bank's global seniority classification. Columns (1)–(2) interact treatment with the office- and team-level leave–one–out (LOO) shares of global seniors. Columns (3)–(4) define low-belief seniors using the median (P50) of the baseline Culture Fit Index among global seniors; Columns (5)–(6) use the 25th percentile (P25). Odd-numbered columns include only the office-level LOO share; even-numbered columns additionally include the corresponding team-level LOO share. All specifications control for the individual baseline Culture Fit Index (measured in semester 2; if missing, semester 1) interacted with semester indicators, and include semester and randomization-strata fixed effects. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B35: Peer Effects: "Relative" Seniors (Social Responsibility)

Social Responsibility	(1) Seniors b/se/p	(2) Seniors b/se/p	(3) Low SR b/se/p	(4) Low SR b/se/p
Treat	0.1236***	0.1279***	0.1237***	0.1323***
	(0.0363)	(0.0358)	(0.0330)	(0.0341)
	[0.0007]	[0.0004]	[0.0002]	[0.0001]
OffSh. Seniors	-0.0997***	-0.1145*		
	(0.0340)	(0.0665)		
	[0.0036]	[0.0859]		
Treat × OffSh. Seniors	-0.0938	-0.1637		
	(0.1008)	(0.1042)		
	[0.3526]	[0.1168]		
TeamSh. Seniors		0.0274		
		(0.0679)		
		[0.6870]		
Treat × TeamSh. Seniors		0.0533		
		(0.0783)		
		[0.4961]		
OffSh. Seniors (SR $\leq$ 3)			-1.0625***	-0.8812***
			(0.3684)	(0.2821)
			[0.0041]	[0.0019]
Treat × OffSh. Seniors (SR $\leq$ 3)			-0.2980	-0.5547
			(0.5891)	(0.4653)
			[0.6132]	[0.2338]
TeamSh. Seniors (SR $\leq$ 3)				-0.0799
				(0.1700)
				[0.6386]
Treat × TeamSh. Seniors (SR $\leq$ 3)				0.0799
				(0.2108)
				[0.7048]
N	9605	9134	7703	7306
N Offices	488	483	456	448
R2	0.2184	0.2186	0.2420	0.2423
Mean DV	3.8224	3.8165	3.8369	3.8304

Notes: *All regressions are estimated at the individual level with Social Responsibility as the dependent variable. Seniors are defined using the relative seniority classification (top two grades within each office). Columns (1)–(2) interact treatment with the office- and team-level leave–one–out (LOO) shares of seniors; Columns (3)–(4) restrict seniors to those with low baseline Social Responsibility (defined as baseline response ≤ 3). Odd-numbered columns include only the office-level LOO share; even-numbered columns additionally include the corresponding team-level LOO share. All specifications control for baseline Social Responsibility (measured in semester 2; if missing, semester 1) interacted with semester indicators, and include semester and randomization-strata fixed effects. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B36: Peer Effects: "Global" Seniors (Social Responsibility)

Social Responsibility	(1) BankC. b/se/p	(2) BankC. b/se/p	(3) Low SR b/se/p	(4) Low SR b/se/p
Treat	0.1272***	0.1289***	0.1352***	0.1381***
	(0.0258)	(0.0268)	(0.0255)	(0.0267)
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
OffSh. Seniors	-0.2181**	-0.3430***		
	(0.0885)	(0.1066)		
	[0.0140]	[0.0014]		
Treat × OffSh. Seniors	-0.3314**	-0.1893*		
	(0.1366)	(0.1085)		
	[0.0155]	[0.0815]		
TeamSh. Seniors		0.1675***		
		(0.0525)		
		[0.0015]		
Treat × TeamSh. Seniors		-0.1772**		
		(0.0736)		
		[0.0164]		
OffSh. Seniors (SR ≤ 3)			-1.8010**	-1.6091**
			(0.8513)	(0.6943)
			[0.0348]	[0.0208]
Treat × OffSh. Seniors (SR ≤ 3)			-1.7078*	-1.9419**
			(1.0174)	(0.8495)
			[0.0938]	[0.0226]
TeamSh. Seniors (SR ≤ 3)				0.0846
				(0.1075)
				[0.4318]
Treat × TeamSh. Seniors (SR ≤ 3)				-0.0614
				(0.1756)
				[0.7268]
N	9977	9481	7986	7570
N Offices	621	616	580	575
R2	0.2189	0.2190	0.2403	0.2407
Mean DV	3.8321	3.8268	3.8454	3.8396

Notes: *All regressions are estimated at the individual level with Social Responsibility as the dependent variable. Seniors are defined using the bank's global seniority classification. Columns (1)–(2) interact treatment with the office- and team-level leave–one–out (LOO) shares of global seniors; Columns (3)–(4) restrict seniors to those with low baseline Social Responsibility (defined as baseline response ≤ 3). Odd-numbered columns include only the office-level LOO share; even-numbered columns additionally include the corresponding team-level LOO share. All specifications control for baseline Social Responsibility (measured in semester 2; if missing, semester 1) interacted with semester indicators, and include semester and randomization-strata fixed effects. Standard errors are clustered at the office level and reported in parentheses; p-values are reported in brackets.* $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

D Additional Figures

Figure C1: Offices

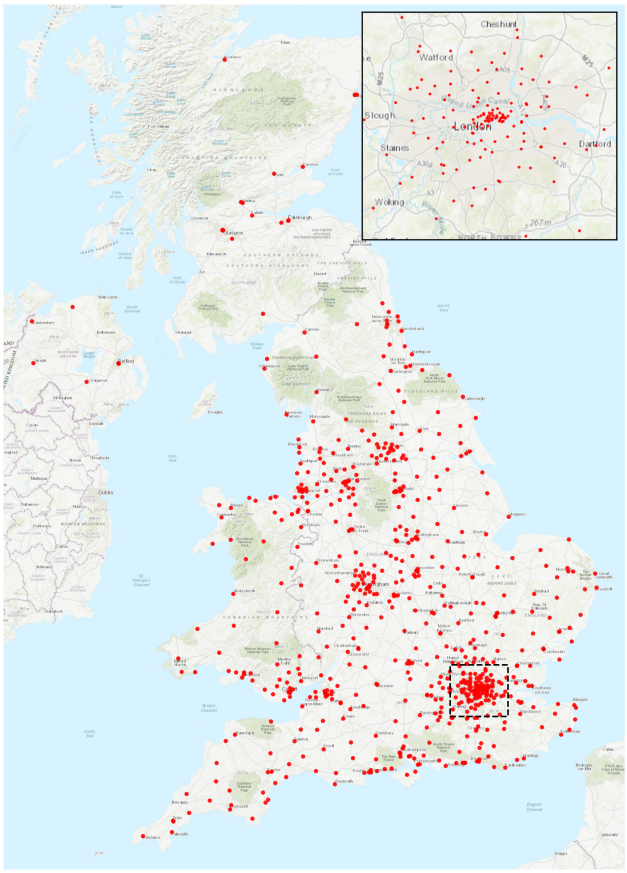


Figure C2: Employees by Area Code

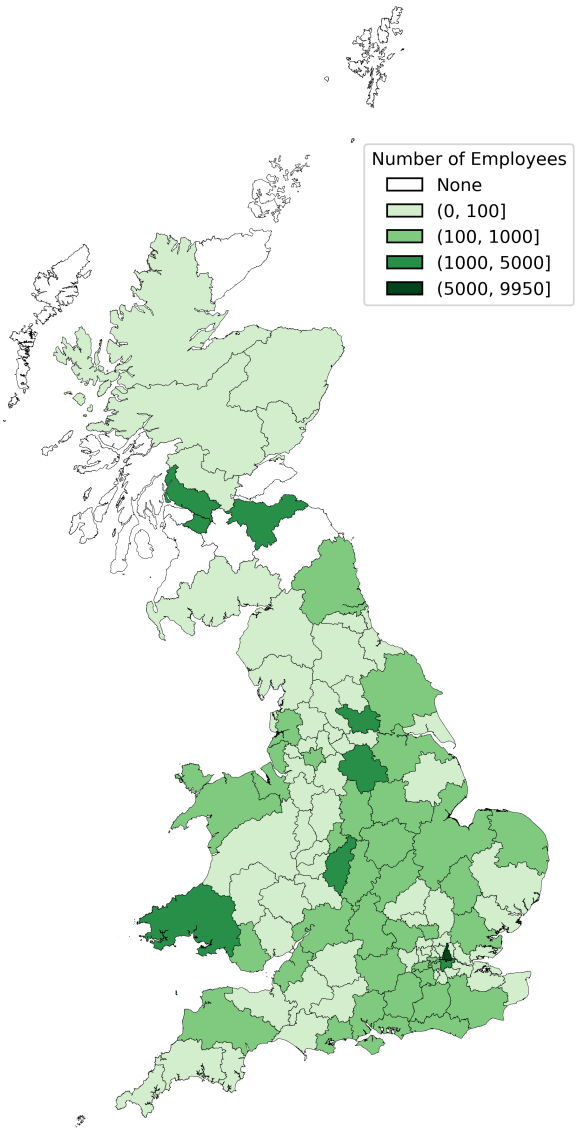


Figure C3: Volunteering Hours

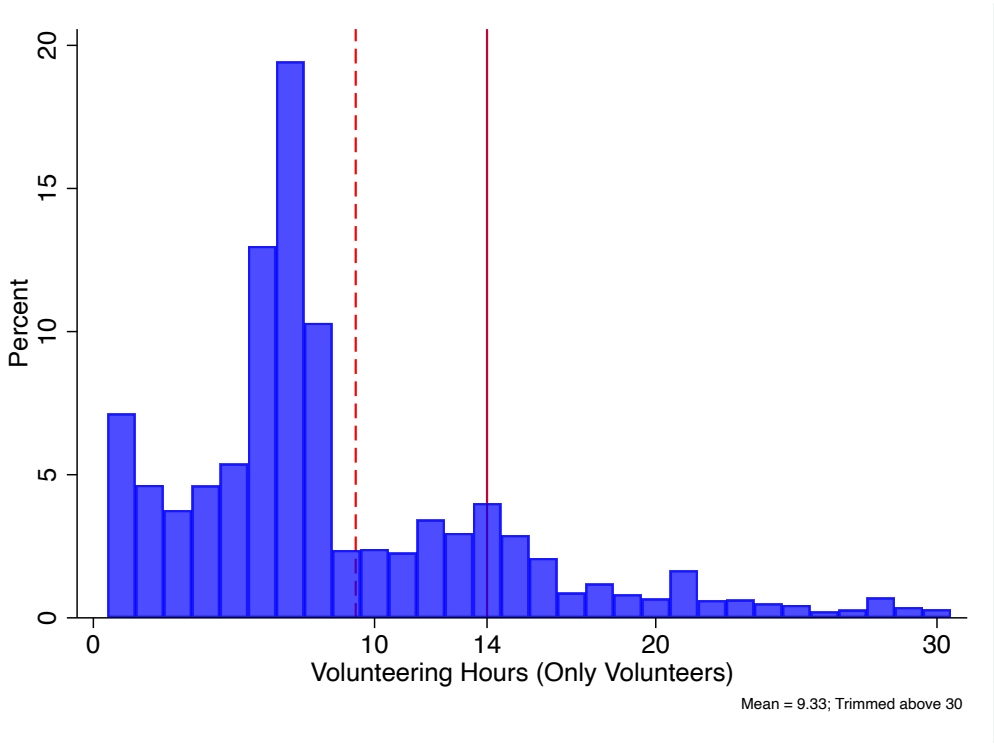


Figure C4: Monotonicity of Treatment Assignment

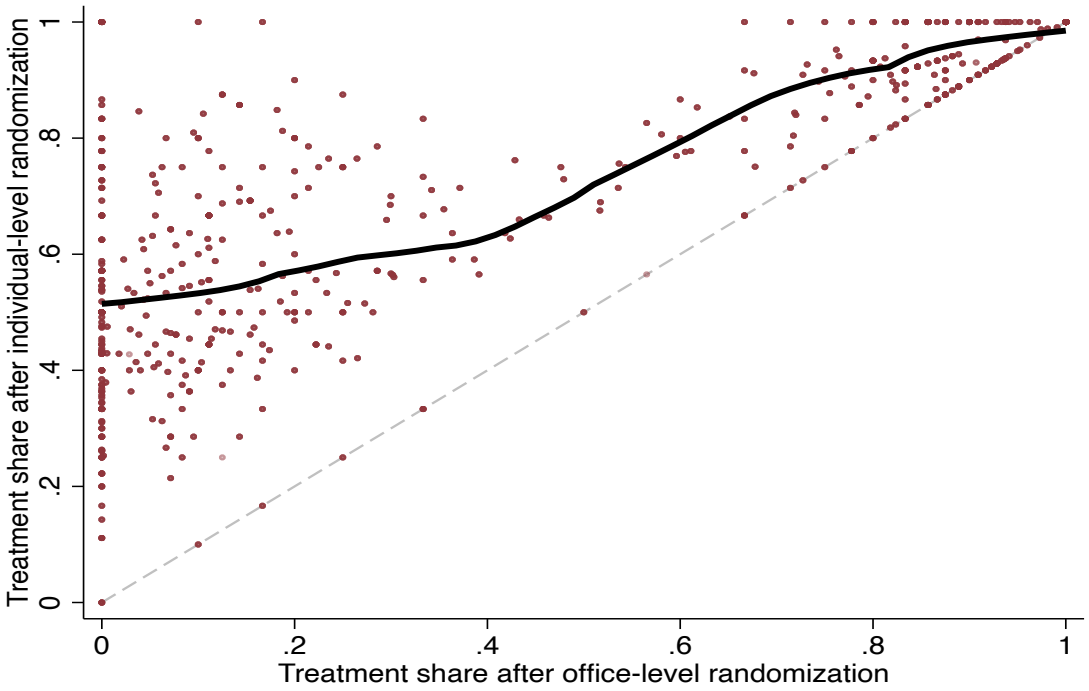


Figure C5: Random Treatment Assignment within Offices

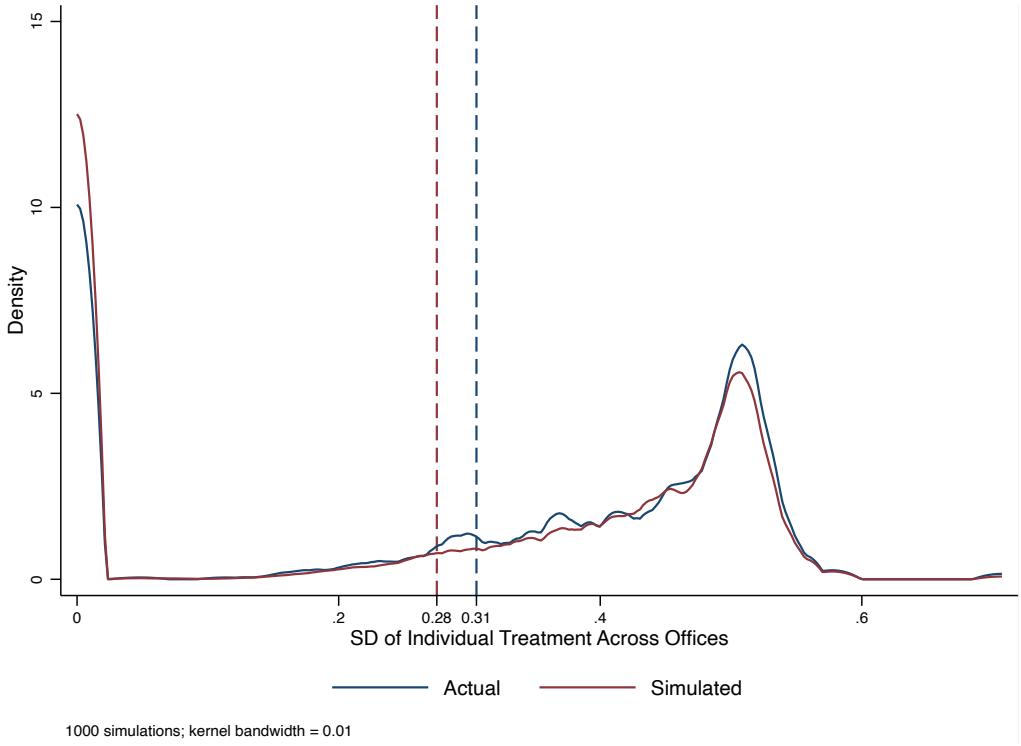


Figure C6: Random Individual Treatment Assignment (Team)

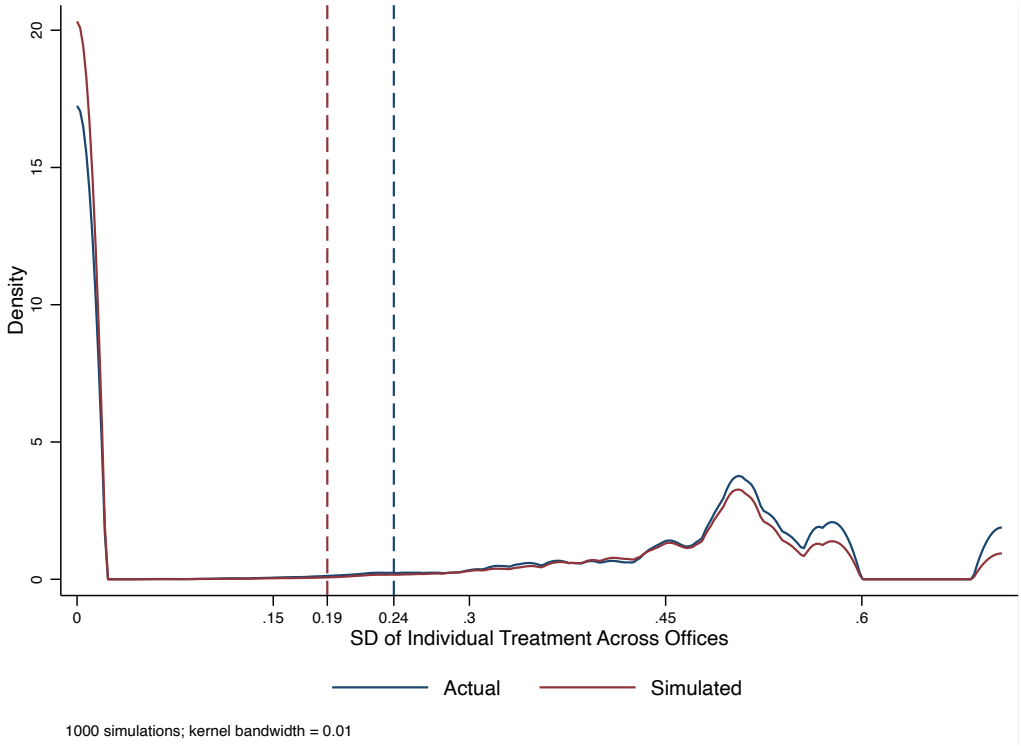


Figure C7: Volunteering Take Up by Office (large offices only)

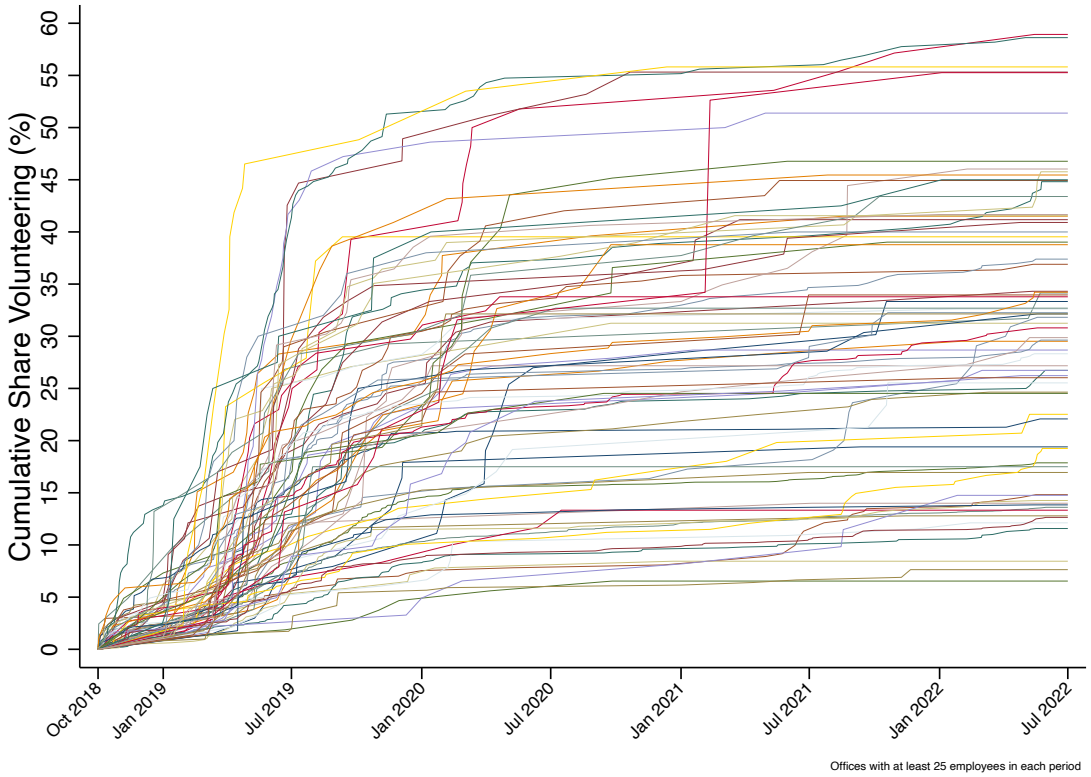


Figure C8: Volunteering: Take Up by Treatment Group

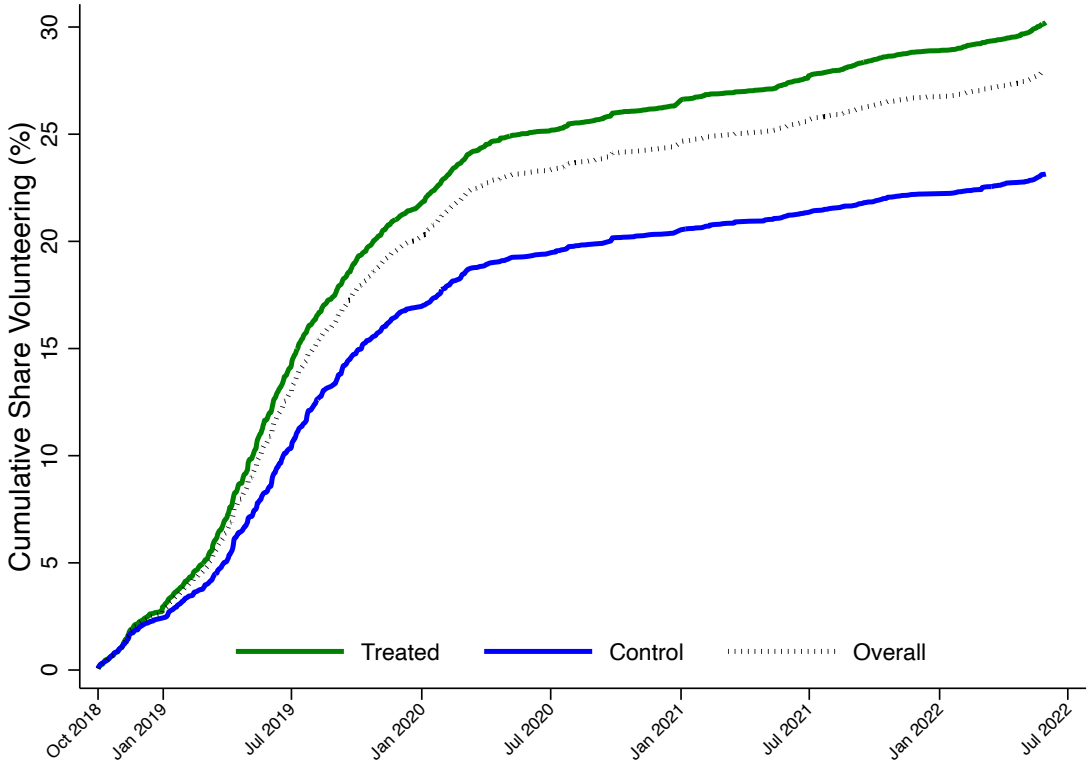


Figure C9: ITT Volunteering (Dynamics)

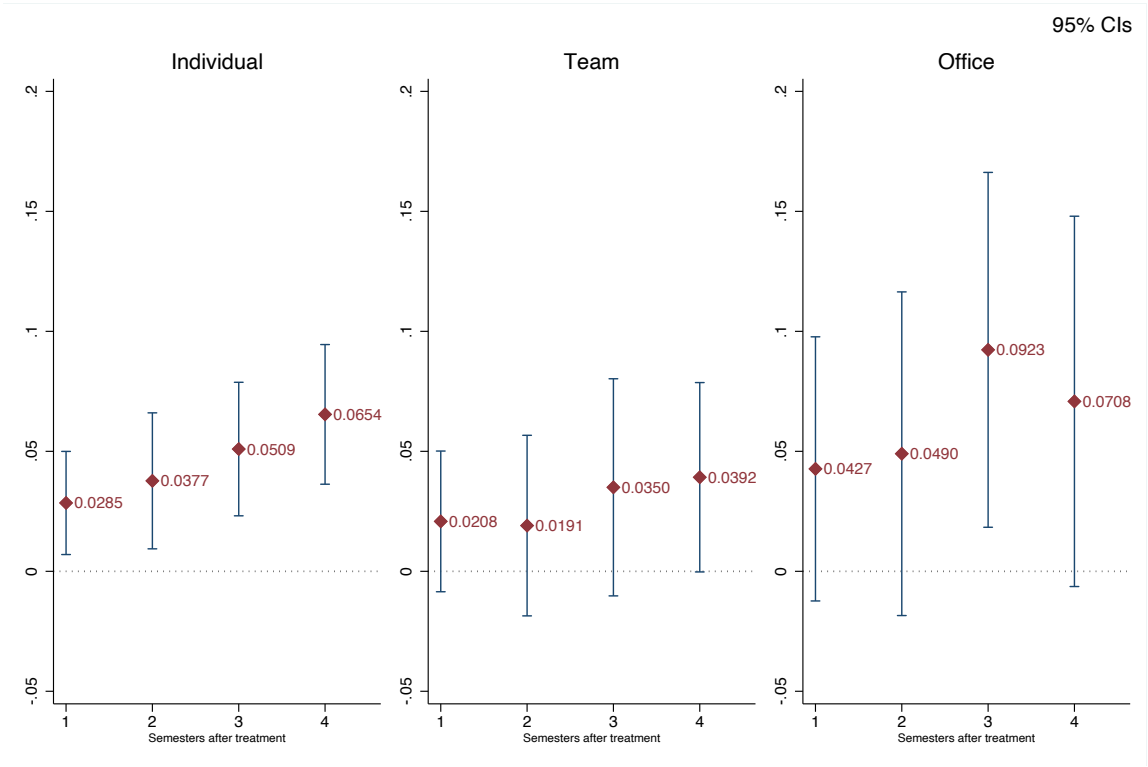


Figure C10: Volunteering Take Up by Seniority and Culture

