Working Paper 22-063

Is Hybrid Work the Best of Both Worlds? Evidence from a Field Experiment

Prithwiraj (Raj) Choudhury
Tarun Khanna
Christos A. Makridis
Kyle Schirmann

Harvard Business School

Is Hybrid Work the Best of Both Worlds? Evidence from a Field Experiment

Prithwiraj (Raj) Choudhury
Harvard Business School

Tarun Khanna
Harvard Business School

Christos A. Makridis
Columbia Business School; Stanford University

Kyle Schirmann
Harvard Business School

Working Paper 22-063

Funding for this research was provided in part by Harvard Business School. These views are the authors' own and do not reflect those of any affiliated institutions.

Is Hybrid Work the Best of Both Worlds? Evidence from a Field Experiment

Prithwiraj (Raj) Choudhury, Tarun Khanna, Christos A. Makridis, Kyle Schirmann *

March 24, 2022

Abstract

Hybrid work is emerging as a novel form of organizing work globally. This paper reports causal evidence on how the extent of hybrid work—the number of days worked from home relative to days worked from the office—affects work outcomes. Collaborating with an organization in Bangladesh, we randomized the number of days that individual employees worked from the office for nine weeks in the summer of 2020. Our results indicate that an intermediate number of days in the office results in more emails sent, a higher number of email recipients, and increased novelty of work products. Our test for underlying mechanisms suggests that hybrid work might represent the "best of both worlds," offering workers greater work-life balance, without the concern of being isolated from colleagues.

Keywords: Hybrid Work; Remote Work; Work-from-Home; Field Experiment.

JEL Codes: J23, J24, O10, O33

*Raj: Harvard Business School, pchoudhury@hbs.edu; Tarun: Harvard Business School, tkhanna@hbs.edu; Christos: Columbia Business School and Stanford University, cmakridi@stanford.edu; Kyle: Harvard Business School, kschirmann@hbs.edu. These views are our own and do not reflect those of any affiliated institutions. We thank Subhradip Sarker for important early contributions and Jeffrey Gu for excellent research assistance.

1 Introduction

"Hybrid work," where employees spend some of their work days in the physical office and the rest of their work days working remotely, is emerging as a novel form of organizing knowledge work globally (Teevan, 2021; Cutter, 2021). Barerro et al. (2022) estimate that 20 percent of full workdays will be supplied from home after the pandemic ends, compared with just five percent before.[1] While some companies and employees are considering more limited work-from-home (WFH) plans that allow workers to WFH for a day or two a week, other companies have announced plans for greater flexibility.[2] Recent commentary has hypothesized that hybrid and remote work have the potential to transform cities and the spatial composition of the workforce (Choudhury, 2020; Eisen, 2019; Poleg, 2021), reminiscent of hypotheses at the turn of the 21st century with the introduction of the Internet and increased use of computers (Glaeser, 1998).

However, prominent voices remain skeptical towards hybrid work. The CEO of Goldman Sachs has been skeptical about its efficacy in "innovative, collaborative" contexts and wants workers back in the office five days a week.[3] A recent paper by Yang et al. (2021) echoes concerns about how remote work might negatively affect the production of creative work if workers communicate with fewer colleagues. This raises questions around whether a transition to hybrid work would negatively affect intrafirm communication and/or the novelty of work output.

Despite these ongoing debates, we lack causal evidence on how hybrid work affects intrafirm

[1] Brynjolfsson et al. (2021) report that, while 30% of respondents reported that they were working fully remotely and 39.1% reported that they were working in a hybrid arrangement as of October 2020, as much as 9.5% and 20.8% reported that they anticipated fully-remote or hybrid work, respectively, in the long run. Relatedly, Bloom et al. (2021) report that the share of new U.S. patent applications that advance WFH technologies more than doubled from January to September 2020.

[2] https://www.flexjobs.com/blog/post/companies-switching-remote-work-long-term/

[3] https://fortune.com/2022/03/10/goldman-sachs-office-hybrid-remote-work-david-solomon/

communication and individual work outcomes, especially related to novelty of work. Pre-pandemic studies, such as Bloom et al. (2014c) and Choudhury et al. (2020), document causal productivity gains of workers transitioning from the office to work-from-home, and from work-from-home to work-from-anywhere, respectively. Other studies conducted during the pandemic, such as Gibbs et al. (2021) and Yang et al. (2021), also focus on firm-wide transitions to remote work. However, to the best of our knowledge, we lack studies that document causal evidence on how the extent of hybrid work—that is the extent to which the worker worked from the office rather than from home—affects patterns of intrafirm communication or novelty of work output.

To address this important and timely question, we report results from a field experiment conducted in the summer of 2020 in collaboration with BRAC, the world's largest non-governmental organization (Khanna and Ramachandran, 2021). We randomized the number of days that 130 workers worked from the physical office over a period of nine weeks. Following the lifting of a national lockdown, a bureaucratic policy mandated by ongoing COVID-19 health and safety concerns restricted the number of employees allowed back in the physical office. Exploiting this policy, daily lotteries determined which workers were directed to work from the office versus working from home. Our sample includes workers in the human resources department, performing administrative tasks at the headquarters of the organization in Dhaka, Bangladesh. While 130 employees took part in the experiment, we have complete data (including email communication data) for 108 HR employees.[4] The randomization protocol ensures that the number of days worked in the office for each worker in our sample is exogenously determined during the nine-week treatment period.

To examine how patterns of hybrid work affect communication and work novelty at the indi-

[4]The sample dropped from 130 to 108 as 108 employees voluntarily consented to their data being shared for the experiment. The data was anonymized prior to being used by the researchers. The relevant IRB and Data Safety and Security approvals are IRB22-0292 and DAT22-0094.

vidual level, we categorize workers into three groups based on how many days they were assigned to work from the office: high WFH (0-8 days in the office, corresponding to 0-23% of work days in the office), intermediate WFH (9-14 days in the office, 23-40%) and low WFH (15+ days in the office, greater than 40%).[5] Using data on 32,745 emails sent among the 108 HR employees in the pre-treatment (lockdown) and treatment periods, we create several individual-level measures of intrafirm communication, including pairwise counts of emails sent by each individual sender, and characteristics of emails (i.e., word length, sentiment, number of unique recipients). Using the text and attachments of these 30,323 emails, we created measures of individual-level work novelty using machine learning and textual analysis methods. These methods—Doc2vec paragraph embeddings, BIRCH clustering, and cosine similarity—are described in detail in the Appendix.

We report four sets of results. First, using a negative binomial model, we find a statistically significant non-linear relationship between number of days in the office and email communication. In particular, we find that intermediate WFH is associated with a 0.814 increase in the number of emails for a given dyad and low WFH is associated with a 0.537 increase, both relative to the baseline of high WFH. We confirm using a t-test that the two point estimates are statistically different from one another. We also find that intermediate WFH is associated with positive increases in the number of email recipients and the sentiment of email text, measured using natural language processing methods on the content of each email.

Second, we study how the extent of hybrid work affects novelty of work output. We find that individuals in the intermediate-WFH category have greater novelty of work products compared to workers in low- and high-WFH categories, and this difference is significant: the novelty measure

[5] The cut-offs for the three groups are determined to balance the number of workers in each sub-group, but we show later that our results are invariant to slight changes around the cut-off.

of intermediate-WFH workers increases by 0.669 standard deviations more than for high-WFH employees. In supplementary analysis reported in the Appendix, we document that managers subjectively rate the work output of workers in the intermediate-WFH category as being of higher quality than workers in the high-WFH and low-WFH groups.

Third, we shed light on mechanisms using prior survey measures from the literature on remote work in organizational theory (Raghuram et al., 2001). This literature has argued that workers with intermediate hybrid might enjoy the "best of both worlds" in relation to two underlying mechanisms: flexibility and isolation. Prior literature has argued that the flexibility to decide how to accomplish a task (Amabile et al., 1996) and the flexibility to schedule one's work week (Elsbach and Hargadon, 2006) positively affect the production of novel work. As Bloom et al. (2014a) document, remote work reduces distractions and commute times and provides greater flexibility to workers. On the flip side, other research has shown that remote work leads to isolation from colleagues (Bartel et al., 2012) and isolation negatively impacts work outcomes (Golden et al., 2008a). Intermediate hybrid work is plausibly the sweet spot, where workers enjoy flexibility and yet are not as isolated compared to peers who are predominantly working from home. We report findings using worker surveys conducted at the end of the experiment that suggest that workers in the intermediate-WFH category reported greater satisfaction with working from home, greater work-life balance, and lower isolation compared to workers in the high- and low-WFH categories.

Finally, we test the external validity of our findings by drawing on new and proprietary data from Gallup. We observe 12,166 individuals in the United States between March 2020 and June 2021. Controlling for a wide array of individual characteristics and income and occupation fixed effects, we find that higher levels of trust, mentorship, and purpose at work moderate job satisfaction and preparedness among full WFH workers, but do not for intermediate WFH or in-person workers. In other words, intermediate WFH arrangements allow workers to capture the benefits of a productive and enjoyable workplace almost as much as those workers who are always in the office. While these results are not causal, they underscore the external validity of our main result: hybrid work is associated with improved worker outcomes.

Our results contribute to the growing literature on remote work and productivity (Bloom et al., 2014b). To the best of our knowledge, there are no prior studies that examine how the extent of hybrid work—the number of days in the office relative to number of days WFH—relates to individual work outcomes and intrafirm communication. Our results consistently suggest that intermediate levels of WFH may result in both enhanced novelty of work products and greater work-related communication. Our results are also important for understanding the importance of physical co-location of workers to productivity outcomes (Storper and Venables, 2004) and the future of work and cities (Gaspar and Glaeser, 1998; Glaeser, 1998).

2 Data and Measurement

2.1 Experimental Design

Our field experiment was conducted in collaboration with BRAC, the world's largest non-governmental organization, headquartered in Bangladesh. Founded over four decades ago, the firm has over 35,000 staff as of September 2020 and over $1 billion in total income. In 2019, 81% of BRAC's revenues came from earned income and women comprise 42% of the total workforce of BRAC (Khanna and Ramachandran, 2021). While BRAC is headquartered in Dhaka, it has operations in multiple countries including Myanmar, Liberia, Sierra Leone, Uganda, and Rwanda. Employees in the BRAC headquarters—the focus of our study—work in a modern office in Dhaka.[6] Prior to the pandemic, these employees had worked all five days in the office.

To study the causal effect of how the extent of hybrid work—low WFH, intermediate WFH, and high WFH—affects workplace communication and work outcomes for individual workers, we randomized the number of days that employees came into the office during a transitional return-to-office period. We focus on a sample of 130 employees from corporate HR working at the Dhaka headquarters. An additional 30 employees from the corporate microfinance team took part in the experiment but were excluded from our analyses as they did not consent to sharing email, our primary data source to measure work outcomes.

We ran our experiment for a total of nine weeks between July 5th and September 3rd, 2020. Using a random number generator, we selected which employees should come to the office each day.[7] Since the decision of who is supposed to come is randomized over the nine weeks, some employees were randomly assigned to come to the office for only a few days, whereas others were assigned to come for a higher number of days. The nine-week treatment period included 35 work days, exclusive of weekends and a mid-summer break during the religious festival of Eid. This is the maximum number of days a worker could be working from office. Each of the 130 employees in the sample followed this schedule closely and their attendance was taken daily by their managers. At the end of each work week, their attendance was verified by the team lead. At the end of each weekend, they were also provided with the randomized attendance assignment sheet to follow for the upcoming work week.[8] Panel A in Figure 1 plots the distribution of days in the office across

[6]See Figure A.1 in Section A.1 of the Appendix for a photograph of the BRAC headquarters.

[7]Bangladesh has a work week that commences on Sunday and ends on Thursday, with Friday and Saturday comprising the weekend.

[8]Given this experiment was conducted within an actual firm with full-time employees, a few exceptions to adhering to the attendance schedule were made for emergency reasons such as friends and family being sick. However, the number of days we use as a right-hand-side variable is the actual number, not the scheduled number,

employees in the human resources department.[9] Panel B plots the distribution of the hyperbolic sine of the number of emails across all dyads, displaying substantial variation.

[INSERT FIGURE 1 HERE]

We categorize workers into three groups based on how many days they were exogenously assigned to work from the physical office: high WFH (0-8 days in the office), intermediate WFH (9-14), and low WFH (15+). The cut-offs for the three groups are determined to equalize the number of total (i.e., HR and microfinance) workers in the experiment within each of the three sub-groups, but in Table A.7 of Section A.3.5 of the Appendix we show that our results are qualitatively robust to an alternative classification based off of equally-spaced bins within the HR unit (the unit for which we have email data).

Figure 2 summarizes the timeline of the experiment. We partition the time series into two periods: lockdown, consisting of 82 days starting March 26th, 2020; and post-lockdown (the "treatment" period), consisting of 62 days starting July 5th, 2020 and ending September 3rd, 2020. Again, owing to holidays and other common breaks, the 62-day treatment period contained 35 work days. In the bulk of the paper, we refer to the post-lockdown period, which contains the hybrid work arrangement, as the treatment period. We compare this to the lockdown period.

[INSERT FIGURE 2 HERE]

2.2 Outcome Measures

To measure intrafirm communication across dyads, we count the number of emails sent between each individual sender and receiver, for the lockdown and treatment periods. Building on the

so we are not just identifying an intent to treat.

[9] Figure A.3 in Section A.2 of the Appendix also presents a similar distribution of days for microfinance workers.

distribution of emails across dyads in Panel B in Figure 1, Figure A.4 in Section A.2 of the Appendix plots the distribution separately across both the lockdown and treatment periods.

To measure novelty of individual work outcomes, we use two sets of measures. Our primary method is based on coding the novelty of the text of employee emails and email attachments, using a clustering process and measuring cosine similarity pf document vectors to a synthetic representative document. Here, we use text analysis and machine learning methods, including Doc2vec (Le and Mikolov, 2014) and BIRCH clustering (Zhang et al., 1997). As an additional check, we also apply a cruder hashing function (see Section A.6 of the Appendix for details).

In supplementary analysis, we also code employee performance based on managerial ratings. These ratings were collected using survey instruments from the prior management literature on remote work, notably those employed by Greenhaus and Parasuraman (1993) and Touliatos et al. (1984). These surveys ask managers to rate their direct reports in terms of the following qualities on a 7-point scale ranging from (1) unsatisfactory to (7) excellent on the following measures: ability, cooperation, job knowledge, creativity, productivity, and quality of work.

3 Hybrid Work and Intra-firm Communication

Recent evidence from Yang et al. (2021) points to the fact that firm-wide WFH might lead to intra-firm communication patterns becoming static and siloed. However, we know little about how hybrid work affects patterns of intra-firm communication.

To quantify the causal effect of hybrid on intra-firm communication, we create a dyadic pair between an employee i and every other employee j over the lockdown and treatment periods, subsequently regressing the number of emails sent from employee i on the intensity of hybrid work while controlling for individual characteristics and allowing for heterogeneity by group:[10]

$$y_{ijt} = \gamma^I WFH_i^I + \gamma^H WFH_i^H + \xi^I(WFH_i^I \times G_i) + \xi^H(WFH_i^H \times G_i) + \beta X_{it} + \epsilon_{ij} \tag{1}$$

where y_{ijt} denotes the number of emails that the (i, j) dyad exchanged during the treatment, WFH^j for $j \in I, H$ denotes an indicator for whether the employee works an "intermediate" ($j = I$) or a "high" ($j = H$) amount from home, G denotes an indicator for whether an individual falls within a particular group (e.g., gender), and X denotes individual demographics, including an indicator for being male, being a non-manager, having a master's or PhD degree, being married, whether the spouse also works from home, and having to care for a child. While we do not need to control for these characteristics given our randomization strategy, they are nonetheless useful and help remove any unobserved heterogeneity that could be spuriously correlated with extent of WFH and email. Table A.2 of the Appendix motivates our empirical setup: co-location (i.e., when both the sender and receiver are in the office) is not strongly correlated with email activity; most of the variation comes from the number of days the sender is in the office.

While our use of dyadic data here comes at a disadvantage of greater sparsity in the number of emails sent between any dyad, we address that limitation by using a negative binomial model (although our results are qualitatively similar using a standard least square estimator).[11] Moreover,

[10] Interactions within a dyad are not independent, since the number of emails that employee i sends to j is dependent in part on the number that i receives from j (Quintane and Kleinbaum, 2011). We focus on the number of messages exchanged within a directed dyad. However, since employee i exchanges with not only j, but also j', there is a common person effect that could cause us to underestimate standard errors (Kenny et al., 2006). We address this challenge by also testing robustness in Table A.5 with robust standard errors that are clustered on both members of a dyad (Cameron et al., 2011; Kleinbaum et al., 2013).

[11] The distribution is highly censored: the median number of emails across dyads is zero and the mean is 1.28 (the standard deviation is 7.19, skewness 17.5, and kurtosis 450).

the dyadic data allows us to exploit pre-pandemic communication between any pair, meaning that we can test for pre-existing communication ties—that is, whether communication prior to the pandemic governs the direction of communication during the treatment period.

Our identifying variation comes from the comparison of email communication among dyads, based on whether the sender of the email is in the low-, intermediate-, or high-WFH group. We employ dummy variables for each, but our coefficients on the days in the office indicators are normalized to an indicator for the employee coming in 0-8 days. To justify this approach, Section A.3.2 of the Appendix presents results showing that the bulk of the relationship between emails and days in the office comes from days that the sender is in the office. We estimate Equation 1 using a negative binomial model to account for the excess of zero emails within dyads.

Panel A in Table 1 documents the results associated with Equation 1. Column 1 documents that non-managers, intuitively, have significantly fewer email exchanges: roughly 1.6 less for any given dyad. For perspective, the median is zero and the mean is 1.2. More importantly, intermediate WFH is associated with a 0.814 increase in the number of emails for a given dyad and low WFH is associated with a 0.537 increase, relative to the baseline of high WFH. We conduct a t-test to examine whether the two are statistically different from one another. We can reject the null hypothesis that they are equal with a p-value of 0.09. The difference becomes more precise as we add additional controls in the columns that follow.

We now proceed by sequentially adding additional demographic characteristics, including: an indicator for being male, an indicator for having a master's or PhD degree (normalized to having just a bachelor's degree), an indicator for being married, an indicator for whether their spouse also works from home, and an indicator for whether they have to care for a child. We find that men generally send more emails than women, but differences in educational attainment do not explain email communication. Married workers send fewer emails, but those who care for children send more. Importantly, however, our marginal effects of intermediate and low WFH remain economically and statistically significant: in our strictest specification (column 6), a sender's intermediate WFH is associated with 0.689 more emails in a dyadic pair and low WFH is associated with 0.364 more emails, relative to high WFH. Furthermore, we can reject the null hypothesis that the two are equal to each other (p-value $= 0.05$).

[INSERT TABLE 1 HERE]

Our estimates are statistically significant, but how economically meaningful are they? We now consider a simple aggregation exercise where we ask how email traffic would change if both the low- and high-WFH groups came in as often as the intermediate-WFH group. To start, we estimate the baseline regression again, but use the intermediate-WFH group as the normalization, producing coefficients of -0.689 for high WFH and -0.325 for low WFH. Next, we sum the number of emails across both groups, yielding 5,070 emails for high WFH and 2,551 for low WFH, and take the sum of their product with the respective elasticities. In sum, this produces 4,322 emails, which is 63% of the total intermediate WFH number of emails (6,805). The aggregation exercise abstracts from the dynamic social adjustment that would occur if everyone had hybrid work, but it still provides a useful interpretation for the economic significance of the effect.

So far, we have examined the effects of WFH on overall email volumes. However, we are also interested in the broader democratizing effects: how many employees are receiving emails from others with whom they did not previously correspond? Panel B in Table 1 documents these results by using the logged average number of unique recipients as the outcome variable. Starting in column 1, we see that intermediate WFH is associated with a statistically significant 58%

(= exp(0.462)) increase in the number of unique email recipients relative to the high-WFH baseline. While low WFH is also positively associated with unique email recipients, it is not statistically significant. Again, we see that non-managers send fewer emails to wider groups of people. We subsequently layer additional demographic controls, which do not alter the treatment effect in any meaningful way. In the strictest specification (column 6), we find that intermediate WFH is associated with a 50% increase in unique email recipients. However, because of the sample size, we cannot statistically reject the null hypothesis that the intermediate- and low-WFH categories have similar effects on the number of unique email recipients. Interestingly, however, we see that more educated employees send fewer emails to unique recipients, which could be consistent with a greater focus on cognitively-intensive activities that require less teamwork.

We now discuss a handful of robustness exercises. First, Table A.4 in Section A.3.4 of the Appendix replicates the main results using an alternative cutoff for intermediate WFH, demonstrating that employees coming into the office between 12-14 and 15-23 days in the treatment period send substantially more emails than their counterparts who come in 0-7 days. Table A.5 replicates the main results with the standard cutoff but uses two-way clustered standard errors on both the recipient and sender following Kleinbaum et al. (2013). Furthermore, Table A.7 allows for additional heterogeneity in the intensity of WFH by creating four equally-spaced bins of days in the office using the standard cutoff off all employees. We find that the bulk of the variation is concentrated in the third quartile, or those with 13-15 days in the office.

Second, we explore several additional parametric assumptions about our estimator and the potential for non-linearities in Section A.3.5 of the Appendix. Table A.8 replicates the main results using a Poisson distribution. Table A.9 investigates several dimensions of heterogeneity. Most of the variation in intermediate-WFH arrangements is driven by non-managers. We also find that dyadic communication during the lockdown prior to the treatment is associated with fewer emails among those in intermediate- and low-WFH arrangements, relative to high WFH, which is significant since it suggests that hybrid work can democratize communication within the firm: communication becomes less siloed among dyadic pairs that communicated prior to the treatment.

Next, Table A.10 documents qualitatively similar results when we collapse our data to the employee level. Because we are no longer working with dyadic pairs, we can now use the log number of emails sent as the outcome variable. These estimates suggest that intermediate WFH leads to 46.6% more emails relative to high WFH, whereas low WFH leads to only 10.8% fewer emails, although the difference is not statistically significant. Finally, Table A.11 shows that there is no statistically significant difference in email activity depending on whether an individual works from home earlier versus later in the week, ruling out a task-planning mechanism.[12]

In summary, the set of results reported above shed light on how intermediate levels of hybrid work (i.e., intermediate WFH) might be more efficient than both low WFH and high WFH in maximizing the number of emails sent by employees and the number of unique recipients of emails (the span or reach of the sender's network). These results complement recent research on how firm-wide remote work affects email communication in the workplace, notably Yang et al. (2021). However, we do not interpret volume of email as a measure of productivity. Instead, we next turn to study how the extent of hybrid work affects worker productivity as measured using the novelty of work products (emails and email attachments) generated by individuals.

[12]Rather than computing the number of days that employees come into the office over the three different periods, we can also exploit daily variation. Section A.3.1 of the Appendix presents regressions of email communication on an indicator for whether the sender is in the office. Unlike earlier with our negative binomial model, we estimate a linear probability model for whether the dyad emailed on a particular day as a function of WFH, controlling for dyad and day fixed effects. We find that WFH is associated with a 0.4 percentage point decline in the probability of whether the employee sent a particular recipient an email.

4 Hybrid Work and Novelty of Work Products

We now examine the effects of extent of hybrid work on novelty of work products by using the text of emails and work-related attachments to emails. Our examination is motivated by managerial concerns around the effectiveness of hybrid work in contexts that require creative and novel work. As with "creativity," the measurement of "novelty" depends on the context. In past literature, novelty has been measured by quantifying the presence of new keywords and combinations of words that have not been seen in prior work products (Boudreau et al., 2016).

We innovate using cutting-edge machine learning and text analyses methods to generate an approximation of novelty. We first pre-process the text of emails and attachments sent by each employee. Then, in a three-step procedure, we score each document for novelty. First, we use Doc2vec paragraph embeddings to represent each document as a fixed-length, numeric vector, much as in principal component analysis. Doc2vec considers not only the rarity of particular words, but also the context, such as surrounding words, to determine the key components of each document. Next, we use a clustering algorithm (BIRCH, a computationally-optimized version of K-means) to group these vectors into different clusters. For example, one cluster might retrospectively represent "strategy reports," although the labels are not pre-determined. Finally, we calculate the cosine similarity between each document and a synthetic centroid document within each cluster, and transform this into a novelty measure by taking the additive inverse and mapping the range to [0, 1]. Thus, a very unusual or novel document would be given a high score for novelty.

The clustering problem involves parametrizations similar to the well-known mean-variance tradeoff. We want to categorize the type of each document so we can compare it to others of its type to determine novelty. But as the number of document types increases, then variation in the novelty scores is lost. Thus, we optimize the vector length and number of clusters to select the most appropriate parameters. Details of the calculation are found in Section A.6 of the Appendix.

The results of this process indicate a striking rise in novelty associated with intermediate WFH, significant at the 5% level. Consider the interpretation of the intermediate-WFH coefficient in column (1) of Table 2. The regression indicates that intermediate-WFH workers increase this novelty measure by 0.529 standard deviations more than a high-WFH worker would. The effect persists even in the presence of demographic, behavioral, and colocation controls.[13]

[INSERT TABLE 2 HERE]

Table A.14 in Section A.4 of the Appendix presents additional evidence on the positive productivity effects of intermediate WFH using managerial performance ratings. We also apply a cruder, alternative novelty measure, detailed in Section A.7.2 of the Appendix, which considers only attachment text and is broadly consistent with these results.

5 Understanding Mechanisms

To test plausible mechanisms, we build on prior literature that has argued that intermediate hybrid work could represent the "best of both worlds" in relation to two underlying mechanisms: flexibility and isolation. Remote work offers workers lower distraction, less commuting, and more flexibility (Bloom et al., 2014a), and greater flexibility is related to employees generating novel and creative work (Hackman et al., 1975; Amabile et al., 1996). On the flip side, remote work has

[13] In unreported results, the effect also holds when we control for the amount of time between the first and last email sent on a given day, which is an unreliable proxy for time spent working.

been shown to increase isolation from colleagues (Raghuram et al., 2001) and isolation has been argued to negatively affect work outcomes (Golden et al., 2008a). Intermediate hybrid might offer the best of both worlds: flexibility without isolation.

To test this, we draw on answers to several survey questions that we designed to understand employee attitudes about remote work. We designed this survey by leveraging prior organizational literature in remote work, notably Raghuram et al. (2001). Each question is ranked on a scale of one to seven (1 = strongly disagree, 7 = strongly agree). First, drawing on validated survey questions from Raghuram et al. (2001), we ask: (i) "Overall, I am satisfied with working from home," (ii) "Working from home allows me to perform my job better than I ever could when I worked in the office," (iii) "If I were now given the choice to return to a traditional office environment (i.e., no longer telework), I would be very unlikely to do so," (iv) "Since I started working from home, I have been able to balance my job and personal life," and (v) "Since I started working from home, my productivity has increased." These questions help us understand whether employee engagement and/or reallocation of tasks are major mechanisms. We also draw on validated survey questions from Golden et al. (2008b): (i) "I feel left out on activities and meetings that could enhance my career," (ii) "I miss out on opportunities to be mentored," (iii) "I feel out of the loop," (iv) "I miss face-to-face contact with coworkers," (v) "I feel isolated," (vi) "I miss the emotional support of coworkers," and (vii) "I miss informal interactions with others." These questions help us understand whether social isolation is at play.

Table 3 documents these results and suggests that workers in the intermediate-WFH group report greater satisfaction with working from home, greater work-life balance, and lower isolation compared to both workers in the high- and low-WFH categories.[14]

[14]In addition, Table A.12 in Section A.4 of the Appendix reports similar results when we define the cutoff for

[INSERT TABLE 3 HERE]

In the Appendix (Table A.11), we also attempt to rule out alternative mechanisms and report results that show cognate outcomes did not differ based on whether the worker attended office earlier versus later in the work week, allowing us to rule out an alternative explanation of our results related to more efficient planning and management of time. Intermediate hybrid work may lead to a more efficient planning and allocation of time towards heterogeneous tasks—those that require social interaction and others that require independent work and concentration.

While all workers were made aware of the randomized weekly schedule on Saturday night (i.e., prior to the work week commencing), workers who attended the office later in the week (i.e., on Wednesday and Thursday) arguably had more time to plan activities such as meetings with colleagues who were also attending the office on the same days as they were, compared to workers who attended office earlier in the work week (i.e., on Sunday and Monday). Yet we find no differences in the number of emails sent based on the day of the week the worker attended the office. We interpret this as evidence suggestive that better planning and more efficient time allocation to heterogeneous tasks were not in play in our setting.[15]

While our RCT allows us to quantify a causal effect of varying degrees of hybrid work on intra-firm communication and novelty of work products at the individual level within a real-world setting, our results are nonetheless obtained from a single organization in a lower middle-income country. Section A.5 of the Appendix reports additional results that draw on nationally-representative data from Gallup, for 12,166 individuals in the United States, between March 2020

remote work based on the set of employees in the human resources division alone (that is, excluding microfinance). Furthermore, Table A.13 replicates the main results using cutoffs that are unique to the HR employees.

[15] In unreported results, we also detected no significant trends relating to the number of times an employee switches work modalities within the week.

and June 2021. Although we do not have a source of exogenous variation, we estimate models that regress one-to-five indices of employee job satisfaction and preparedness on the degree of remote work (high, medium, low) interacted with measures of organizational culture, controlling for a wide array of demographic characteristics, income, occupation, and time fixed effects. Table A.17 shows that, while remote work is associated with higher levels of job satisfaction and preparedness, organizational culture is an important mediator. For example, among employees who feel more of a sense of purpose or receive more mentorship, high WFH is much less associated with job satisfaction and preparedness, whereas intermediate-WFH workers resemble those who work in person. This is consistent with our experimental findings that intermediate WFH is positively associated with employee satisfaction towards remote work and other individual outcomes.

6 Conclusion

The COVID-19 pandemic has led to a fundamental transformation in the way work is conducted, with hybrid work emerging as an option for organizing work within firms. While there is a vigorous managerial and policy debate around hybrid work, to the best of our knowledge, there is not yet causal evidence on how hybrid work on workplace communication and work outcomes.

We provide the first causal evidence on the effects of the extent of hybrid work on intrafirm communication and novelty of work products generated by workers. Our results exploit a field experiment conducted in collaboration with a large firm in Bangladesh, randomizing the number of days that each employee comes into the office over a nine-week period. Drawing on every email that was sent during the treatment period, as well as the pre-treatment (lockdown) period by workers in the experiment, we measure intrafirm communication. We also use a novel measure of worker performance using text of emails and email attachments. To recap, we find that intermediate levels of WFH correlate with the highest number of emails sent, highest number of email recipients, more positive sentiment of emails, and greater novelty of work products. Our mechanism test suggests that hybrid might represent the best of both worlds: flexibility without isolation.

While our results are not without limitations, specifically, novelty of work-related products coded from email text is not a perfect measure of worker productivity, they provide important guidance for the transition to hybrid work. Among other questions, future research should explore the productivity effects of hybrid work in a wide variety of contexts; study whether, and under what conditions, intermediate levels of WFH correlate with effective mentoring outcomes for workers; and explore how adoption of intermediate WFH might change the geography of work.

References

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., and Herron, M. (1996). Assessing the work environment for creativity. *Academy of management journal*, 39(5):1154–1184.

Barerro, J., Bloom, N., and Davis, S. J. (2022). WHY WORKING FROM HOME WILL STICK. *NBER working paper*.

Bartel, C. A., Wrzesniewski, A., and Wiesenfeld, B. M. (2012). Knowing Where You Stand: Physical Isolation, Perceived Respect, and Organizational Identification Among Virtual Employees. *Organization Science*, 23(3):743–757.

Bloom, N., Davis, S. J., and Zhestkova, Y. (2021). COVID-19 Shifted Patent Applications Toward Technologies that Support Working from Home. SSRN Scholarly Paper ID 3695191, Social Science Research Network, Rochester, NY.

Bloom, N., Garicano, L., Sadun, R., and Van Reenen, J. (2014a). The Distinct Effects of Information Technology and Communication Technology on Firm Organization. *Management Science*, 60(12):2859–2885.

Bloom, N., Lemos, R., Sadun, R., Scur, D., and Van Reenen, J. (2014b). The new empirical economics of management. *Journal of European Economic Association*, 12(4).

Bloom, N., Liang, J., Roberts, J., and Ying, Z. J. (2014c). Does working from home work? *Quarterly Journal of Economics*, 130(1):165–218.

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., and Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, 62(10):2765–2783.

Bramoulle, Y., Currarinic, S., Jackson, M. O., Pin, P., and Rogers, B. W. (2012). Homophily and long-run integration in social networks. *Journal of Economic Theory*, 147(5):1754–1786.

Brynjolfsson, E., Horton, J., Makridis, C. A., Mas, A., Ozimek, A., Rock, D., and TuYe, H.-Y. (2021). Measuring the Incidence of Remote Work. *Mimeo*.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust Inference With Multiway Clustering . *Journal of Business and Economic Statistics*, 29(2):238–249.

Choudhury, P. (2020). Our work-from-anywhere future. *Harvard Business Review*, 98(6):58–+.

Choudhury, P., Foroughi, C., and Larson, B. (2020). Work-from-anywhere: The productivity effects of geographic flexibility. *Strategic Management Journal*.

Cutter, C. (2021). Remote Work May Now Last for Two Years, Worrying Some Bosses. *Wall Street Journal*.

Eisen, B. (2019). Workers Are Fleeing Big Cities for Smaller Ones–and Taking Their Jobs With Them . *Wall Street Journal*.

Elsbach, K. D. and Hargadon, A. B. (2006). Enhancing creativity through “mindless” work: A framework of workday design. *Organization science*, 17(4):470–483.

Gaspar, J. and Glaeser, E. L. (1998). Information technology and the future of cities. *Journal of Urban Economics*, 43:136–156.

Gibbs, M., Mengel, F., and Siemroth, C. (2021). Work from Home & Productivity: Evidence from Personnel & Analytics Data on IT Professionals. *BFI working paper.*

Glaeser, E. L. (1998). Are cities dying. *Journal of Economic Perspectives*, 12(2):139–160.

Golden, T. D., Veiga, J. F., and Dino, R. N. (2008a). The impact of professional isolation on teleworker job performance and turnover intentions: does time spent teleworking, interacting face-to-face, or having access to communication-enhancing technology matter? *Journal of Applied Psychology*, 93(6):1412.

Golden, T. D., Veiga, J. F., and Dino, R. N. (2008b). The impact of professional isolation on teleworker job performance and turnover intentions: Does time spent teleworking, interacting face-to-face, or having access to communication-enhancing technology matter? *Journal of Applied Psychology*, 93(6):1412–1421.

Golub, B. and Jackson, M. O. (2012). How Homophily Affects the Speed of Learning and Best-Response Dynamics. *Quarterly Journal of Economics*, 127(3):1287–1338.

Greenhaus, J. H. and Parasuraman, S. (1993). Job performance attributions and career advancement prospects: An examination of gender and race effects. *Organizational Behavior and Human Decision Processes*, 55(2):273–297.

Hackman, J. R., Oldham, G., Janson, R., and Purdy, K. (1975). A new strategy for job enrichment. *California Management Review*, 17(4):57–71.

Hutto, C. J. and Gilbert, E. E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Eighth International Conference on Weblogs and Social Media (ICWSM-14), Ann Arbor, MI, June 2014.*

Kenny, D. A., Kashy, D. A., and Cook, W. L. (2006). Dyadic data analysis. *Guilford Press*.

Khanna, T. and Ramachandran, S. (2021). BRAC in 2020. *HBS case study*.

Kleinbaum, A. M., Stuart, T. E., and Tushman, M. L. (2013). Discretion Within Constraint: Homophily and Structure in a Formal Organization. *Organization Science*, 24(5):1316–1336.

Le, Q. and Mikolov, T. (2014). Distributed representations of sentences and documents. In *International conference on machine learning*, pages 1188–1196. PMLR.

Poleg, D. (2021). The Future of Offices When Workers Have a Choice. *New York Times*.

Quintane, E. and Kleinbaum, A. M. (2011). Matter over mind? Email data and the measurement of social networks. *Connections*, 31(1):22–46.

Raghuram, S., Garud, R., Wiesenfeld, B. M., and Gupta, V. (2001). Factors contributing to virtual work adjustment. *Journal of Management*, 27(3):383–405.

Rivest, R. (1992). The md5 message-digest algorithm. Technical Report RFC 1321, Internet Engineering Task Force.

Storper, M. and Venables, A. J. (2004). Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, 4(4):351–370.

Teevan, J. (2021). The new future of work: Research from Microsoft into the Pandemic's Impact on Work Practices. *Microsoft*.

Touliatos, J., Bedeian, A. G., Mossholder, K. W., and Barkman, A. I. (1984). Job-related perceptions of male and female government, industrial, and public accountants. *Social Behavior and Personality: an international journal*, 12(1):61–68.

Yang, L., Holtz, D., Jaffe, S., Suri, S., Sinha, S., Weston, J., Joyce, C., Shah, N., Sherman, K., Hecht, B., and Teevan, J. (2021). The effects of remote work on collaboration among information workers. *Nature Human Behaviour*.

Zhang, T., Ramakrishnan, R., and Livny, M. (1997). Birch: A new data clustering algorithm and its applications. *Data Mining and Knowledge Discovery*, 1(2):141–182.

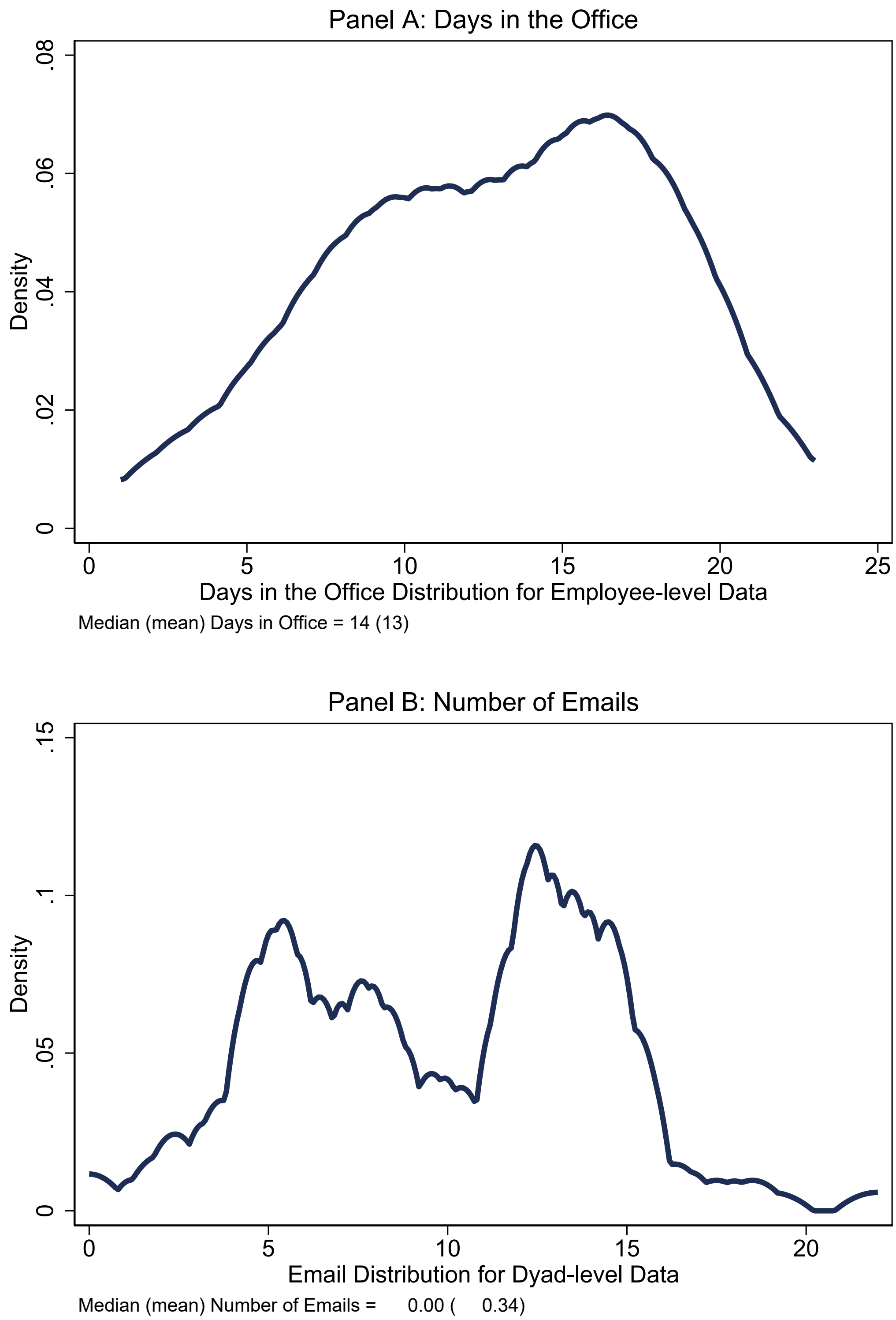

Figure 1: Distribution of Days in the Office and Emails

Notes.—Source: Authors. Panel A in the figure plots the distribution of the number of days that a person comes into the office. Panel B in the figure plots the hyperbolic sine of the number of emails sent in the dyadic data. Our sample is restricted to the employees in the human resources department.

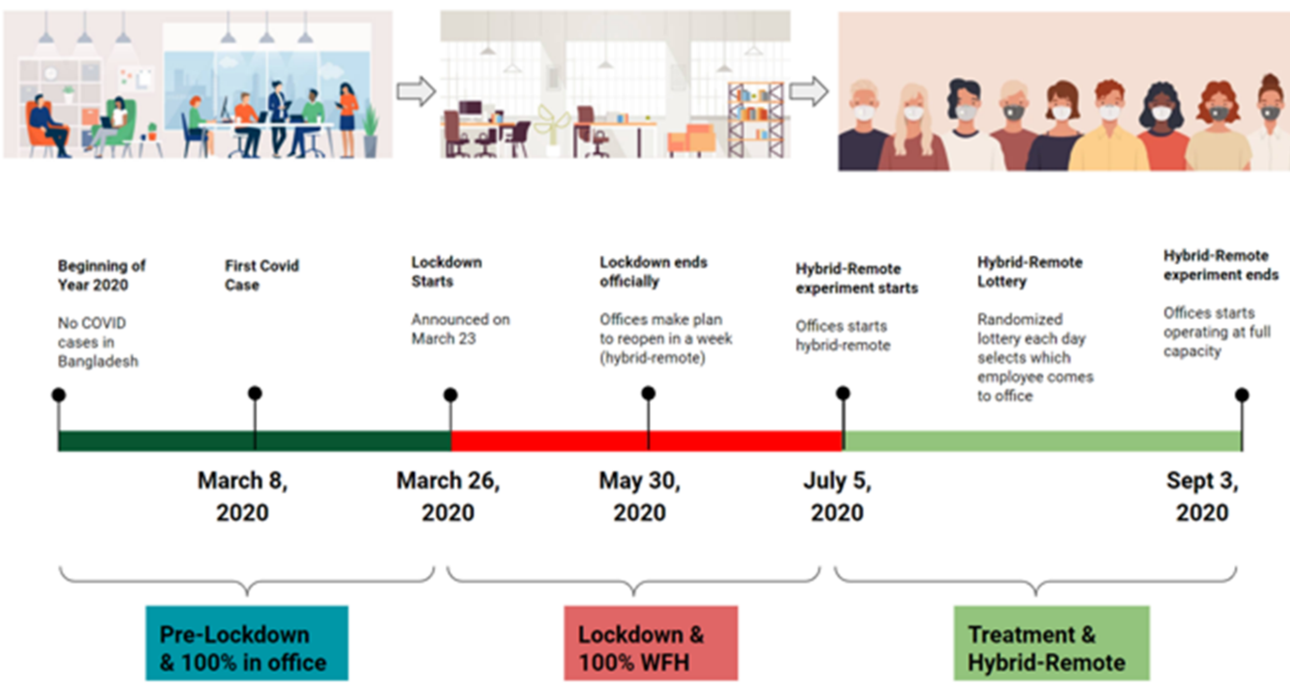

Figure 2: Timeline of the Experimental Design

Notes.—Source: Authors. The figure plots the timeline of the experimental design, ranging from the pre-lockdown period to the hybrid work period.

Table 1: Intensity of Working-from-Home and Email Behavior

	Dep. var. = Number of Emails Sent					
	(1)	(2)	(3)	(4)	(5)	(6)
Dyadic Data						
Intermediate WFH	.814***	.781***	.758***	.716***	.710***	.689***
	[.178]	[.182]	[.186]	[.171]	[.170]	[.171]
Low WFH	.537***	.493***	.457***	.421***	.379**	.364**
	[.151]	[.152]	[.162]	[.160]	[.161]	[.161]
Non-Manager	-1.608***	-1.558***	-1.555***	-1.563***	-1.538***	-1.419***
	[.217]	[.215]	[.217]	[.216]	[.209]	[.205]
Male		.185	.192	.259*	.260*	.334**
		[.136]	[.140]	[.134]	[.133]	[.137]
Masters/PhD			-.139	-.003	-.018	-.062
			[.180]	[.204]	[.203]	[.208]
Married				-.384	-.337	-.554*
				[.273]	[.274]	[.289]
Spouse WFH					-.127	-.111
					[.138]	[.138]
Caring for Child						.447***
						[.135]
Sample Size	10600	10600	10600	10600	10600	10600
	Dep. var. = log(Unique Recipients of Emails)					
Employee Data						
Intermediate WFH	.462***	.449***	.404**	.425***	.422**	.406**
	[.155]	[.162]	[.164]	[.160]	[.161]	[.164]
Low WFH	.299	.271	.212	.227	.262	.246
	[.200]	[.214]	[.209]	[.209]	[.207]	[.208]
Non-Manager	-.612**	-.587**	-.555**	-.554**	-.561**	-.502*
	[.275]	[.270]	[.272]	[.273]	[.268]	[.267]
Male		.096	.130	.104	.103	.144
		[.153]	[.152]	[.151]	[.152]	[.154]
Masters/PhD			-.357**	-.396**	-.394**	-.400**
			[.163]	[.176]	[.178]	[.175]
Married				.184	.133	.047
				[.214]	[.221]	[.219]
Spouse WFH					.141	.133
					[.159]	[.161]
Caring for Child						.180
						[.152]
Sample Size	99	99	99	99	99	99

Notes.—Source: Authors. Panel A in the table reports the coefficients associated with negative binomial regressions of the number of emails sent during the treatment period on an indicator for whether the email sender came in between 9-14 days in the office (intermediate WFH), an indicator for 15-23 days in the office (low WFH), controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. Panel B in the table reports comparable models, but now using the log number of unique email recipients that a given employee has emailed with the same right-hand-side variables as Panel A. Standard errors are clustered at the dyad-level in Panel A and at the employee-level in Panel B.

Table 2: Relationship Between Work Novelty and the Intensity of Working-from-Home

	Dep. var. = Change in Cosine Similarity-Based Work Product Index						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intermediate WFH	0.529***	0.505***	0.590***	0.580**	0.583**	0.583**	0.669***
	(0.197)	(0.192)	(0.218)	(0.237)	(0.238)	(0.236)	(0.211)
Low WFH	0.244	0.152	0.228	0.227	0.200	0.195	0.283
	(0.206)	(0.211)	(0.225)	(0.225)	(0.240)	(0.246)	(0.517)
Non-Manager	−0.793***	−0.749***	−0.875***	−0.882***	−0.875***	−0.839***	−0.886***
	(0.266)	(0.258)	(0.254)	(0.255)	(0.259)	(0.268)	(0.326)
Male		0.257	0.181	0.190	0.189	0.220	0.281
		(0.179)	(0.178)	(0.170)	(0.169)	(0.181)	(0.231)
Masters/PhD			−0.004	0.011	0.013	0.012	0.063
			(0.331)	(0.301)	(0.304)	(0.305)	(0.294)
Married				−0.074	−0.037	−0.102	−0.082
				(0.310)	(0.326)	(0.304)	(0.308)
Spouse WFH					−0.095	−0.106	−0.096
					(0.190)	(0.200)	(0.229)
Caring for Child						0.148	0.131
						(0.204)	(0.255)
Colocation Intensity							−0.090
							(0.219)

Notes.—Source: Authors. This table reports the coefficients associated with OLS regressions of the change in the work novelty measure on an indicator for whether the email sender came to the office between 9-14 days (intermediate WFH), 15-23 days (low), or fewer than nine days (high, omitted category) controlling for the following demographic characteristics: male, non-manager (employee), education (masters/PhD–normalized to having a bachelor's), married, spouse works from home, employee has to care for a child, and standardized colocation intensity (the standardized number of days a person is physically colocated in the office with another team member). The dependent variable represents the change in standard deviations from the baseline to the experimental period. Text data was classified into eight clusters (BIRCH-based) from a Doc2vec vector length of four; see Section A.7.1 of the Appendix for a discussion of the parameter fit.

Table 3: Intensity of Working-from-Home and Employee Attitudes and Engagement

	Job Satisfaction	Better Balance	Prefer WFH
	(1)	(2)	(3)
Panel A			
Intermediate WFH	.668**	.803**	-.168
	[.308]	[.355]	[.389]
Low WFH	-.199	-.215	-.116
	[.356]	[.393]	[.415]
R-squared	.16	.11	.12
Sample Size	143	143	143
	Feel Left Out	Miss Mentorship	Feeling Isolated
	(1)	(2)	(3)
Panel B			
Intermediate WFH	.287	-.018	-.784*
	[.388]	[.390]	[.434]
Low WFH	.446	.670	.262
	[.413]	[.429]	[.500]
R-squared	.08	.05	.11
Sample Size	143	143	143

Notes.—Source: Authors. The table reports the coefficients associated with regressions of indices of employee preferences (ranging from one to seven) on an indicator for whether the email sender came in between 9-14 days in the office (intermediate WFH), an indicator for 15-23 days in the office (low WFH), controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. In Panel A, we draw on indices from Raghuram et al. (2001): "Overall, I am satisfied with working from home," "Since I started working from home, I have been able to balance my job and personal life," and "If I were now given the choice to return to a traditional office environment (i.e., no longer telework). In Panel B, we draw on indices from Golden et al. (2008b): "I feel left out on activities and meetings that could enhance my career," "I miss out on opportunities to be mentored," and "I feel isolated." Standard errors are heteroskedasticity-robust.

A Online Appendix

Contents

A.1 Data Preparation

BRAC is the world's largest non-governmental organization, headquartered in Bangladesh. Founded over four decades ago, the firm has over 35,000 staff as of September 2020 and over $1 billion in total income. In fact 81% of BRAC's revenues came from earned income in 2019 (Khanna and Ramachandran, 2021). While it is headquartered in Dhaka, BRAC has operations in multiple countries including Myanmar, Liberia, Sierra Leone, Uganda, and Rwanda. BRAC headquarters workers, the focus of our study, work in a modern office in Dhaka. Figure A.1 gives a visual representation of the entry to the BRAC headquarters.

Figure A.1: BRAC Headquarters

Notes.—The figure plots the entry to the BRAC headquarters.

Email data was provided by the firm's IT department. Because users frequently manipulate their inboxes and other incoming mail folders, we looked exclusively at emails in the sent mailbox—that is, emails sent by a given employee. In addition to being generally more reliable, this also prevented any accidental duplicate-counting of a given message. The email text was then processed to remove disclaimers and boilerplate text, individual employees' email signatures, and inline replies. Consequently, the word counts reported in this paper represent substantive text in the body of each message. In the case of multipart messages (e.g., with messages containing both HTML and text representations of the body), we utilized a common Python library that selects the part most likely to contain the true body. The small number of emails containing Bangla text, identified by the character set, were translated to English using Google Translate APIs.

Generally speaking, email counts and other metrics reported at the level of a sender (employee) count each email message only once, regardless of the number of recipients. Because emails can contain multiple recipients within and across the To, CC, and BCC fields, the dyadic data multiply counts single emails if they are sent to multiple recipients (although, within a given dyad, no individual message is ever counted more than once).

A.2 Descriptive Statistics

Figure A.2 documents summary statistics for the sample during the treatment period, which ran from 7/5/20 until 9/3/20. We see that 53% of the sample is male. We also find an average of 136 emails. However, the standard deviation is nearly as large as the mean, showing that there is significant heterogeneity across employees. The mean number of days an employee is in the office is 9.86 (and the median is 10). Most employees are married and highly educated, but fewer have a spouse who also works from home. We also show the correlations across the variables.

Figure A.2: Summary Statistics in Email Data over Lockdown

	Mean	SD	Male	Non-Manager	Masters/ PhD	Married	Spouse WFH	Care for Child	Email Novelty	No. of Emails	Unique Recipients	Days in Office	High WFH	Intermediate WFH	Low WFH
Male	0.53	0.5	1												
Non-Manager	0.87	0.34	-0.12	1											
Masters/PhD	0.86	0.35	0.08	0.012	1										
Married	0.86	0.35	0.2	-0.07	0.25	1									
Spouse WFH	0.29	0.49	0.028	-0.01	0.07	0.26	1								
Care for Child	0.45	0.5	-0.14	-0.19	0.07	0.31	0.13	1							
Email Novelty	-0.018	0.96	0.05	-0.19	-0.14	-0.13	-0.1	0.04	1						
No. of Emails	136	138	0.15	-0.43	-0.03	-0.05	-0.09	0.19	0.51	1					
Unique Recipients	21.5	15	0.05	-0.26	-0.18	-0.01	0.05	0.19	0.49	0.69	1				
Days in Office	9.86	4.55	0.17	0.22	-0.16	-0.18	-0.24	-0.14	0.18	0	0.06	1			
High WFH	0.4	0.49	-0.15	-0.29	0.15	0.15	0.11	0.06	-0.18	-0.06	-0.13	-0.85	1		
Intermediate WFH	0.4	0.49	0.013	0.19	-0.08	-0.14	0.06	-0.02	0.25	0.1	0.12	0.32	-0.66	1	
Low WFH	0.2	0.4	0.17	0.12	-0.08	-0.01	-0.22	-0.05	-0.08	-0.05	0.02	0.64	-0.41	-0.41	

Notes.—Source: Authors' experimental design. Using the individual-level (during the intervention period), the table reports the mean and standard deviation for demographic characteristics, as well as the correlations, across all the variables: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child, the number of emails, unique recipients, the number of days in the office (from the sender), and the share under high WFH (1-8 days in the office), intermediate WFH (9-14 days in the office), and low WFH (15-23 days in the office).

Next, Figure A.3 builds on the results from Figure 1 in the main text, but presents the distribution of days in the office for both microfinance and human resources departments. Importantly, the distributions overlap significantly, mitigating concerns about differential selection and/or lack of external validity for our treatment effects between the two sets of w orkers.

Next, Figure A.4 plots the distribution separately across our three major time periods. We see that there are substantially more emails sent by the average employee during the lockdown versus pre-lockdown period, but almost as many post-lockdown as during. We also observe a much larger variance across employees during the lockdown (129) versus before or after (121 and 89).

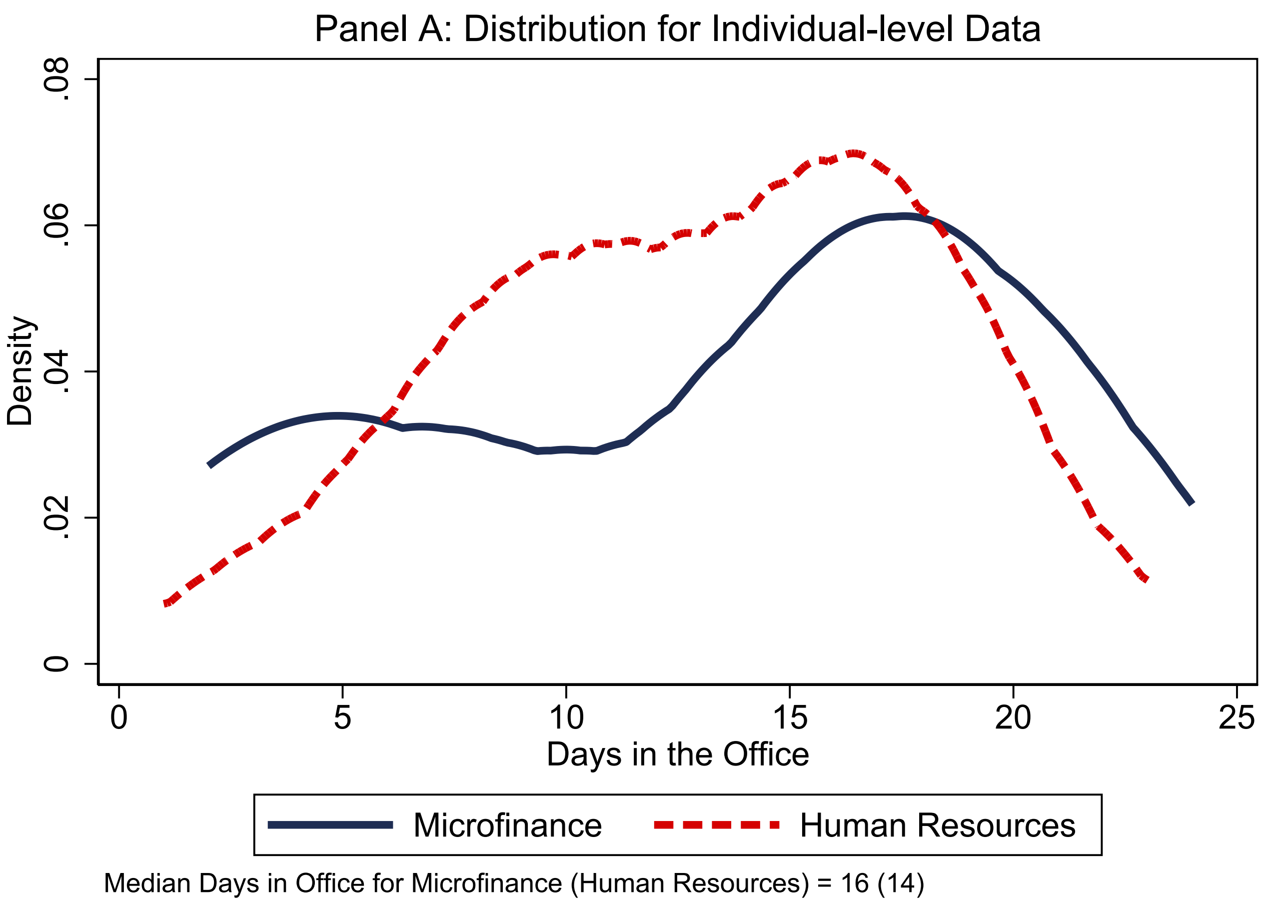

Figure A.3: Distribution of Days in the Office

Notes.—Sources: The figures plot the distribution of the number of days that a person comes into the office in the individual-level data for both microfinance and human resources departments.

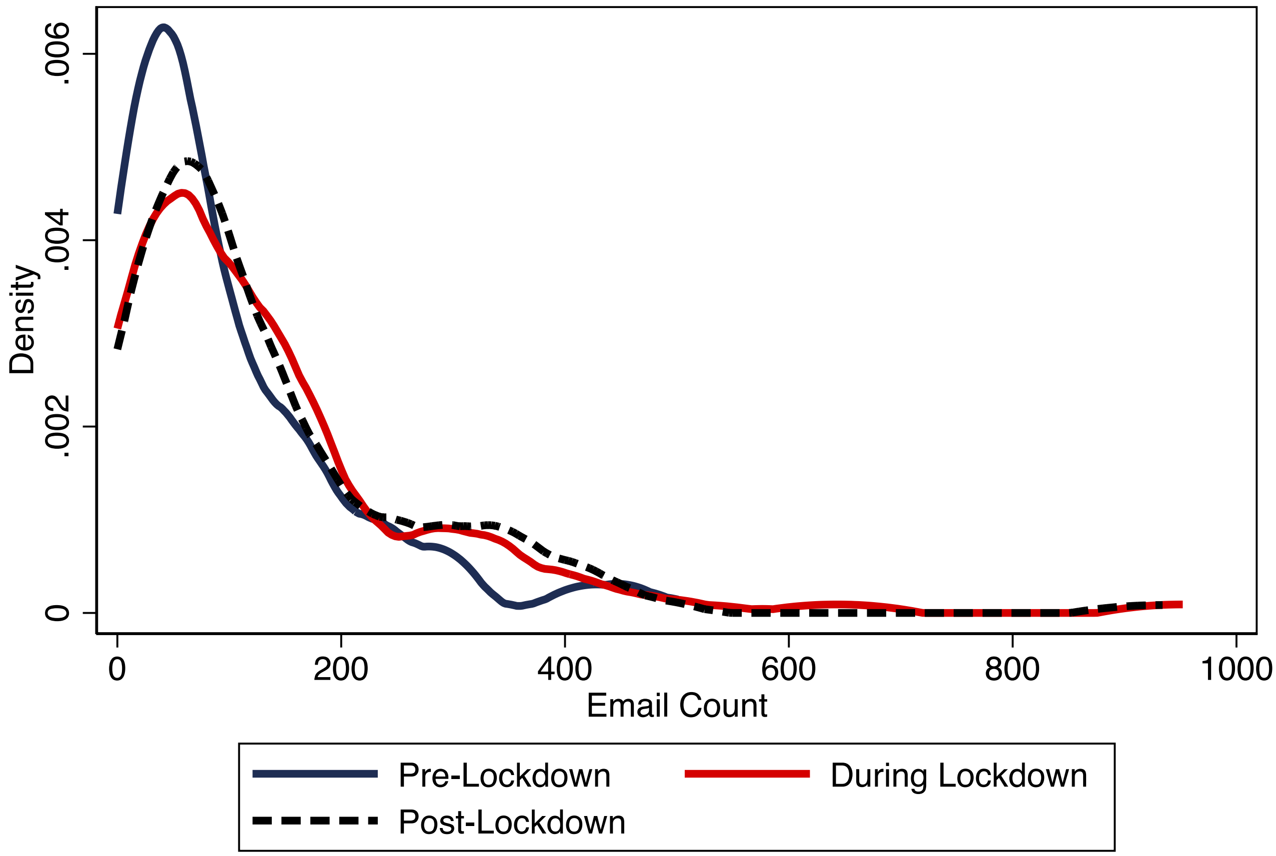

Figure A.4: Distribution of Emails Across Employees Over Time

Notes.—Source: Authors. The figure plots the distribution of emails sent pre-, during, and post-lockdown.

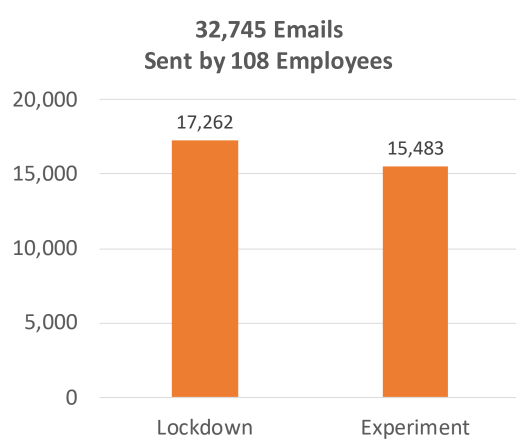

Figure A.5: Email Count

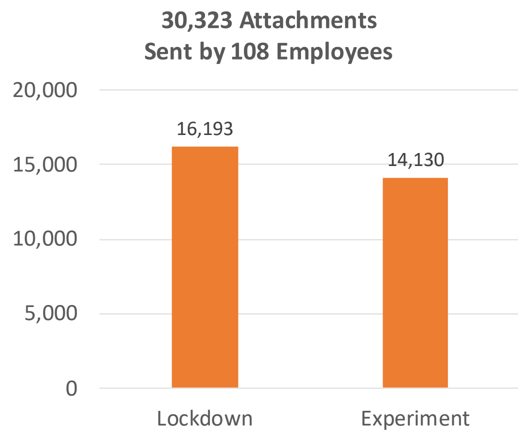

Figure A.6: Attachment Count

*Omits 128 malformed attachments (garbled or unrecognizable type)

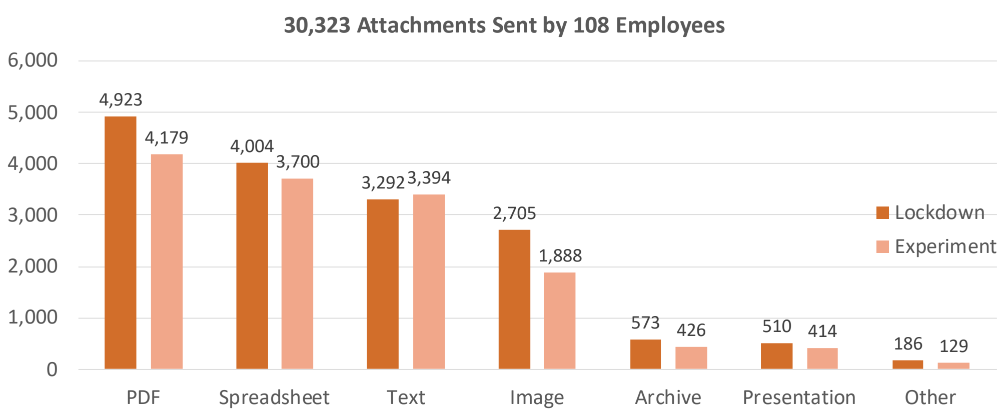

Figure A.7: Attachment Types

*Text analysis was performed only on extractable content from PDF, spreadsheet (e.g., Excel), text (e.g., Word), and presentation (i.e., PowerPoint) files. No OCR was attempted on images or embedded images.

A.3 Supplement to the Main Results with Email Traffic

A.3.1 Daily Patterns in Remote Work and Emails

While the main text focuses on remote work patterns over the entire treatment period, we now explore how remote work and email communication vary on a daily basis. In particular, we create a dyadic pair between each employee i and every other employee j, subsequently regressing an indicator for whether a dyadic pair interact on a given day on an indicator for whether the sender is working remotely that day, controlling for dyadic and time fixed effects:

$$y_{ijt} = \gamma r_{it} + \eta_{ij} + \lambda_t + \epsilon_{ijt} \tag{2}$$

where y_{ijt} denotes an indicator for whether the (i, j) dyad has exchanged emails on day t during the intervention, r denotes an indicator for whether the sender is working remotely, and η and λ denote fixed effects on dyad and time. Because of the randomization that we have built into our experiment, our estimate of γ has a causal interpretation. The identifying variation comes from the comparison of employees who were randomly assigned to come in on certain days over others, inducing variation in the pairs of employees that entered jointly as well.

Table A.1 documents these results. Starting with column 1, which presents the raw correlation, we see that employees are 0.4 percentage points less likely to send an email on a day that they are working remotely, suggesting that in-person work and email communication are complements, rather than substitutes. Column 2 introduces dyad and time fixed effects, thereby exploiting variation among pairs of employees who do not communicate via email on some days, but do on others. Here, we find a 0.2 p.p. decline in the probability of sending an email. Column 3 allows for heterogeneity in the gender of the sender, but there are no statistically significant differences between men and women. Column 4 allows for heterogeneity in the recipient of the email, specifically whether the email is sent to a manager. While managers are much more likely to receive emails in general, we find an 8 p.p. lower probability of sending an email among employees when they work remotely. In contrast, column 5 shows that email communication is slightly higher towards direct reports on remote work days, but the coefficient is not statistically significant.

Table A.1: Relationship Between Daily Email Communication and Remote Work

	Sent Emails Today				
	(1)	(2)	(3)	(4)	(5)
Sent to Manager				.292***	
				[.032]	
Sent to Direct Report					.112***
					[.036]
Remote Day	-.004***	-.002***	-.003***	-.001***	-.002***
	[.001]	[.000]	[.001]	[.000]	[.000]
× Male			.002		
			[.001]		
× Sent to Manager				-.082***	
				[.021]	
× Sent to Direct Report					-.005
					[.025]
R-squared	.00	.01	.01	.05	.02
Sample Size	648570	648570	648570	648570	648570
Dyad FE	No	Yes	Yes	Yes	Yes
Day FE	No	Yes	Yes	Yes	Yes

Notes.—Source: Authors. The table reports the coefficients associated with regressions of an indicator for whether person i and recipient j traded emails on day t on an indicator for whether the employee came into the office on day t and dyad and day-of-the-year fixed effects. The sample is restricted to the dates of the experiment in July and August. Standard errors are clustered at the dyad level.

A.3.2 Examining the Role of the Sender and Recipient of Emails

The main text presents the baseline results that allow non-linearities between the number of days the sender is in the office and emails sent. We now provide two alternative measurement strategies: (a) the number of days that employee i is in the office with employee j during the treatment period, and (b) the number of days that the sender is in the office. We argue that both these results point towards the focus on non-linearities through our three-bin approach.

Table A.2 documents these results. Starting with demographics, we see that men tend to send an additional 0.49 more emails within a dyadic pair, whereas non-managers tend to send 1.06 fewer emails within a dyadic pair. Those with a master's or PhD degree are not statistically more likely to send more emails, but those who are married send 0.82 fewer emails. Interestingly, caring for a child is associated with 0.53 *more* emails between any dyadic pair.

We now focus on our main remote work variables. Starting with our measure of co-location, increases in the number of days that a dyad is jointly in the office is negatively associated with the number of emails exchanged, but just barely and not in a statistically significant way. We find a slight positive interaction effect between co-location and male gender, but again it is not statistically significant. We find a positive interaction between co-location for non-managers, but it is only statistically significant at the 10% level. We find no meaningful variation when we allow for heterogeneity between the treatment (post-lockdown) and pre-treatment (lockdown) periods, but we do find a strong negative interaction effect for co-location among dyads that had talked prior to the pandemic, which is again consistent with a democratizing effect of remote work on email communication. These results are in contrast to what we might normally expect based on theories of homophily whereby similar individuals associate more (e.g., those who communicated prior to the treatment) (Bramoulle et al., 2012; Golub and Jackson, 2012).

Next, we turn towards the number of days the sender is in the office. Unlike our measure of co-location, we generally find a strong positive interaction effect across our specifications. For example, an additional day in the office is associated with an additional 0.038 emails sent. We find no heterogeneity between men and women or between the lockdown and post-lockdown periods. However, we find a strong positive interaction effect between our non-manager indicator and days the sender was in the office. In short, these results suggest that the number of days the sender is in the office matters much more than the number of days a dyad co-locates.

Table A.2: The Effects of Co-location and Days Sender in the Office on Email Communication

	Number of Emails Sent									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
# Days Jointly in Office	-.011	-.039	-.162**	.006	.063*					
	[.031]	[.049]	[.082]	[.030]	[.038]					
× Male		.057								
		[.063]								
× Non-Manager			.161*							
			[.088]							
× Post-Lockdown				-.017						
				[.018]						
× Emailed Pre-Lockdown					-.173***					
					[.057]					
# Days Sender in Office						.038**	.033	-.067*	.057***	.009
						[.016]	[.023]	[.035]	[.017]	[.017]
× Male							.009			
							[.028]			
× Non-Manager								.123***		
								[.041]		
× Post-Lockdown									-.005	
									[.010]	
× Emailed Pre-Lockdown										-.036
										[.025]
Male	.492***	.283	.490***	.627***	-.140	.428***	.332	.405***	.569***	-.104
	[.145]	[.285]	[.145]	[.147]	[.145]	[.146]	[.328]	[.144]	[.150]	[.151]
Non-Manager	-1.061***	-1.081***	-1.548***	-1.017***	-.344**	-1.246***	-1.259***	-2.263***	-1.273***	-.324**
	[.181]	[.183]	[.359]	[.179]	[.152]	[.195]	[.193]	[.440]	[.195]	[.162]
Masters/PhD	-.098	-.086	-.107	.100	-.652***	-.010	-.005	-.032	.221	-.673***
	[.211]	[.211]	[.210]	[.193]	[.191]	[.219]	[.219]	[.211]	[.203]	[.208]
Married	-.820**	-.836**	-.828**	-.659**	-.260	-.791**	-.801**	-.780**	-.611*	-.285
	[.331]	[.330]	[.335]	[.306]	[.227]	[.339]	[.340]	[.334]	[.313]	[.236]
Spouse WFH	-.060	-.048	-.026	-.032	.091	.018	.023	.157	.085	.069
	[.144]	[.144]	[.147]	[.149]	[.150]	[.150]	[.149]	[.161]	[.158]	[.150]
Caring for Child	.530***	.530***	.520***	.497***	.216	.525***	.524***	.475***	.466***	.233
	[.141]	[.141]	[.141]	[.140]	[.146]	[.144]	[.145]	[.141]	[.142]	[.148]
Post-Lockdown				.045					.055	
				[.085]					[.125]	
Emailed Pre-Lockdown					4.678***					4.443***
					[.396]					[.430]
Sample Size	10176	10176	10176	20352	10176	10176	10176	10176	20352	10176

Notes.—Source: Authors. The table reports the coefficients associated with negative binomial regressions of the number of emails sent during the treatment period on an indicator for being male, a non-manager (employee), the post-lockdown period, whether the (i, j) dyad exchanged at least one email prior to the treatment, the number of days dyad (i, j) is jointly in the office, the number of days the sender is in the office, and their interactions. We also control for whether an individual has a masters/PhD (normalized to having a bachelor's), married, the spouse works from home, and the employee has to care for a child. Standard errors are clustered at the dyad-level.

A.3.3 Sentiment Analysis

Additionally, we study how extent of hybrid work relates to characteristics of emails sent by workers, notably the length and sentiment of email text. Table A.3 documents these results. These results show that employees spending significant time at the office are more parsimonious in their emails, with an additive effect when co-located in the office with a recipient.

We assess sentiment using a simple algorithm from a widely-used open-source software package (Hutto and Gilbert, 2014). After preprocessing the email bodies to remove extraneous elements, such as email signatures, we assigns proportions of positive, negative, and neutral sentiment to each email body. This analysis finds that in-person attendance is strongly associated with an increase in the average sentiment of emails, with positive sentiment increasing significantly and neutral sentiment decreasing somewhat; there is little effect on negative sentiment. Though we do not explore the mechanisms by which co-location affects sentiment, such a reading is consistent with co-location engendering more collegiality.

Table A.3: Relationship Between Email Metrics and Remote Work

	Word Count	Positive	Negative Sentiment	Neutral
	(1)	(2)	(3)	(4)
Days Sender in Office	-2.829^{**}	0.003	-0.0001	-0.003
	(1.347)	(0.003)	(0.0002)	(0.003)
Non-Manager	6.466	-0.091^{***}	-0.0002	0.092^{***}
	(13.424)	(0.035)	(0.002)	(0.035)
Post-Lockdown	3.340	0.007	-0.001	-0.005
	(7.396)	(0.021)	(0.001)	(0.021)
Gender FE	Yes	Yes	Yes	Yes
R^2	0.050	0.039	0.035	0.039
Adjusted R^2	0.032	0.020	0.016	0.020

Notes: Notes.—Source: Authors. The table reports the coefficients associated with regressions of word count and sentiment metrics against the number of days the sender is in the office. Standard errors are clustered at the sender-level.

A.3.4 Robustness Over the Intermediate Work Cutoff Threshold

Our baseline results partition employees into one of three bins based on the overall distribution of days in the office—that is, drawing from the distributions for both the microfinance and human resources departments. However, since our email data consists only of messages sent by HR employees, we replicate our main results from Table 1 in an additional series of results below, Table A.4. In reality, there is no statistically significant differenc. The intermediate cutoff changes to 9-13 days, rather than 9-14 days, which is not surprising given the overlap in the distribution of days in the office we presented earlier in Figure A.3.

Next, Table A.5 presents the main results (using the original cutoffs from the full sample) using two-way clustering from Kleinbaum et al. (2013). The standard errors increase, but the main effects remain statistically significant.

Next, Table A.6 replicates the results from the main text by altering the cutoff on the remote work classifications using only the employees in human resources. Again, there is no statistically difference between these two sets of results.

Next, we allow for even greater heterogeneity in the intensity of WFH by classifying employees into four bins: 0-7 days in the office, 8-12 days, 13-15 days, and 16-23 days. Our goal here is to allow for greater heterogeneity. We also estimate our regressions separately by group to ease the interpretation of marginal effects, rather than having to present many interaction effects.

Table A.7 documents these results. We see, in the pooled sample, that the results are clustered among those who come into the office 13-15 days: they send 0.914 more emails for a given dyad.

Table A.4: Intensity of Working-from-Home and Number of Emails Sent (Alternative Cutoff)

	Number of Emails Sent					
	(1)	(2)	(3)	(4)	(5)	(6)
8-11 Days in Office	.193	.113	.105	.081	.079	.084
	[.197]	[.196]	[.200]	[.198]	[.197]	[.200]
12-14 Days in Office	1.046***	1.002***	.985***	.929***	.937***	.887***
	[.208]	[.212]	[.218]	[.198]	[.201]	[.200]
15-23 Days in Office	.562***	.490***	.466***	.423**	.438**	.406**
	[.160]	[.161]	[.171]	[.171]	[.178]	[.177]
Non-Manager	-1.611***	-1.534***	-1.533***	-1.525***	-1.535***	-1.414***
	[.224]	[.219]	[.221]	[.218]	[.218]	[.212]
Male		.252*	.258*	.302**	.300**	.364***
		[.135]	[.138]	[.137]	[.136]	[.141]
Masters/PhD			-.091	-.009	-.005	-.047
			[.170]	[.195]	[.196]	[.199]
Married				-.249	-.260	-.455
				[.263]	[.268]	[.283]
Spouse WFH					.038	.036
					[.141]	[.140]
Caring for Child						.398***
						[.139]
Sample Size	10600	10600	10600	10600	10600	10600

Notes.—Source: Authors. The table reports the coefficients associated with negative binomial regressions of the number of emails sent during the treatment period on an indicator for whether the email sender came in between 8-11, 12-14, and 15-23 days in the office, normalized to 1-7 days in the office, controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. Here, cutoff for the number of days in the office is generated based off the employees in the human resources department. Standard errors are clustered at the dyad-level.

Table A.5: Intensity of Working-from-Home and Number of Emails Sent (Two-way Clustering)

	Number of Emails Sent					
	(1)	(2)	(3)	(4)	(5)	(6)
Intermediate WFH	.814**	.781**	.758**	.716**	.710**	.689**
	[.337]	[.340]	[.340]	[.322]	[.318]	[.313]
Low WFH	.537	.493	.457	.421	.379	.364
	[.364]	[.371]	[.360]	[.355]	[.345]	[.318]
Non-Manager	-1.608***	-1.558***	-1.555***	-1.563***	-1.538***	-1.419***
	[.454]	[.435]	[.439]	[.434]	[.408]	[.370]
Male		.185	.192	.259	.260	.334
		[.228]	[.224]	[.240]	[.238]	[.230]
Masters/PhD			-.139	-.003	-.018	-.062
			[.294]	[.266]	[.265]	[.266]
Married				-.384	-.337	-.554
				[.404]	[.420]	[.419]
Spouse WFH					-.127	-.111
					[.206]	[.193]
Caring for Child						.447*
						[.241]
Sample Size	10600	10600	10600	10600	10600	10600

Notes.—Source: Authors. The table reports the coefficients associated with negative binomial regressions of the number of emails sent during the treatment period on an indicator for whether the email sender came in between 9-14 days in the office (intermediate WFH), an indicator for 15-23 days in the office (low WFH), controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. Standard errors are clustered at both the sender and receiver levels following Kleinbaum et al. (2013).

Table A.6: Intensity of Working-from-Home and Number of Unique Recipients (Alternative Cutoff)

	log(Unique Recipients of Emails)					
	(1)	(2)	(3)	(4)	(5)	(6)
8-11 Days in Office	.210	.194	.152	.173	.160	.149
	[.191]	[.191]	[.193]	[.194]	[.192]	[.199]
12-14 Days in Office	.605***	.590***	.540***	.615***	.629***	.611***
	[.183]	[.190]	[.192]	[.181]	[.183]	[.186]
15-23 Days in Office	.332	.300	.236	.276	.328	.311
	[.214]	[.228]	[.223]	[.225]	[.219]	[.221]
Non-Manager	-.603**	-.577**	-.543**	-.547**	-.559**	-.502*
	[.270]	[.263]	[.267]	[.268]	[.261]	[.259]
Male		.103	.136	.091	.087	.128
		[.150]	[.148]	[.145]	[.147]	[.150]
Masters/PhD			-.352**	-.412**	-.409**	-.414**
			[.154]	[.173]	[.176]	[.174]
Married				.305	.242	.157
				[.213]	[.218]	[.224]
Spouse WFH					.193	.185
					[.159]	[.161]
Caring for Child						.174
						[.150]
Sample Size	99	99	99	99	99	99

Notes.—Source: Authors. The table reports the coefficients associated with regressions of the logged number of unique recipients during the treatment period on an indicator for whether the email sender came in between 8-11, 12-14, and 15-23 days in the office, normalized to 1-7 days in the office, controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. Standard errors are clustered at the employee-level.

The other two indicators are both positive, but not statistically significant, relative to the omitted group of high WFH (0-7 days in the office). Next, we see that the marginal effect is greater for men than for women on the 13-15 days in the office indicator, but a high number of days in the office is negatively associated with emails sent by women. Next, we see that our effects are roughly as large for employees with a spouse who works from home, relative to their counterparts. However, we do find important heterogeneity when we split on caring for a child. For them, the 13-15 days category is strongly associated with more sent emails. Our other effects are not statistically significant at this level of granularity.

Table A.7: Intensity of Working-from-Home and Emails Sent (Four-Bins)

	Number of Emails Sent						
	All	Male	Female	Spouse WFH	No Spouse WFH	Care for Child	No Care for Child
8-12 Days in Office	.259	.575*	-.022	.193	.233	-.625*	.577**
	[.206]	[.306]	[.243]	[.262]	[.314]	[.358]	[.258]
13-15 Days in Office	.914***	1.391***	.758***	.942***	1.069***	1.275***	.761***
	[.189]	[.378]	[.209]	[.322]	[.312]	[.365]	[.221]
16-23 Days in Office	.107	.606*	-.752*	.370	.229	.105	.005
	[.224]	[.324]	[.457]	[.474]	[.354]	[.444]	[.289]
Sample Size	10176	5618	4558	3074	7102	4452	5724

Notes.—Source: Authors. The table reports the coefficients associated with Poisson regressions of the number of emails sent during the treatment period on an indicator for whether the email sender came in between 8-12 days in the office, an indicator for 13-15 days in the office, and 16-23 days in the office, normalized to 1-7 days in the office, separately by group. We also control for whether an individual is male, is an employee (versus a manager), has a masters/PhD (normalized to having a bachelor's), married, the spouse works from home, and the employee has to care for a child. Standard errors are clustered at the dyad-level.

A.3.5 Robustness Exercises Over Main Non-linearity Results

The main text presents the coefficients associated with negative binomial regressions of number of emails exchanged with dyadic data in Table 1. Now, we examine whether these results are robust to using a Poisson distribution, rather than a negative binomial distribution.

Table A.8 documents these results. Generally speaking, we see less statistically significant results. For example, while the baseline marginal effects of intermediate and low WFH were 0.814 and 0.537, here they are 0.657 and 0.393, respectively. As we add additional controls, the marginal effect on low WFH becomes less statistically significant, but remains positive. Importantly, the marginal effect on intermediate WFH remains economically and statistically significant across every specification. The rationale for the decline in significance stems from the large share of zero email communication across dyads, making the negative binomial a better fit.

Table A.8: Intensity of Working-from-Home and Emails Sent (Poisson Robustness)

	Number of Emails Sent					
	(1)	(2)	(3)	(4)	(5)	(6)
Intermediate WFH	.657***	.614***	.599***	.576***	.595***	.538***
	[.152]	[.156]	[.161]	[.155]	[.155]	[.157]
Low WFH	.393***	.324**	.312**	.294*	.278*	.268*
	[.140]	[.147]	[.151]	[.151]	[.153]	[.158]
Non-Manager	-1.456***	-1.395***	-1.387***	-1.403***	-1.406***	-1.273***
	[.176]	[.174]	[.179]	[.181]	[.183]	[.182]
Male		.215	.221	.280*	.264*	.353**
		[.146]	[.152]	[.155]	[.156]	[.160]
Masters/PhD			-.075	-.009	-.044	-.051
			[.182]	[.189]	[.188]	[.187]
Married				-.331	-.241	-.497*
				[.250]	[.253]	[.264]
Spouse WFH					-.220	-.195
					[.140]	[.140]
Caring for Child						.429***
						[.141]
Sample Size	10600	10600	10600	10600	10600	10600

Notes.—Source: Authors. The table reports the coefficients associated with Poisson regressions of the number of emails sent during the treatment period on an indicator for whether the email sender came in between 9-14 days in the office (intermediate WFH), an indicator for 15-23 days in the office (low WFH), controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. Standard errors are clustered at the dyad-level.

Next, we allow for heterogeneity across several different dimensions in Table A.9. Column 1 begins by replicating the main result. We subsequently interact each of our demographic characteristics with our two WFH indicators, as well as an indicator for whether the dyad communicated at all during the lockdown (prior to the treatment).

We document these results in Table A.9. Column 1 begins by replicating our main result. Column 2 shows that there is not much heterogeneity between male versus female employees; the interactions are statistically insignificant. Column 3 shows that there is significant heterogeneity among employees and non-managers: all of the effect of intermediate WFH on emails is driven by employees, although there is not much heterogeneity on low WFH. Turning towards heterogeneity in educational attainment, column 4 shows that those with a master's or PhD degree are less likely to email, which may reflect their focus on more cognitively-demanding tasks (which may require less coordination and involve more independent work). Columns 5, 6, and 7 show that there is little heterogeneity across marital status, whether the spouse also works from home, and whether the respondent has to care for a child. If anything, the spouse WFH exhibits complementarity for those in intermediate WFH arrangements and substitutability for those in low WFH arrangements, whereas caring for a child interacts negatively for both intermediate and low WFH. Furthermore, those who are married send fewer emails when they are in low WFH arrangements.

Finally, column 8 allows for an interaction with whether the dyad communicated during the lockdown prior to the treatment. Perhaps most importantly, we find a negative interaction effect on the interaction effect for both intermediate and low WFH arrangements, meaning that intermediate and low levels of WFH helps democratize the workplace and break temporally static patterns in workplace communication (Bramoulle et al., 2012; Golub and Jackson, 2012). Given that one of the major concerns about WFH is that it will cause more siloes to emerge relative to workplace communication (Yang et al., 2021), these results suggest that intermediate levels of hybrid work is a potential panacea because it balances between the two extremes.

Table A.9: Heterogeneity in the Intensity of Working-from-Home and Emails Sent

	Number of Emails Sent							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intermediate WFH	.689***	.671***	-.095	1.571***	1.295**	.656***	.881***	.863***
	[.171]	[.242]	[.357]	[.419]	[.568]	[.219]	[.240]	[.190]
× Male		.044						
		[.295]						
× Non-Manager			.838**					
			[.393]					
× Masters/PhD				-.929**				
				[.425]				
× Married					-.644			
					[.562]			
× Spouse WFH						.147		
						[.310]		
× Care for Child							-.413	
							[.296]	
× Emailed Pre-Lockdown								-.554**
								[.235]
Low WFH	.364**	.583**	.326	1.635***	1.429***	.424**	.512**	.332
	[.161]	[.250]	[.406]	[.355]	[.443]	[.192]	[.212]	[.202]
× Male		-.317						
		[.302]						
× Non-Manager			.065					
			[.456]					
× Masters/PhD				-1.447***				
				[.388]				
× Married					-1.214***			
					[.462]			
× Spouse WFH						-.556		
						[.403]		
× Care for Child							-.351	
							[.309]	
× Emailed Pre-Lockdown								-.510**
								[.251]
Male	.334**	.366*	.338**	.319**	.357**	.343**	.326**	-.438***
	[.137]	[.195]	[.135]	[.137]	[.141]	[.139]	[.136]	[.149]
Non-Manager	-1.419***	-1.417***	-1.532***	-1.444***	-1.411***	-1.483***	-1.358***	-.774***
	[.205]	[.201]	[.240]	[.199]	[.201]	[.221]	[.200]	[.167]
Masters/PhD	-.062	-.033	-.103	.873**	-.049	-.018	-.086	-.342**
	[.208]	[.202]	[.207]	[.359]	[.205]	[.211]	[.205]	[.151]
Married	-.554*	-.519*	-.532*	-.547*	.091	-.590**	-.450	-.196
	[.289]	[.300]	[.287]	[.286]	[.444]	[.294]	[.287]	[.193]
Spouse WFH	-.111	-.116	-.094	-.135	-.135	-.129	-.166	.136
	[.138]	[.138]	[.144]	[.139]	[.139]	[.199]	[.145]	[.140]
Caring for Child	.447***	.451***	.461***	.412***	.478***	.446***	.673***	.205
	[.135]	[.135]	[.137]	[.135]	[.135]	[.141]	[.203]	[.132]
Emailed During Lockdown								4.705***
								[.333]
Sample Size	10600	10600	10600	10600	10600	10600	10600	10600

Notes.—Source: Authors. The table reports the coefficients associated with negative binomial regressions of the number of emails sent during the treatment period on an indicator for whether the email sender came in between 9-14 days in the office (intermediate WFH), an indicator for 15-23 days in the office (low WFH), an indicator for being male, a non-manager (employee), the post-lockdown period, whether the (i, j) dyad exchanged at least one email prior to the treatment during the lockdown, and their interactions. We also control for whether an individual has a masters/PhD (normalized to having a bachelor's), whether married, whether the spouse works from home, and whether the employee has to care for a child. Standard errors are clustered at the dyad-level.

Additionally, we examine whether our results are robust to conducting regressions at the employee level, rather than dyad level presented in the main text. We did so at the dyad level for two reasons. First, it provides us with additional variation since we can exploit differences in connectivity for the same sender towards many recipients. Second, it allows us to exploit pre-pandemic connectivity to examine the role of homophily. Nonetheless, other results (especially performance ratings) are based on employee-level data, so we want to ensure that our results are robust to this alternative specification.

Table A.10 documents these results. Starting with column 1, we see that intermediate WFH leads to 44.5% higher emails sent, whereas low WFH leads to 8.6% fewer emails sent, although it is not statistically significant at the 10% level. Next we begin partitioning based on different dimensions of demographic heterogeneity. We find that intermediate WFH is most associated with increases in emails sent, although the marginal effects are not always statistically significant because of the small sample size. Men tend to have larger marginal effects for intermediate WFH, but women have larger marginal effects for low WFH.

Finally, we ask whether it matters if a person works from home on a particular day over another—for example, earlier in the week versus later in the week. Here, we exploit our daily variation, regressing the log number of emails and words per email on an indicator for whether the employee is WFH, whether it is a Wednesday or Thursday (the end of the Bangladesh work week), their interaction, and day-of-month and person fixed effects. While columns 1 and 4 report the raw correlations for completeness, columns 2 and 5 include the fixed effects. We find no statistically significant interaction effect between the latter part of the week and WFH. For completeness, we also omit the interaction effect in columns 3 and 6, showing that employees send fewer emails on days that they are working from home. We conduct all our analysis using least squares regressions,

Table A.10: Intensity of Working-from-Home and Logged Emails Sent (Employee-level)

	log(Number of Emails Sent)						
	All	Male	Female	Spouse WFH	No Spouse WFH	Care for Child	No Care for Child
Intermediate WFH	.445**	.498**	.340	.531	.491*	.314	.579**
	[.213]	[.231]	[.397]	[.447]	[.291]	[.420]	[.239]
Low WFH	.086	-.058	.378	.205	.105	-.047	.219
	[.270]	[.352]	[.483]	[.462]	[.332]	[.560]	[.335]
Male	.518***			.542	.392	.670**	.365
	[.193]			[.454]	[.245]	[.326]	[.273]
Non-Manager	-1.226***	-1.164***	-1.369***	-1.194**	-1.312***	-1.200***	-1.204***
	[.206]	[.259]	[.437]	[.433]	[.305]	[.319]	[.284]
Masters/PhD	-.135	-.304	.144	.140	-.281	.297	-.375
	[.230]	[.246]	[.377]	[.490]	[.295]	[.352]	[.322]
Married	-.355	-.636	-.209			-.588	-.183
	[.337]	[.379]	[.461]			[.465]	[.377]
Spouse WFH	-.088	-.047	-.128			-.287	.033
	[.201]	[.236]	[.428]			[.376]	[.229]
Caring for Child	.321	.468*	.123	.112	.318		
	[.196]	[.256]	[.365]	[.479]	[.255]		
R-squared	.28	.41	.14	.28	.28	.34	.24
Sample Size	94	51	43	29	65	40	54

Notes.—Source: Authors. The table reports the coefficients associated with regressions of the logged number of emails sent during the treatment period on an indicator for whether the email sender came in between 9-14 days in the office (intermediate WFH), an indicator for 15-23 (low WFH), separately by group. We also control for whether an individual is male, is an employee (rather than manager), has a masters/PhD (normalized to having a bachelor's), married, the spouse works from home, and the employee has to care for a child. Standard errors are clustered at the dyad-level.

rather than negative binomial, because of the importance of including person and time fixed effects.

A.4 Robustness Over Productivity Results

Table A.12 begins by replicating the main results from Table 3 in the main text using only the employees in the human resources division. We see a broad similarity between the two—even slightly larger in absolute value coefficient estimates in certain cases. This shows that our results are not a remnant of composition effects and who is specifically included in our sample.

Next, Table A.13 replicates the main results, but this time using only human resources employees to determine the cutoff; each employee is assigned to one of four bins. As before, we find that intermediate WFH (here, between 12-14 days in the office) is linked with higher job satisfaction

Table A.11: Evaluating Heterogeneous Treatment Effects by Day-of-Week

	log(Number of Emails)			log(Words/Emails)		
	(1)	(2)	(3)	(4)	(5)	(6)
Working from Home (WFH)	-.128*	-.041	-.072*	.101	.021	-.013
	[.066]	[.047]	[.038]	[.068]	[.043]	[.029]
Wednesday/Thursday	-.065			.076		
	[.061]			[.058]		
WFH × Wednesday/Thursday	.075	-.073		-.075	-.081	
	[.080]	[.059]		[.072]	[.071]	
R-squared	.00	.43	.43	.00	.46	.46
Sample Size	2710	2710	2710	2632	2632	2632
Person FE	No	Yes	Yes	No	Yes	Yes
Time FE	No	Yes	Yes	No	Yes	Yes

Notes.—Source: Authors. The table reports the coefficients associated with least squares regressions of the number of emails and words in emails at a daily frequency on an indicator for working from home (WFH), an indicator for whether we see a Thursday or Friday, their interaction, conditional on person and day by month fixed effects. Standard errors are clustered at the person-level.

and better work-life balance.

We now examine how WFH intensity is related to managerial performance ratings as an alternative proxy for productivity. Table A.14 documents these results. While our estimates are noisy because of sample size considerations, we nonetheless find robust evidence that intermediate WFH is positively associated with increases in managerial ratings of employee performance: a 0.192 unit increase in ability, a 0.136 unit increase in cooperation, a 0.134 unit increase in knowledge, a 0.191 unit increase in creativity, a 0.263 unit increase in productivity, and a 0.339 unit increase in quality relative to high WFH (p-value = 0.063). Although only the last outcome variable (quality) is statistically significant at the 10% level, we nonetheless view these as useful diagnostics. The low WFH indicator is economically and statistically insignificant across all specifications.

Next, Table A.15 replicates these results using only the set of employees in the human resources department. We see similar positive effects, although they remain statistically insignificant. Fur-

Table A.12: Intensity of Working-from-Home and Employee Attitudes and Engagement

	Job Satisfaction	Better Balance	Prefer WFH
	(1)	(2)	(3)
Panel A			
Intermediate WFH	.726**	.827**	-.410
	[.354]	[.383]	[.421]
Low WFH	-.168	-.167	-.682
	[.486]	[.477]	[.484]
R-squared	.12	.10	.14
Sample Size	118	118	118
	Feel Left Out	Miss Mentorship	Feeling Isolated
	(1)	(2)	(3)
Panel B			
Intermediate WFH	.332	.062	-.631
	[.436]	[.423]	[.470]
Low WFH	.536	.669	.829
	[.546]	[.540]	[.585]
R-squared	.05	.05	.11
Sample Size	118	118	118

Notes.—Source: Authors. The table reports the coefficients associated with regressions of indices of employee preferences (ranging from one to seven) on an indicator for whether the email sender came in between 9-14 days in the office (intermediate WFH), an indicator for 15-23 days in the office (low WFH), controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. In Panel A, we draw on indices from Raghuram et al. (2001): "Overall, I am satisfied with working from home," "Since I started working from home, I have been able to balance my job and personal life," and "If I were now given the choice to return to a traditional office environment (i.e., no longer telework). In Panel B, we draw on indices from Golden et al. (2008b): "I feel left out on activities and meetings that could enhance my career," "I miss out on opportunities to be mentored," and "I feel isolated." Standard errors are heteroskedasticity-robust.

Table A.13: Intensity of Working-from-Home and Employee Attitudes and Engagement (Alternative Cutoffs)

	Job Satisfaction	Better Balance	Prefer WFH
	(1)	(2)	(3)
8-11 Days in Office	.390	.464	-.403
	[.434]	[.476]	[.539]
12-14 Days in Office	.717*	.847*	.117
	[.416]	[.470]	[.513]
15-23 Days in Office	-.167	-.062	-.488
	[.541]	[.518]	[.551]
R-squared	.11	.08	.14
Sample Size	117	117	117
	Feel Left Out	Miss Mentorship	Feeling Isolated
	(1)	(2)	(3)
8-11 Days in Office	.444	.395	-.232
	[.559]	[.562]	[.608]
12-14 Days in Office	.050	.301	-.501
	[.485]	[.489]	[.550]
15-23 Days in Office	.321	.700	.856
	[.576]	[.581]	[.637]
R-squared	.04	.05	.10
Sample Size	117	117	117

Notes.—Source: Authors. The table reports the coefficients associated with regressions of indices of employee preferences (ranging from one to seven) on an indicator for whether the email sender came in between 8-11, 12-14, and 15-23 days in the office, normalized to 1-7 days in the office, controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. In Panel A, we draw on indices from Raghuram et al. (2001): "Overall, I am satisfied with working from home," "Since I started working from home, I have been able to balance my job and personal life," and "If I were now given the choice to return to a traditional office environment (i.e., no longer telework). In Panel B, we draw on indices from Golden et al. (2008b): "I feel left out on activities and meetings that could enhance my career," "I miss out on opportunities to be mentored," and "I feel isolated." Standard errors are heteroskedasticity-robust.

Table A.14: Intensity of Working-from-Home and Managerial Performance Ratings

	Ability	Cooperation	Knowledge	Creativity	Productivity	Quality
	(1)	(2)	(3)	(4)	(5)	(6)
Intermediate WFH	.192	.136	.134	.191	.263	.339*
	[.158]	[.223]	[.188]	[.154]	[.187]	[.181]
Low WFH	.070	.019	.083	.066	-.060	.117
	[.171]	[.204]	[.194]	[.194]	[.191]	[.185]
Non-manager	-.103	-.439**	-.087	.027	-.312	.047
	[.196]	[.197]	[.189]	[.157]	[.203]	[.214]
Male	.160	.011	.242	.349**	.256	.087
	[.134]	[.172]	[.161]	[.154]	[.166]	[.163]
Masters/PhD	-.064	-.130	.336	-.282	.044	.104
	[.200]	[.274]	[.222]	[.226]	[.222]	[.201]
Married	.199	.291	.638**	-.234	-.233	-.058
	[.235]	[.257]	[.250]	[.218]	[.241]	[.228]
Spouse WFH	-.123	-.112	.008	.022	-.038	.172
	[.136]	[.189]	[.161]	[.147]	[.163]	[.153]
Cares for Child	.102	.078	-.036	.217	.164	.079
	[.132]	[.181]	[.160]	[.161]	[.166]	[.158]
R-squared	.05	.08	.12	.10	.08	.05
Sample Size	143	143	143	143	143	143

Notes.—Source: Authors. The table reports the coefficients associated with regressions of managerial productivity ratings (ranging from one to seven) on an indicator for whether the email sender came in between 9-14 days in the office (intermediate WFH), an indicator for 15-23 days in the office (low WFH), controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. Standard errors are heteroskedasticity-robust.

thermore, Table A.16 replicates these results using the four-bin cutoff for human resources employees. Here, we generally see larger positive effects for those who come in 12-14 days. For example, the gradients when the outcome variables are scores for creativity or productivity are especially large and greater than the other indicators for 8-11 or 15-23 days in the office, relative to those who come in 1-7 days.

Table A.15: Intensity of Working-from-Home and Managerial Performance Ratings (Human Resources Only)

	Ability	Cooperation	Knowledge	Creativity	Productivity	Quality
	(1)	(2)	(3)	(4)	(5)	(6)
Intermediate WFH	.099	.274	.132	.104	.232	.291
	[.165]	[.238]	[.202]	[.169]	[.204]	[.184]
Low WFH	.186	.125	.124	-.058	-.010	.216
	[.217]	[.253]	[.251]	[.255]	[.245]	[.241]
Non-manager	-.006	-.630***	-.133	.020	-.348	.039
	[.221]	[.223]	[.231]	[.200]	[.252]	[.269]
Male	.205	-.010	.207	.400**	.326*	.151
	[.144]	[.188]	[.188]	[.183]	[.194]	[.182]
mastersPhD	-.184	-.210	.225	-.452	.015	.011
	[.219]	[.310]	[.286]	[.276]	[.266]	[.251]
Married	.289	.134	.804***	-.179	-.334	.032
	[.262]	[.292]	[.290]	[.257]	[.275]	[.254]
Spouse WFH	-.010	.078	.058	.019	.069	.226
	[.145]	[.196]	[.174]	[.158]	[.178]	[.171]
Cares for Child	.166	.258	-.063	.189	.299	.156
	[.136]	[.191]	[.176]	[.189]	[.195]	[.181]
R-squared	.07	.10	.15	.09	.09	.06
Sample Size	117	117	117	117	117	117

Notes.—Source: Authors. The table reports the coefficients associated with regressions of managerial productivity ratings (ranging from one to seven) on an indicator for whether the email sender came in between 9-14 days in the office (intermediate WFH), an indicator for 15-23 days in the office (low WFH), controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. The sample is restricted to employees in the human resources department. Standard errors are heteroskedasticity-robust.

A.5 External Validity Using Gallup's COVID Panel

Table A.16: Intensity of Working-from-Home and Managerial Performance Ratings (Alternative Cutoffs)

	Ability	Cooperation	Knowledge	Creativity	Productivity
	(1)	(2)	(3)	(4)	(5)
8-11 Days in Office	.063	.199	.182	.140	.237
	[.205]	[.256]	[.238]	[.213]	[.227]
12-14 Days in Office	.228	.305	.230	.261	.404
	[.192]	[.296]	[.242]	[.182]	[.249]
15-23 Days in Office	.242	.145	.194	.030	.090
	[.234]	[.270]	[.269]	[.256]	[.271]
R-squared	.08	.10	.15	.10	.11
Sample Size	117	117	117	117	117

Notes.—Source: Authors. The table reports the coefficients associated with regressions of managerial productivity ratings (ranging from one to seven) on an indicator for whether the email sender came in between 8-11 days in the office, 12-14, and 15-23, normalized to 1-7 days in the office, controlling for the following demographic characteristics: male, a non-manager (employee), education (masters/PhD - normalized to having a bachelor's), married, spouse works from home, and employee has to care for a child. The sample is restricted to employees in the human resources department. Standard errors are heteroskedasticity-robust.

A.5.1 Data and Measurement Strategy

We draw on individual survey data that come from Gallup's COVID Tracking Survey, which began on March 13, 2020 and collected roughly 1000 responses per day through April 26th, when the sample declined to roughly 500 responses per day. We consider responses through June 2021, although the survey continued. The sample contains 148,422 responses from 52,459 unique individuals who completed the survey online (by computer or by smartphone) after receiving an emailed invitation. However, certain questions, particularly those on job satisfaction and managerial quality, are only present for a subset of observations. For much of our analysis, we focus on a sample of 16,357 respondents, consisting of 12,166 unique individuals, surveyed up to four times (median = 1 and 75th percentile = 2). We have roughly 1,500 responses per month from September 2020 until May 2021.

Our primary measure of remote work comes from the following question: "Still thinking about your job, to what extent are you taking the following steps to avoid catching or spreading the coronavirus?” This question has three sub-components: (A) Working remotely, (B) Using personal protective equipment at workplace (including masks, gloves, or face shields), (C) Trying to maintain at least 6 feet of distance between myself and other workers or customers, and (D) Adopting new or more frequent cleaning practices (including wiping surfaces, frequent handwashing, using disinfectant). We focus on sub-component (A) on remote work, which can take three possible values: always, sometimes, or never.

We also observe three other questions relating to work, including:

- Approximately how many hours, in the past 7 days, did you spend working remotely or in a location different from your co-workers?
- Approximately how many hours, in the past 7 days, did you spend working remotely or in a location different from your co-workers?
- Approximately how many hours, in the past 7 days, did you spend working from home?

The answers to these three questions allow us to compute an additional measure of remote work to gauge the extent of hybrid work, which we use as well. However, there is slightly more measurement error and 4,000 fewer responses to it, relative to the other question about precautions relating to the virus, so we opt for the former as our baseline approach.

There are a handful of measures that proxy for the employee-manager relationship, including the following that range between integer values of one (strongly disagree) and five (strongly agree):

- How satisfied are you with your current place of employment as a place to work.?
- I know what is expected of me at work.

- I have the equipment and materials I need to do my work right.
- I have an opportunity to do what I do best everyday.
- There is someone at work who encourages my development.
- At work, my opinions seem to count.
- The mission or purpose of my organization makes me feel my job is important.
- My organization gives all employees equal opportunity to advance to senior management.
- My manager trusts me.
- I feel prepared to do my job.

Moreover, we observe a measure of managerial feedback from “How often do you receive feedback from your manager?” with the following possible answers: daily, a few times a week, a few times a month, a few times a year, once a year, less often than once a year, never.

Next, thinking about everything you’ve done in the past 24 hours, which of the following comes closest to describing your in-person contact with people outside your household?

1. Completely isolated yourself, having no contact with people outside your household
2. Mostly isolated yourself, having very little contact with people outside your household
3. Partially isolated yourself, having some contact with people outside your household
4. Isolated yourself a little, still having a fair amount of contact with people outside your household

5. Did not make any attempt to isolate yourself from people outside your household

How concerned are you about being exposed to coronavirus at your place of work? (A) Not concerned at all, (B) Not too concerned, (C) Moderately concerned, and (D) Very concerned.

How worried are you that you or someone in your family will be exposed to coronavirus? (A) Not worried at all, (B) Not too worried, (C) Somewhat worried, and (D) Very worried.

A.5.2 Empirical Strategy and Results

Using these data on job satisfaction and preparedness, we consider regressions of the form:

$$y_{it} = \gamma m_{it} + \zeta_1 WFH^S_{it} + \zeta_2 WFH^A_{it} + \xi_1(m_{it} \times WFH^S_{it}) + \xi_2(m_{it} \times WFH^A_{it}) + \beta X_{it} + \phi_o + \lambda_t \quad (3)$$

where y_{it} denotes a standardized z-score of job satisfaction or job preparedness for individual i in year-month t, m denotes a standardized z-score of managerial or workplace practices, WFH^S and WFH^A denote indicators for "sometimes remote" and "always remote" (normalized to working from the office), X denotes a vector of individual covariates (including income fixed effects), and ϕ_o denotes fixed effects on occupation o. Standard errors are heteroskedasticity-robust.

We recognize that our coefficients of interest, primarily ξ^S and ξ^A, are not causally identified—we only have cross-sectional variation at our disposal. However, we control for a wide array of potentially confounding factors, such as income and occupation fixed effects. These proxy for unobserved heterogeneity that prompt more productive individuals to sort into different work arrangements. Nonetheless, these estimates are instructive and suggestive of external validity.

Table A.17 documents the results associated with Equation 3. Across every specification, we see a robust association between workplace practices and job satisfaction and job preparedness. For example, a 1 s.d. rise in having someone encouraging at work is associated with a 0.576 s.d. and 0.481 s.d. rise in job satisfaction and preparedness (columns 1 and 7), respectively, and a comparable increase is associated with a 0.603 s.d. and 0.505 s.d. increase for a marginal change in having a sense of purpose (columns 3 and 9). Similarly, trust between the employee and manager predicts job satisfaction and preparedness (columns 6 and 12).

More interestingly, however, we find that working remote always is most associated with both job satisfaction and preparedness, whereas working remote only sometimes is statistically associated with job satisfaction. This reflects the greater autonomy present in full remote work arrangements. While there is greater autonomy, fully remote work comes at the cost of less social interaction and learning in the workplace. We now focus on the interaction effects between our measures of the workplace and remote work.

Starting with changes in the degree of mentorship at work, we see statistically significant negative interaction effects—for example, a 1 s.d. rise in mentorship for a fully remote job is associated with a 0.08 s.d. rise in job satisfaction and a 0.09 s.d. increase in job preparedness, but a 0.12 s.d. and 0.22 s.d. rise for those in full in-person arrangements. We also see similar negative interaction effects on fully remote work for marginal changes in the degree in purpose.

We see no such negative interaction effects for intermediate WFH (i.e., the "sometimes remote" arrangements), meaning that marginal changes in the workplace practices have comparable effects as fully in-person work arrangements. In other words, while the relation between working remote always and job satisfaction is moderated by workplace quality, like mentorship, there is no comparable weakening in the relation between sometimes working remotely and job satisfaction.

Finally, when we focus on managerial quality and whether trust exists with the employee, we see that the benefits are concentrated among the intermediate WFH work arrangements: a 1 s.d. rise in managerial trust is associated with an additional 0.082 s.d. increase in job satisfaction for partially remote jobs (column 6). We find no additional statistically significant effect on job preparedness for these partially remote jobs, but we do find a negative effect on fully remote jobs.

Table A.17: The Mediating Effects of the Workplace with Remote Work on Job Satisfaction & Performance

	Job Satisfaction (z-score)						Job Preparedness (z-score)					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sometimes Remote	.083***	.066**	.071***	.049**	.117***	.097***	.041	.064**	.026	.041	.058**	.084***
	[.024]	[.026]	[.023]	[.024]	[.028]	[.031]	[.027]	[.028]	[.025]	[.026]	[.028]	[.031]
Always Remote	.177***	.120***	.203***	.121***	.170***	.129***	.223***	.192***	.231***	.180***	.201***	.186***
	[.023]	[.026]	[.021]	[.024]	[.028]	[.032]	[.025]	[.029]	[.023]	[.027]	[.027]	[.031]
Mentorship (z-score)	.577***	.579***					.482***	.489***				
	[.016]	[.017]					[.019]	[.020]				
× Sometimes Remote	.019	.006					-.045	-.059*				
	[.027]	[.028]					[.030]	[.031]				
× Always Remote	-.043*	-.051**					-.136***	-.143***				
	[.024]	[.025]					[.027]	[.028]				
Purpose (z-score)			.603***	.604***					.505***	.513***		
			[.014]	[.015]					[.018]	[.019]		
× Sometimes Remote			.003	.007					-.020	-.017		
			[.024]	[.025]					[.028]	[.029]		
× Always Remote			-.093***	-.090***					-.199***	-.194***		
			[.021]	[.023]					[.026]	[.026]		
Trust (z-score)					.449***	.444***					.516***	.520***
					[.018]	[.018]					[.020]	[.019]
× Sometimes Remote					.080**	.082**					.020	.000
					[.032]	[.033]					[.035]	[.035]
× Always Remote					.023	.016					-.060*	-.070**
					[.031]	[.032]					[.032]	[.033]
R-squared	.329	.340	.362	.375	.259	.270	.204	.237	.232	.269	.294	.316
Sample Size	26648	24032	27886	25094	14379	12868	25949	23410	27193	24474	14466	12932
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Income FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year and Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes.—Source: Gallup COVID Panel, March 2020 - June 2021. The table reports the coefficients associated with a standard normal index of job satisfaction and job performance (on a scale of one to five) on an indicator for whether the respondent works remotely always, whether the respondent works remotely sometimes (normalized to non-remote work), an index of managerial quality, their interactions, and a vector of demographic characteristics, including: education fixed effects (less than high school, high school, technical, some college, college, and college plus), male, race (White, Black, Asian, Hispanic), full-time employment, marital status, and whether the respondent has children. The two outcome variables are defined based on the response to the questions: "How satisfied are you with your current place of employment as a place to work" and "I feel prepared to do my job." The three workplace measures are defined based on the responses to "There is someone at work who encourages my development," "The mission or purpose of my organization makes me feel my job is important," and "My manager trusts me." Standard errors are heteroskedasticity-robust and observations are weighted by the survey weights.

How do our results with the Gallup panel compare with those obtained from our randomized controlled experiment? First, we see that hybrid and fully remote work are associated with greater job satisfaction and preparedness relative to fully in-person work arrangements. While we cannot rule out selection effects for these set of results, our results are heavily robust to the inclusion of industry and income fixed effects, together with a wide array of demographic characteristics. Second, we see an economically meaningful mediating effect of management on job satisfaction and preparedness for intermediate WFH work, but not high WFH work. Given our results from the RCT highlight the value of intermediate WFH work and employee performance as well as intrafirm communication, our empirical results here provide supplementary evidence that intermediate levels of WFH might be productivity enhancing across a variety of workplace settings.

A.6 Methodology for the Novelty Measure

We apply a new procedure to generate an approximation of novelty. We first preprocess all email text to remove email signature, inline replies, and the like. Then, we combine this text with the plaintext extracted from document attachments[16]. We then generate word vectors for each document (email text or attachment text) using Doc2vec (Le and Mikolov, 2014), a natural language processing model implementing in Gensim and extended from Word2vec. Doc2vec makes use of neural networks to produce a numeric vector representing a document (Figure A.8). Beyond a mere bag-of-words approach, this attempts to leverage semantic relationships. Thus, each document admits a fixed-length, numeric vector representation.

We then apply the BIRCH (Balanced Iterative Reducing and Clustering Using Hierarchies)

[16]Specifically, we examine Word documents, PowerPoint presentations, and PDF files.

algorithm (Zhang et al., 1997), an unsupervised learning process in the scikit-learn package, to cluster all such vectors into document types. Then, returning to the original vectors, we generate a "representative" document for each sender $\times$ document type by taking the simple average. For example, this would generate a synthetic "representative salary report sent by employee A." Next, for each document, we compute the cosine similarity against this synthetic representative document (Figure A.9). The cosine similarity measure ranges from -1 to 1 for documents orthogonal or identical, respectively, to the synthetic representative document. By halving this value and subtracting from 0.5, we produce a novelty measure between 0 and 1, with 1 indicating more novelty.[17]

To collapse these values into a tractable form, we take the mean of all novelty measures per sender during the lockdown and treatment periods (Figure A.10). Finally, we take the difference in each sender's scores from the lockdown to the treatment period and standardize this value.

To interpret these as a measure of productivity, we would need an additional identifying assumption is that producing more novel documents implies that an employee has produced more work; we do not do so here, leaving it as simply "novelty." Although the details of the calculation outlined above are involved, the interpretation of this novelty index is natural. Consider the case in which an employee simply attaches the same document to many emails. The cosine similarity of these documents will be very high, and so the novelty measure will be low. Therefore, we interpret the value of the difference index to mean that little productive work was required for this

[17]For example, the following string received a novelty score of 0.013 (low novelty): "Dear [Name] As per your requirement here I have attached the documents. Regards [Name]". This string received a higher novelty score of 0.315: "i) Staff reflection/comments on overall actual achievements made so far including challenges and areas to be improved based on objectives (MID-YEAR):... In this pandemic situation i ensured virtual communication with [#] field staffs regarding COVID awareness and positive case identify and regular HR service. It could have been better if I could give more attention to develop my team members. Rest of the year I will focus on building a productive and high performing team through proper guideline and involvement in this pandemic situation...."

document. Now, consider a case where an employee creates a salary report each week. Although the documents are broadly similar, and therefore would be classified as the same document types, we would anticipate deviations in the content of each document. Therefore, the cosine similarity index would be lower, and the novelty measure higher. One potential interpretation is that this represents more work performed by the sender.

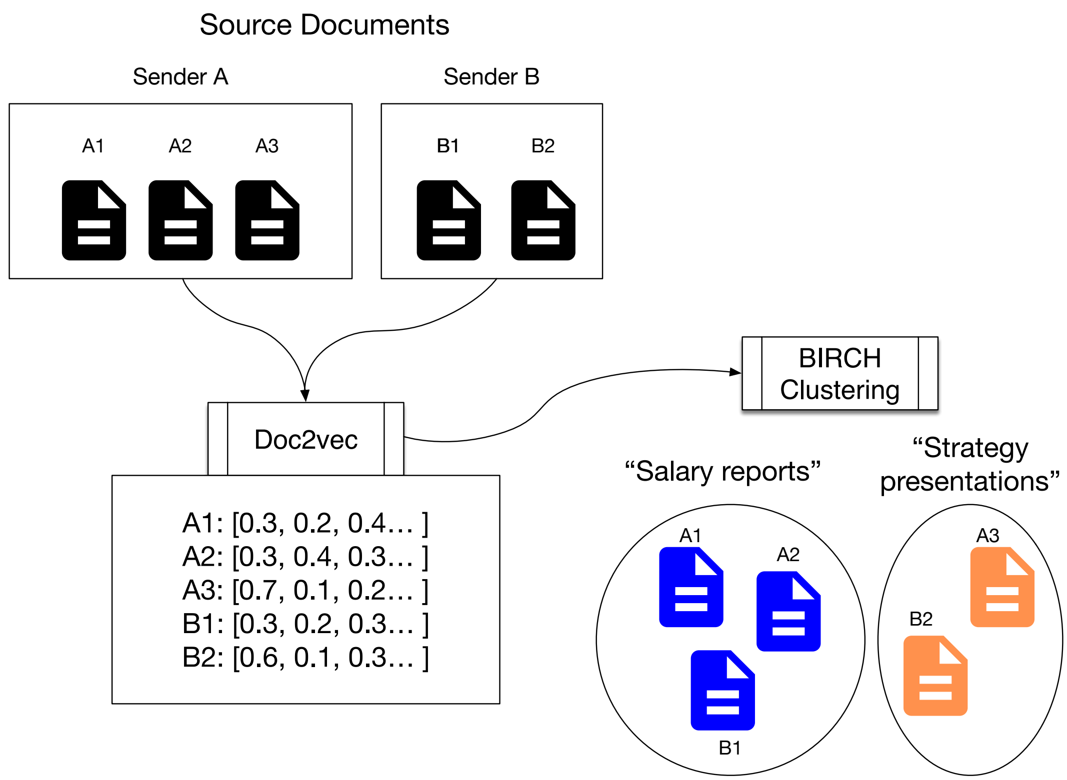

Figure A.8: Clustering Process

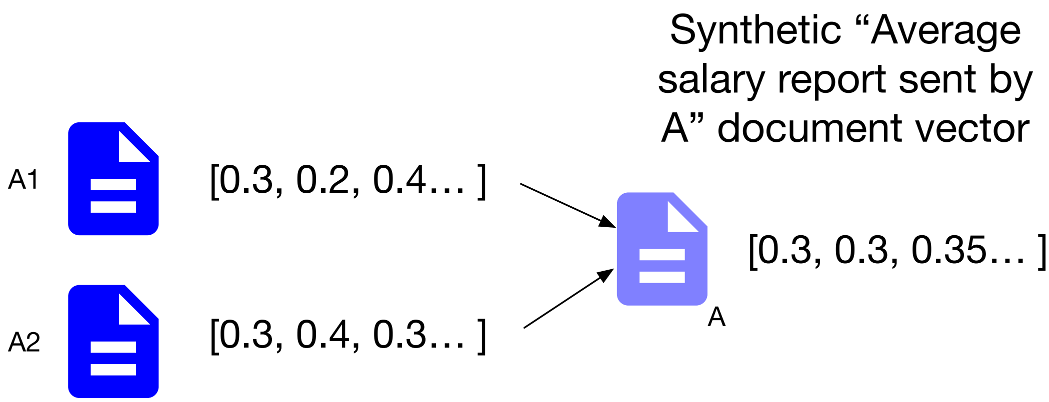

Figure A.9: The Synthetic Document

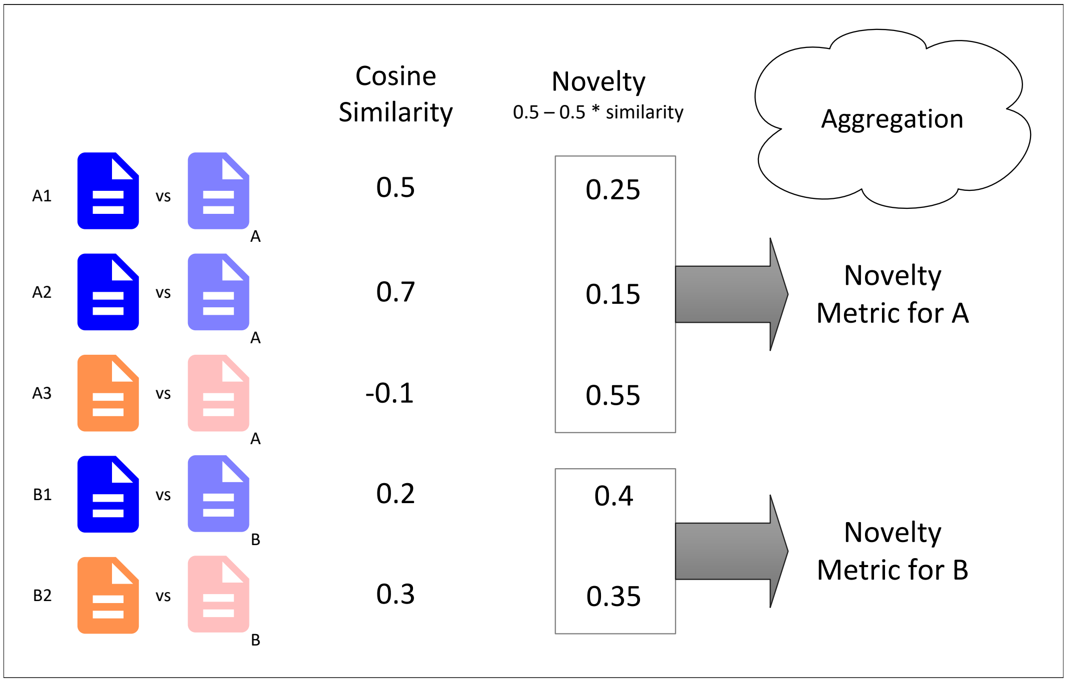

Figure A.10: Novelty Measure

A.7 Robustness of the Novelty Measure

A.7.1 Parametrization of the Clustering Model

As Doc2vec and BIRCH leverage unsupervised machine learning methods, the choice of appropriate parameters is critical to deriving robust and useful results. For example, increasing the number of epochs may improve the fit of a model while increasing computational time; similarly, modifying the learning rate can dampen the influence of outliers. Beyond these computational considerations, two parameters are particularly salient: the length of the vector Doc2vec uses to represent a document (akin to the number of principal components), and the number of clusters BIRCH returns.

Because our data are unlabeled—we do not manually tag or cluster documents, but rather use these unsupervised methods—we must rely on heuristics to determine an appropriate parametrization. One such heuristic is the "silhouette score," which ranges from -1 to 1 and is implemented in scikit-learn. A higher value indicates that 1) clusters are clearly separated from each other in the metric space and 2) clusters are relatively tight relative to the distance between clusters overall. By ranging both the vector size and number of clusters from 4 to 20, inclusive, and running the Doc2vec/BIRCH process, we observe the range of silhouette scores reported in Figure A.11. Note that increasing granularity (that is, increasing the vector size or number of clusters) is generally detrimental for this silhouette score, given our data. Although this may appear counterintuitive, increasing granularity can result in "overfitting" the data and thereby undermining the clusters.

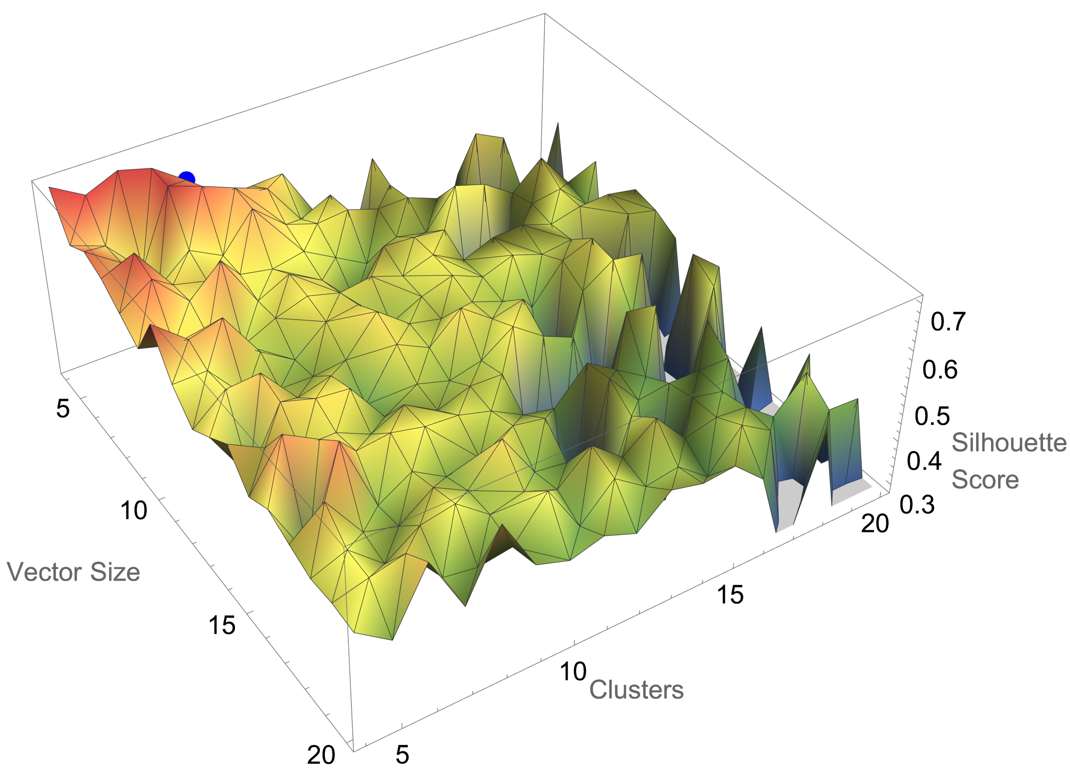

Figure A.11: Silhouette scores for various possible parameters for the vector size and number of clusters. The blue dot represents the parametrization used in Table 2.

This process results in the cluster distribution illustrated by Figures A.12 and A.13.

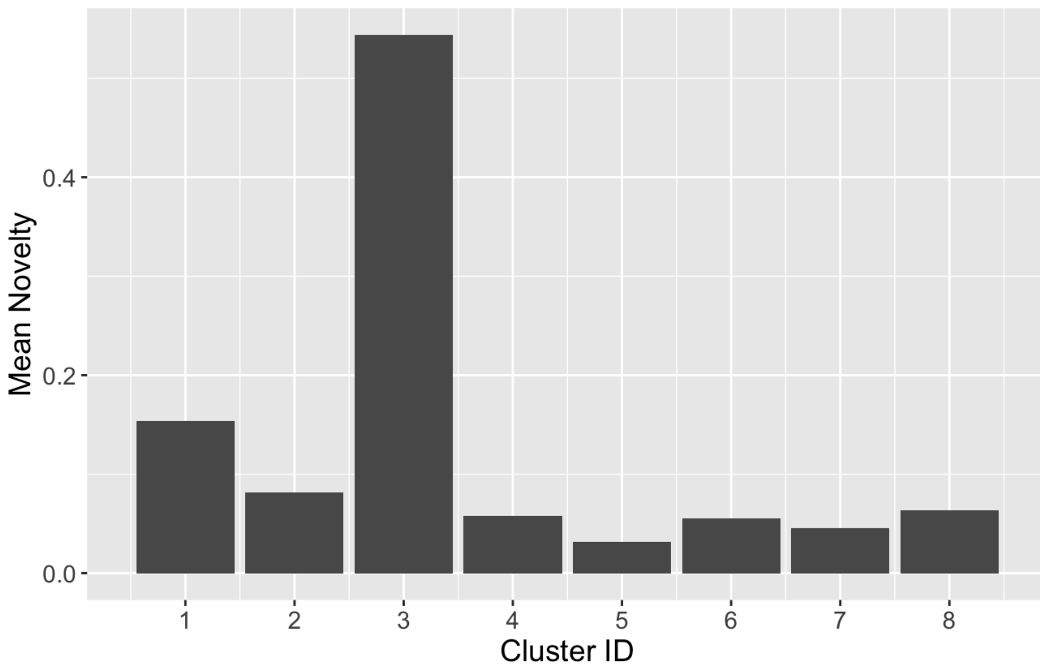

Figure A.12: Distribution of mean novelty scores per cluster (treatment period only)

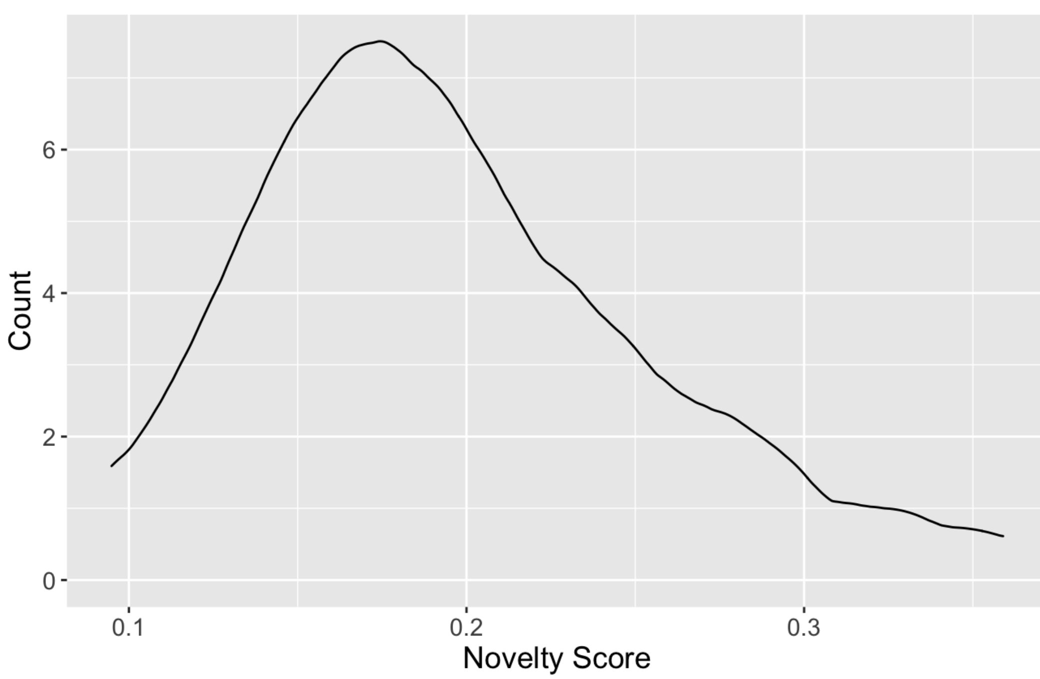

Figure A.13: Distribution of mean novelty scores, across all clusters, by sender (treatment period only)

We also ran several other variations of the model in Table A.18, all of which are broadly consistent with our reported results. Specifically, we examine using (1) only attachment text, (2) only email text, (3) K-means rather than BIRCH clustering, and (4) a sender $\times$ cluster-level synthetic document used as the centroid, rather than a common cluster-level synthetic document. (Note that, in all cases, individual clusters are calculated on a universal basis, not on a per-sender basis.) Additionally, we applied alternate bin specifications in Table A.19 and found robust results for an intermediate WFH category. We also tested the effect of email counts, as the novelty measure captures both volume and novelty; our results remain robust, although slightly less significant (Table A.20). Next, rather than examining the standardized change in novelty from the lockdown to treatment period, we look at the standardized value during the treatment period only in Table A.21. Finally, we also examine heterogeneity within specific clusters in Table A.24, as each document is tagged with a single cluster from 1 to 8. The subsample analysis is somewhat noisy as it ignores the aggregation at a per-sender level across all documents. The effect is most strongly pronounced in cluster 2 and 3, with the coefficients in other clusters statistically insignificant from zero. Assigning an intuitive meaning to these labels is inherently subjective, but a review of the data reveals that cluster 2 is primarily related to "salary adjustment reports" and cluster 3 to "policy clarifications."

Table A.18: Relationship Between Work Novelty and the Intensity of Working-from-Home, Alternate Specifications

	Change in Cosine Similarity-Based Work Novelty Measure			
	(1)	(2)	(3)	(4)
Intermediate WFH	0.619***	0.094	0.380*	0.634***
(9-14 days in office)	(0.224)	(0.223)	(0.217)	(0.225)
Low WFH	0.099	−0.302	0.130	0.303
(15+ days in office)	(0.232)	(0.241)	(0.249)	(0.551)
Gender FEs	Yes	Yes	Yes	Yes
Manager/Worker FEs	Yes	Yes	Yes	Yes
Other Demographic FEs	Yes	Yes	Yes	Yes

Notes: See Table 2 for a discussion of the novelty measure. The specific robustness checks represented above involve using (1) only attachment text, (2) only email text, (3) K-means rather than BIRCH clustering, and (4) a sender × cluster-level synthetic document used as the centroid, rather than a common cluster-level synthetic document.

Table A.19: Relationship Between Work Novelty and the Intensity of Working-from-Home, Alternative WFH Classifications

	Change in Cosine Similarity-Based Work Novelty Measure	
	(1)	(2)
8-11 Days in Office	0.345	0.334
	(0.227)	(0.227)
12-14 Days in Office	0.736***	0.673***
	(0.226)	(0.239)
15+ Days in Office	0.044	0.037
	(0.204)	(0.236)
Gender FEs	Yes	Yes
Manager/Worker FEs	Yes	Yes
Other Demographic FEs	No	Yes
N	105	99

Notes: Source: Authors. See text for discussion of the construction of the similarity measure. The dependent variable measures the change in the novelty measure from the baseline to treatment period. "Other Demographic FEs" are controls for education (bachelor's, master's/PhD), marriage status, whether the employee's spouse works from home, and whether the employee cares for a child at home. Text data was classified into eight clusters from a Doc2vec vector length of four; see Section A.7.1 of the Appendix for a discussion of the parameter fit.

Table A.20: Relationship Between Work Novelty and the Intensity of Working-from-Home, Controlling for Email Volume

	Change in Cosine Similarity-Based Work Product Index	
	(1)	(2)
Intermediate WFH	0.373^{**}	0.329^{*}
(9-14 days in office)	(0.184)	(0.200)
Low WFH	−0.114	−0.115
15+ days in office	(0.196)	(0.217)
Email Volume	0.002^{***}	0.003^{**}
	(0.001)	(0.001)
Gender FEs	Yes	Yes
Manager/Worker FEs	Yes	Yes
Other Demographic FEs	No	Yes

Notes: Source: Authors. See text for discussion of the construction of the similarity measure. The dependent variable measures the change in the novelty measure from the baseline to treatment period. "Other Demographic FEs" are controls for education (bachelor's, master's/PhD), marriage status, whether the employee's spouse works from home, and whether the employee cares for a child at home. Text data was classified into eight clusters from a Doc2vec vector length of four; see Section A.7.1 of the Appendix for a discussion of the parameter fit.

Table A.21: Relationship Between Work Novelty and the Intensity of Working-from-Home, Treatment Period Only

	Standardized Cosine Similarity-Based Work Product Index During Treatment Period	
	(1)	(2)
Intermediate WFH	0.505***	0.570**
(9-14 days in office)	(0.192)	(0.235)
Low WFH	0.152	0.269
(15+ days in office)	(0.211)	(0.237)
Gender FEs	Yes	Yes
Manager/Worker FEs	Yes	Yes
Other Demographic FEs	No	Yes

Notes: Source: Authors. The dependent variable reflects the standardized score of each sender's novelty's index with reference to the treatment period only (not a change from the control to treatment period).

Table A.22: Relationship Between Work Novelty and the Intensity of Working-from-Home

	Dep. var. = Change in Cosine Similarity-Based Work Product Index (K-Means)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intermediate WFH	0.392*	0.373*	0.351	0.352	0.347	0.347	0.380*
	(0.209)	(0.204)	(0.231)	(0.254)	(0.254)	(0.253)	(0.217)
Low WFH	0.155	0.080	0.075	0.075	0.137	0.130	0.146
	(0.214)	(0.217)	(0.242)	(0.243)	(0.245)	(0.249)	(0.481)
Non-Manager	−0.855***	−0.820***	−0.910***	−0.909***	−0.925***	−0.876***	−1.001***
	(0.309)	(0.303)	(0.296)	(0.295)	(0.283)	(0.283)	(0.342)
Male		0.210	0.201	0.200	0.203	0.246	0.284
		(0.181)	(0.183)	(0.178)	(0.178)	(0.185)	(0.233)
Masters/PhD			−0.274	−0.275	−0.278	−0.279	−0.249
			(0.375)	(0.345)	(0.340)	(0.340)	(0.322)
Married				0.005	−0.079	−0.169	−0.161
				(0.320)	(0.330)	(0.312)	(0.317)
Spouse WFH					0.212	0.197	0.248
					(0.204)	(0.210)	(0.245)
Caring for Child						0.203	0.141
						(0.199)	(0.250)
Colocation Intensity							−0.067
							(0.215)

Notes.—Source: Authors. The same notes as in Table 2 apply, except that K-Means, rather than BIRCH, was used for the clustering.

Table A.23: Relationship Between Work Novelty and the Intensity of Working-from-Home

	Dep. var. = Change in Central Cluster, Cosine Similarity-Based Work Product Index						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intermediate WFH	0.497**	0.472**	0.565**	0.563**	0.564**	0.565**	0.634***
	(0.203)	(0.197)	(0.226)	(0.248)	(0.249)	(0.248)	(0.225)
Low WFH	0.262	0.165	0.247	0.247	0.228	0.224	0.303
	(0.219)	(0.222)	(0.236)	(0.237)	(0.253)	(0.260)	(0.551)
Non-Manager	−0.722***	−0.676***	−0.791***	−0.793***	−0.788***	−0.758***	−0.787**
	(0.253)	(0.246)	(0.244)	(0.244)	(0.249)	(0.264)	(0.314)
Male		0.270	0.193	0.196	0.195	0.222	0.287
		(0.182)	(0.180)	(0.172)	(0.171)	(0.183)	(0.238)
Masters/PhD			0.035	0.039	0.040	0.040	0.091
			(0.343)	(0.312)	(0.314)	(0.315)	(0.304)
Married				−0.020	0.005	−0.049	−0.030
				(0.319)	(0.338)	(0.314)	(0.320)
Spouse WFH					−0.063	−0.073	−0.049
					(0.201)	(0.211)	(0.242)
Caring for Child						0.124	0.099
						(0.211)	(0.264)
Colocation Intensity							−0.088
							(0.233)

Notes.—Source: Authors. The same notes as in Table 2 apply, except that the synthetic document is calculated at the cluster level (i.e., "Cluster 1" uses the same synthetic document for all users).

Table A.24: Relationship Between Work Novelty and the Intensity of Working-from-Home

	Cosine Similarity-Based Work Product Index							
	Subsample Analysis (Cluster ID)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intermediate WFH	0.174	2.320***	0.659***	0.521	−0.488	0.253	0.322	0.250
	(0.158)	(0.381)	(0.221)	(0.381)	(0.433)	(0.314)	(0.399)	(0.354)
Low WFH	−0.069	4.088***	0.282	0.248	−0.350	−1.772	0.550	−0.536
	(0.360)	(0.595)	(0.532)	(0.604)	(0.476)	(1.243)	(0.858)	(0.676)
Non-Manager	−0.724**	−0.630***	−0.847***	−0.130	−0.210	−0.236	−0.702	0.068
	(0.336)	(0.222)	(0.322)	(0.256)	(0.446)	(0.344)	(0.587)	(0.360)
Male	0.194	−1.664***	0.267	0.257	−0.006	0.133	0.181	0.124
	(0.174)	(0.325)	(0.235)	(0.252)	(0.435)	(0.456)	(0.257)	(0.287)
Masters/PhD	0.221	−1.919***	0.036	−0.259	2.075	−0.708*	0.550*	0.176
	(0.156)	(0.302)	(0.305)	(0.607)	(1.534)	(0.363)	(0.328)	(0.230)
Married	−0.239	2.650***	−0.051	0.295	−2.308	0.374	−0.132	−0.132
	(0.224)	(0.539)	(0.312)	(0.459)	(1.535)	(0.464)	(0.285)	(0.312)
Spouse WFH	−0.215	0.481**	−0.064	0.729	−0.248	−0.356	−0.347	−0.532*
	(0.174)	(0.230)	(0.237)	(0.533)	(0.212)	(0.469)	(0.522)	(0.275)
Caring for Child	0.192	−2.246***	0.138	−0.126	0.212	0.580*	0.368	−0.046
	(0.191)	(0.400)	(0.261)	(0.303)	(0.381)	(0.339)	(0.519)	(0.312)
Colocation Intensity	−0.009	−1.312***	−0.083	0.056	−0.137	0.545	−0.543	−0.055
	(0.121)	(0.308)	(0.227)	(0.226)	(0.183)	(0.364)	(0.497)	(0.255)

A.7.2 Hash-Based Novelty Measure

As a robustness check, we apply a cruder novelty measure and find broadly consistent results. This measure is based on the count of novel work related email attachments sent by workers, determined by a unique hexadecimal string calculated from each attachment. Specifically, we use a hashing function—MD5 (Rivest, 1992)—to code the number of novel work-related attachments from individual emails and treat this variable as a proxy for individual productivity.

To generate this measure, we first extract each email attachment from all emails sent by the individual and classify the attachment by file types, i.e., PDF files, Excel files, Word files, etc. Next, we apply the MD5 hash function to determine whether the attachment is unique for the given employee, i.e., whether the employee sent the attachment for the first time ever. Each hash is represented as a fixed-length hexadecimal string, such as "5e8c7faa7914883775d4009c1850e67e". Changing even a single bit of data in a file will produce a different hash value, so this hashing process allows us to determine whether an attachment is truly unique. We then tag each sent email as containing a novel work product if at least one of the attachments has never been sent before by a given employee. The results are reported in Table A.25. Of particular note are regressions (2) and (4), which show evidence that this measure of novelty is highest for employees in the intermediate WFH category.[18] These results suggest that individuals with intermediate WFH send roughly 19 more novel email work products compared to those in the high WFH group.

[18]For this table, the sample size is determined by the fact that we have 107 employees for whom we have email data and two periods (lockdown, treatment). Some of the regressions have fewer observations because we don't have survey data for every person.

Table A.25: Relationship Between Work Novelty and the Intensity of Working-from-Home

	Count of Novel Email Attachments			
	(1)	(2)	(3)	(4)
Intermediate WFH		0.421**		0.278*
		(0.168)		(0.155)
Low WFH		0.287*		0.202
		(0.171)		(0.137)
Male	0.389***	0.309***	0.374***	0.312***
	(0.104)	(0.101)	(0.117)	(0.118)
Non-manager	−0.595***	−0.695***	−0.499***	−0.575***
	(0.135)	(0.129)	(0.134)	(0.145)
Bachelor's Degree	−0.157	0.047		
	(0.228)	(0.275)		
Master's/PhD	−0.185	−0.016		
	(0.169)	(0.194)		
Married	−0.326*	−0.335		
	(0.179)	(0.230)		
Spouse WFH	−0.022	−0.005		
	(0.118)	(0.114)		
Cares for Child	0.105	0.129		
	(0.106)	(0.101)		
N	202	198	214	210

Notes: Source: Authors. The parameters were estimated by a negative binomial model. An attachment is considered "novel" if it has never been sent before by a given sender. We use a hash function, which varies based on the content (rather than name) of the attachment, to determine novelty. The coefficients for the treatment period and constant are omitted from the table. Standard errors are clustered at the sender level.