

Online Appendix (For Online Publication Only)
Cullen and Perez-Truglia, “How Much Does Your Boss Make?”
September 4, 2021

A Survey Instrument

Below we provide screenshots of the translated survey instrument. The business language of the firm is English so both written and spoken English skills are extremely high, allowing us to pilot the survey with many bilingual employees. After drafting the survey in English, a team of two (employees of the bank) translated the survey so that both versions were available on Qualtrics in English and the local language. We then did all pilot testing of the survey in the local language, meaning we piloted the survey with individuals willing to take the translated version and give detailed feedback about clarity and interpretation in English to us so that we could respond directly with further questions and then instructions for translation edits.

When we elicited the probability beliefs over a series of bins around the respondent’s guess, shown in the 9th screen of the survey instrument, the bin cutoffs displayed were adaptive to the reported prior. We created 5 bins, with the middle bin centered at the prior with cutoffs $\pm 2.5\%$ around the prior; the next two bins on either side ranged from -2.5 to -10% of the prior and from $+2.5\%$ to $+10\%$ of the prior; and the most extreme bins included anything greater than $+10\%$ or less than -10% from the prior, respectively. The respondent can type probabilities for each bin, and those numbers were forced to sum up to 100% . The box called “Total” (the one that must be 100% for the response to be valid and proceed to the next screen) cannot be edited, and it is updated automatically to provide real-time feedback.

Dear colleagues,

You are invited to participate in a survey study conducted by [Researcher Names] from [Bank Name] in collaboration with a group of academic researchers from Harvard University and other universities from the United States. This survey is intended to teach us more about how [Bank Name] employees learn about their workplace, earnings and career prospects. The purpose is to find ways we can improve our communication about salaries and promotion, and to understand your beliefs about your future career with [Bank Name]. This study is aligned with one of three platforms in the five-year strategy of [Bank Name].

This survey should take less than 30 minutes to complete. All the information provided in this survey is 100% truthful. As a token of our appreciation for your participation, you will be able to earn a minimum of \$9.75 and up to \$700, based on your performance in a game included in this survey.

The rewards will be deposited in your payroll account by the end of Q2.

ALL SURVEY RESPONSES ARE COMPLETELY CONFIDENTIAL. Contact the Office of the Chief Economist should any issue arise.

Thank you in advance for your participation!

Sincerely,

Chief Economist, [Bank Name]

I confirm that I am [Respondent Name] and I would like to take part in this study

>>

To get a general picture of the people answering this survey, we need to know a few things about your background.

Where did you grow up?

Recent research on decision making shows that choices are affected by the context in which they are made. Differences in how people feel, in their previous knowledge and experience, and in their environment can influence the choices they make. To help us understand how people make decisions, we are interested in information about you, specifically whether you actually take the time to read the instructions; if you don't, some results may fail to tell us very much about decision making in the real world. To help us confirm that you have read these instructions, please select the "none of the above" option below. Thank you very much.

In this survey, you will be asked to guess the answer to some questions, and will be rewarded according to the accuracy of your answers. Take the following example:

What is the average height of women in this country (in centimeters)? [Note: we will reward you up to \$2.61 for accuracy]

cm

Note the message "we will reward you up to \$2.61 for accuracy." What we mean by that is that we will use a formula with the ACTUAL average height to reward you. The more accurate your answer is, the more money you will get, up to \$2.61.

Go ahead and provide your guess. This is a practice question, so it will not be scored.



Your guess was 150 cm. The truth is 153 cm. As a result, you would have been awarded \$2.20.

- If you had responded exactly the truth (153 cm), you would have been awarded \$2.61.
- If you had responded 10% above or below the truth (138 cm or 168 cm), then you would have been awarded \$1.05.
- If you had responded 20% above or below the truth (123 cm or 183 cm), then you would have been awarded just \$0.05.

This formula was designed by economists. According to this formula, it is in your best interest to respond honestly.

We will reward you for guessing averages, and also for making other types of guesses. Whenever you see that there is a reward for your guess, please remember that it is in your best interest to respond what you truly believe.

>>

Now, we want to ask you some questions related to salaries. In this survey, we always refer to the monthly basic salary: that is, your monthly salary WITHOUT specific allowances, WITHOUT bonus payments and WITHOUT tax & other deductions. This is the salary specified in your contract.

To make sure that you understand this definition, please try to recall your basic salary and report it here, so we can show you how your answer compares to our records. Please be as exact as possible when reporting this amount, using "." for the decimal separator.

What is your current monthly basic salary from March of 2017? [Note: we will reward you up to \$2.61 for accuracy]

There is a discrepancy between the amount that you reported and our records: you reported a monthly basic salary of \$782, while the administrative records from [Bank Name] indicate an amount of \$730.

Remember, our definition of basic salary EXCLUDES specific allowances, EXCLUDES tax & other deductions and corresponds to March of 2017.

Do you agree with the amount of \$730 shown in our records?

- Yes
- No



**For the remainder of the survey, please keep in mind that all salaries correspond to this same definition:
monthly basic salary for March 2017, WITHOUT specific allowances and WITHOUT tax & other deductions.**



Consider the other employees from the bank who work in your same position (Teller) and unit (Branch 10). According to our records, there are around 50 employees in this group.

What is the average monthly basic salary among all employees in your same position and unit as of March 2017?
[Notes: we will reward you up to \$2.61 for accuracy]

In the previous question, you reported to believe that the average monthly basic salary among employees similar to you was \$848 in March 2017. The next question is designed to assess how confident you feel about your response.

With what probability do you think that the real average could fall in each of the following bins? The probabilities must sum up to 100%. [Note: we will reward you up to \$2.61 for accuracy]

Below \$763.20	<input type="text" value="0"/>	%
Between \$763.20 and \$827.13	<input type="text" value="0"/>	%
Between \$827.13 and \$868.88	<input type="text" value="0"/>	%
Between \$868.88 and \$932.81	<input type="text" value="0"/>	%
Above \$932.81	<input type="text" value="0"/>	%
Total	<input type="text" value="0"/>	%



We have taken a random sample of 5 employees who hold your same position (Teller) and work in your same unit (Branch 10), and calculated the average basic salary among them. With the following set of questions, we want to assess how much you would be willing to pay to obtain this information about average salary.

Below you are presented with 5 hypothetical scenarios. In each scenario, you will be given the choice of either seeing the information about average salary OR receiving extra money as part of your reward for responding to the survey.

We will randomly choose 20 survey respondents. If you are one of these 20 lucky respondents, one of the 5 scenarios will be randomly chosen to be implemented. As a result, it is in your best interest to respond honestly to these scenarios.

Please make your hypothetical choices below, and in the next screen you will find out if your responses will be implemented or remain hypothetical.

Scenario 1: Between the next two options, which one would you prefer?

Information about average salary	\$1.30
<input type="radio"/>	<input type="radio"/>

Scenario 2: Between the next two options, which one would you prefer?

Information about average salary	\$6.52
<input type="radio"/>	<input type="radio"/>

Scenario 3: Between the next two options, which one would you prefer?

Information about average salary	\$26.09
<input type="radio"/>	<input type="radio"/>

Scenario 4: Between the next two options, which one would you prefer?

Information about average salary	\$130.46
----------------------------------	----------

Scenario 5: Between the next two options, which one would you prefer?

Information about average salary	\$652.32
----------------------------------	----------

You have NOT been selected among the 20 participants who will have one of their 5 scenarios implemented. As a result, your choices in the 5 scenarios remain hypothetical.

Please go to the next screen to continue with the survey.



Next, a group of individuals participating in this survey will be chosen to receive some information about the average salary in their same position and unit.

Please continue to the next screen to find out if you will be selected to receive this information.



You have been selected to receive the following information.

We have randomly chosen a random sample of 5 employees who work in your same position (Teller) and unit (Branch 10). The following is the average basic salary in this sample of 5 employees as of March of 2017: \$861.

Please take some time to read and understand this information carefully. When you are ready, proceed to the next screen.

>>

We want to give you the opportunity to re-assess your answer to one of the previous questions. This opportunity is given automatically to all survey participants, regardless of their responses.

What is the average monthly basic salary among all employees who work in your same position (Teller) and unit (Branch 10) as of March 2017? [Notes: we will reward you up to \$2.61 for accuracy]

In the previous question, you reported to believe that the average monthly basic salary among employees similar to you was \$913 in March 2017. The next question is designed to assess how confident you feel about your response.

With what probability do you think that the real average could fall in each of the following bins? The probabilities must sum up to 100%. [Note: we will reward you up to \$2.61 for accuracy]

Below \$821	<input type="text" value="0"/> %
Between \$821 and \$891	<input type="text" value="0"/> %
Between \$891 and \$936	<input type="text" value="0"/> %
Between \$936 and \$1,004	<input type="text" value="0"/> %
Above \$1,004	<input type="text" value="0"/> %
Total	<input type="text" value="0"/> %



Now consider the position Teller Supervisor, which is above your current position.

How many times do you think you would need to be promoted to reach that position (or another position in the same level)?

What is the likelihood that you will be promoted to position Teller Supervisor (or another position in the same or higher level) in the next 5 years?



Consider all employees from the bank who hold position Teller Supervisor. What was their average monthly basic salary as of March 2017? [Note: we will reward you up to \$2.61 for accuracy]

0

>>

In the previous question, you reported to believe that the average monthly basic salary among employees in position Teller Supervisor was \$2,609 in March 2017. The next question is designed to assess how confident you feel about your response.

With what probability do you think that the real average could fall in each of the following bins? The probabilities must sum up to 100%. [Note: we will reward you up to \$2.61 for accuracy]

Below \$2,348	<input type="text" value="0"/>	%
Between \$2,348 and \$2,544	<input type="text" value="0"/>	%
Between \$2,544 and \$2,674	<input type="text" value="0"/>	%
Between \$2,674 and \$2,870	<input type="text" value="0"/>	%
Above \$2,870	<input type="text" value="0"/>	%
Total	<input type="text" value="0"/>	%



We have taken a random sample of 5 employees in position Teller Supervisor, and calculated the average basic salary among them. With the following set of questions, we want to assess how much you would be willing to pay to obtain this information about average salary.

Below you are presented with 5 hypothetical scenarios. In each scenario, you will be given the choice of either seeing the information about average salary OR receiving extra money as part of your reward for responding to the survey.

We will randomly choose 20 survey respondents. If you are one of these 20 lucky respondents, one of the 5 scenarios will be randomly chosen to be implemented. As a result, it is in your best interest to respond honestly to these scenarios.

Please make your hypothetical choices below, and in the next screen you will find out if your responses will be implemented or remain hypothetical.

Scenario 1: Between the next two options, which one would you prefer?

Information about average salary	\$1.30
<input type="radio"/>	<input type="radio"/>

Scenario 2: Between the next two options, which one would you prefer?

Information about average salary	\$6.52
<input type="radio"/>	<input type="radio"/>

Scenario 3: Between the next two options, which one would you prefer?

Information about average salary	\$26.09
<input type="radio"/>	<input type="radio"/>

Scenario 4: Between the next two options, which one would you prefer?

Information about average salary	\$130.46
<input type="radio"/>	<input type="radio"/>

Scenario 5: Between the next two options, which one would you prefer?

Information about average salary	\$652.32
<input type="radio"/>	<input type="radio"/>

>>

You have NOT been selected among the 20 participants who will have one of their 5 scenarios implemented. As a result, your choices in the 5 scenarios remain hypothetical.

Please go to the next screen to continue with the survey.



Next, a group of individuals participating in this survey will be randomly chosen to receive some information about the average salary in position Teller Supervisor.

Please continue to the next screen to find out if you will be selected to receive this information.



You have been selected to receive the following information.

We have randomly chosen a random sample of 5 employees in position Teller Supervisor. The following is the average basic salary in this sample of 5 employees as of March of 2017: \$2,087.

Please take some time to read and understand this information carefully. When you are ready, proceed to the next screen.



We want to give you the opportunity to re-assess your answer to one of the previous questions. This opportunity is given automatically to all survey participants, regardless of their responses.

Consider all employees from the bank who hold position Teller Supervisor. What was their average monthly basic salary as of March 2017? [Note: we will reward you up to \$2.61 for accuracy]

In the previous question, you reported to believe that the average monthly basic salary among employees in position Teller Supervisor was \$2,348 in March 2017. The next question is designed to assess how confident you feel about your response.

With what probability do you think that the real average could fall in each of the following bins? The probabilities must sum up to 100%. [Note: we will reward you up to \$2.61 for accuracy]

Below \$2,113	<input type="text" value="0"/> %
Between \$2,113 and \$2,289	<input type="text" value="0"/> %
Between \$2,289 and \$2,407	<input type="text" value="0"/> %
Between \$2,407 and \$2,583	<input type="text" value="0"/> %
Above \$2,583	<input type="text" value="0"/> %
Total	<input type="text" value="0"/> %



Now, we want to ask you a few questions about your job at [Bank Name].

Recall that as of March of 2017, your monthly basic salary was \$730.

What do you expect your basic salary to be one year later, in March of 2018?

[Note: we will compare your response to our own projection of your future salary, and we will reward you up to \$2.61 if your response is close to our projection]

And what do you expect your basic salary to be five year later, in March of 2022?

[Note: we will compare your response to our own projection of your future salary, and we will reward you up to \$2.61 if your response is close to our projection]

How satisfied are you with your current salary at [Bank Name]?

- Very satisfied
- Somewhat satisfied
- Neither satisfied nor dissatisfied
- Somewhat dissatisfied
- Very dissatisfied

Across the thousands of [Bank Name] employees, salaries vary with the nature of work, education, experience, responsibilities, etc. What do you think of wage differentials in the company today?

- They are too large
- They are adequate
- They are too small

Taking all the aspects of your job into account, how satisfied are you with your current job at [Bank Name]?

- Very satisfied
- Somewhat satisfied
- Neither satisfied nor dissatisfied
- Somewhat dissatisfied
- Very dissatisfied



What percentage of employees in your same position and unit were assigned to each of the following KPI ratings as of year 2016?

The probabilities must sum up to 100%.

[Note: we will reward you up to \$2.61 for accuracy]

A1	<input type="text" value="0"/>	%
A2	<input type="text" value="0"/>	%
A3	<input type="text" value="0"/>	%
B	<input type="text" value="0"/>	%
C	<input type="text" value="0"/>	%
Total	<input type="text" value="0"/>	%

In comparison to others, are you a person who is generally willing to give up something today in order to benefit from that in the future or are you not willing to do so?

Please use a scale from 1 to 10, where a 1 means you are “completely unwilling to give up something today” and a 10 means you are “very willing to give up something today”. You can also use the values in-between to indicate where you fall on the scale.

1 (Completely Unwilling)	2	3	4	5	6	7	8	9	10 (Very Willing)
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

The survey is almost over. Now we want to assess your attitudes towards transparency. Remember that all your responses are confidential.

How often do you talk about salaries with coworkers?

-
-
-
-
-

If the bank shared with you data on the average pay for all positions. Which positions would you be most interested to look at? Please rank the following options from 1 (most interesting) to 4 (least interesting) by moving the boxes upward or downward:

You own position

Positions right above your level

Positions two levels above of your own position

Other positions



Currently at [Bank Name], salaries are confidential information. Please consider the following two hypothetical scenarios.

Scenario A: the bank created a website showing the average salary by position/unit, for all positions within the bank.

Would you be in favor or against the creation of a website like this?

- Strongly in favor
- In favor
- I would not care
- Against
- Strongly against

Scenario B: the bank created a website with the list of names and salaries of all its employees, including your name and your salary. As a result, you could look up the incomes of any other employee, and any employee could look up your own income.

Would you be in favor or against the creation of a website like this?

- Strongly in favor
- In favor
- I would not care
- Against
- Strongly against



Thanks for completing the survey!

Your total reward for this survey will be the sum of three amounts:

†+ A fixed fee of \$6.52.

†+ The total rewards for the accuracy of your responses during the survey.

†+ A surprise amount, picked at random from the range \$3.26-\$14.35.

We will transfer your total reward to your [Bank Name] account after the survey collection is finalized, which may take up to 10 weeks. You do not need to contact us any further -- rest assured that we will notify you by email when the reward is deposited to your account.

Did you have any technical or language-related problems when doing the survey?

- Yes
- No



B Email with Invitation to Survey

Dear [Employee's Full Name],

We would like to invite you to participate in a survey for [Bank's Name]'s employees. It takes less than 30 minutes to complete the survey and, as a token of our appreciation, you will be receiving a monetary reward – the average reward is around \$30.

[Follow this link to take the survey](#)

This survey is conducted by [Bank's Name] in collaboration with researchers from U.S. universities such as Harvard University. It will help us understand how to communicate with our employees.

You were selected at random to receive this invitation, and all your responses will be completely confidential.

If you have any difficulty responding to this survey, please reply to this email or use the following contact points:

[Bank's Contact 1]

[Bank's Contact 2]

[Bank's Contact 3]

If the link does not work, just copy and paste the following URL to your Internet browser: [Survey's URL]

Thank you for your participation. Your contribution will help to make [Bank's Name] a better place.

Sincerely,

Chief Economist, [Bank's Name]

C Additional Details and Results

C.1 Pay Inequality

This section provides some summary statistics related to pay inequality in the firm where the experiment was conducted and compares it to the inequality in other organizations.

We start by measuring the overall within-firm inequality. We compute a measure used in other studies (Song et al., 2019): the ratio of the 10th to 90th percentile of the distribution of base salary is 0.21. This degree of inequality is quite similar to that of medium sized firms in the United States: the ratio is 0.19 for the average firm with 5,000–10,000 employees (Song et al., 2019). The results are similar if we use different criteria, such as the ratio between the 90th percentile earner and the median earner.

Next, we can decompose the inequality by the horizontal and non-horizontal variation. Let $S_{i,p}$ be the salary of employee i in peer group p (i.e., pair position-unit). By construction:

$$S_{i,p} \equiv (S_{i,p} - \bar{S}_p) + \bar{S}_p \quad (\text{C.1})$$

Then, we can compute the variance of both sides of the equation:

$$\text{var}_i(S_{i,p}) = \text{var}_i(S_{i,p}|i \in p) + \text{var}_p(\bar{S}_p) \quad (\text{C.2})$$

The total dispersion in salaries ($\text{var}_i(S_{i,p})$) is the sum of the horizontal dispersion ($\text{var}_i(S_{i,p}|i \in p)$), weighted by the share of employees in each position, plus the non-horizontal dispersion ($\text{var}_p(\bar{S}_p)$) across positions. Then, we can express the share of horizontal inequality (S_h) as:

$$S_h = \frac{\text{var}_i(S_{i,p}|i \in p)}{\text{var}_i(S_{i,p}|i \in p) + \text{var}_p(\bar{S}_p)} \quad (\text{C.3})$$

Using data on the universe of employees in the firm where the experiment was conducted, we find that only 4.5% of the differences in base salary are horizontal. We can compare the contribution of horizontal inequality with the ones reported in other organizations. We start with the organization from Card et al. (2012). We obtained data on the regular pay in 2014 for all at the employees at the different campuses of University of California. We define peer groups as the combination of position title and department (e.g., one peer group could be the Assistant Professors at the Business School of UCLA). In our study, horizontal inequality accounts for 4.5% of the overall inequality in base salary. We find estimates in the same order of magnitude across the different UC campuses, ranging from a minimum of 7.0% in UC Merced to a maximum of 19.4% in UC San Francisco.

We can also provide a rough comparison to Baker et al. (1994). Using data from a large

U.S. firm, they report that dummies for the eight job levels (i.e., a coarse measure of the vertical position within the firm) explain around 70% of the variation in logarithm of salary. This evidence suggests that vertical inequality probably explains the vast majority of the salary differences within their firm. We used data from our firm to replicate that regression from Baker et al. (1994). We estimated a regression of the logarithm of base salary on a set of dummies for the nine paybands. We find that, like in Baker et al. (1994), a coarse measure of vertical rank picks up the vast majority of the pay inequality. In data from our firm, the level dummies explain 84% of the variation in salaries, which is comparable in magnitude to the corresponding 70% reported in Baker et al. (1994).

We can also provide some comparisons of the country context. While making these comparisons, however, we have to be careful because the employees in our sample are by no means a representative sample of the whole country. On the contrary, our sample represents the richest, most educated segment of the country. For example, in the country where the firm is located, less than 10% of the population aged 25 or older had a College degree in 2019. This share is much lower than for the United States, where the U.S. Census Bureau estimates that share at around 36.0% in 2019. However, among the employees who participated in our study, 86% of them had a College degree. In other words, our sample at hand is almost two and a half times more educated than the average American.

With those caveats in mind, the results for the country context are presented in Figure C.1. Each panel presents a histogram for one specific preference measure from the *Global Preferences Study* across the 76 countries covered in the data. The preference measures are: positive reciprocity, negative reciprocity, altruism, trust, risk taking and patience. For reference, each histogram highlights in blue the bar corresponding to the country where the firm is located. All six panels from Figure C.1 show that the country we study is about average for the 76 countries covered in the data.

C.2 More Details about the Measures of Effort and Performance

In this section, we provide more details and descriptive statistics about the measures of effort and performance: the number of hours worked, emails sent and sales performance.

The data on emails sent is available for all 2,060 employees in the subject sample. The number of hours worked and sales performance, however, is only available for different subsamples of employees: the hours worked can only be measured for employees working in one of the two headquarter offices, while the sales performance is only defined for employees who have some kind of sales role. We observe the email data for all employees. As a result, there is by construction a full overlap between the number of emails sent and the other two outcomes. To illustrate the overlap between the availability of data on hours worked and

sales performance, Table C.1 offers a simple cross-tabulation. The columns (*Sales Data*) indicate whether we observe a sales outcome during the 3 months post-treatment. The rows (*Hours Data*) indicate whether we observe an outcome for hours worked during the 3 months post-treatment. There is almost no overlap between the measures of hours worked and sales performance: we observe data on both of these outcomes for only 67 out of the 2,060 employees. The reason for the lack of overlap is straightforward: only a minority of employees who work in the headquarter offices (for whom we observe the hours worked) have a sales role, and thus we cannot measure sales performance for the majority of such employees.

Figure C.2 shows the distribution of each of the three measures of effort and performance (i.e., the exact same variables used as dependent variables in columns (1) through (3) of Table 2). Figure C.2.a corresponds to the number of hours worked, while Figure C.2.b corresponds to the number of emails sent and Figure C.2.c corresponds to the sales performance. These outcomes are defined exactly as in the regression analysis: i.e., the logarithm of the daily averages over the 3 months post-treatment.

Figure C.3 provides binned scatterplots showing the association between these outcomes. If these outcomes are all picking up effort and performance, they are expected to be positively correlated with each other. Before presenting these results, however, we must point out that we report the raw associations, and as such they are subject to a number of caveats. First, the timing of the association between these outcomes does not need to coincide exactly: for example, it is possible that an increase in the number of hours worked generates higher sales (and thus a higher sales performance index) months later rather than immediately. Second, there is quite a bit of measurement error and volatility in these outcomes, which can generate significant attenuation biases when looking at their associations. Third, the raw correlations do not control for any other employee characteristics (e.g., due to the nature of the tasks, some positions may involve more emails than others). Despite all these caveats however, we should at least expect the associations between the pairs of outcomes to have the right sign.

Figure C.3.a corresponds to the association between the number of emails sent and the number of hours worked, based on data for the 602 subjects for whom we observe both of these outcomes. We find a positive association: a 1% increase in the number of emails sent is associated to a 0.127% increase in the number of hours worked, and this relation is statistically significant (p-value=0.019). Figure C.3.b corresponds to the association between the number of emails sent and the sales performance index, based on data for the 791 subjects for whom we observe both of these outcomes. We observe a positive association here too: a 1% increase in the number of emails sent is associated with a 0.076% increase in the sales performance (p-value=0.011). In sum, and as expected, we find a positive and statistically

significant association between the different measures of effort and performance.⁷⁰

C.3 Randomization Balance and Descriptive Statistics

In this section, we present some descriptive statistics about the subject pool.

Of the 3,841 invitations sent out, 2,060 individuals completed the main module of the survey. This final sample excludes some participants based on pre-treatment characteristics. We excluded 23 participants who were randomly assigned to have their choices in the information-shopping scenarios executed and, as a result, their surveys were programmed to be automatically terminated. Among these 23 participants, 52% had been randomly assigned to receive information about peer salary and 52% had been assigned to receive information about manager salary. The final sample also excludes 1 subject who failed the training test for the definition of own salary: this subject reported an own salary that was over twice as large as the true own salary and then responded as disagreeing with our definition of own salary. This subject was assigned to later receive information on peer salary and manager salary. Last, we excluded 14 subjects with the most extreme prior misperceptions about peer salary (misperceptions over 100%) and manager salary (misperceptions over 400%): of these 14 subjects, 1 were assigned to receive feedback about peer salary and 3 were assigned to receive feedback about manager salary.

Table C.2 presents some descriptive statistics. Column (1) corresponds to the entire subject pool. On average, subjects are 29 years old and have been working at the firm for five years. 73% of them are female and 86% have a college or higher degree. We can check whether there is balance in observables across treatment groups. Subjects were cross-randomized to receive information about manager and peer salary, which resulted in four treatment groups. In columns (2) through (5) of Table C.2, we break down the average characteristics by each of the four treatment groups. The last column reports p-values for the null hypothesis that each average characteristic is constant across the four treatment groups. The results show that, as is consistent with successful random assignment, the observable characteristics are balanced across treatments.

Table C.3 presents descriptive statistics for different samples of employees. Column (1) corresponds to the universe of employees, while columns (2)–(5) correspond to different subsets of the sample: columns (2) and (3) provide summary statistics for the sample of individuals who were not invited and were invited to the survey, respectively; columns (4) correspond to the employees who were invited but did not respond to the survey; and column (5) corresponds to the final sample of 2,060 survey respondents. By comparing columns (1) and (5),

⁷⁰Figure C.3 does not include the association between the hours worked and sales performance because, as shown in Table C.1 and discussed above, there is almost no overlap between those two measures.

it follows that our sample is quite representative of the universe of employees. Even though some of the differences in gender, age, education and tenure are statistically significant, they are always economically small (given the large sample size, we have enough statistical power to detect even small differences). For instance, the subject pool is 73% female vs. 71% female in the universe, the mean ages are 29.2 vs. 30.1 years old, the shares of College graduates are 86% vs. 87%, and the mean tenures are 4.99 vs. 5.09 years.

The only non-trivial difference between the subject pool and the universe of employees is with respect to salary: the average salary in the subject pool is 28% lower than in the universe of employees. We can use the results from the rest of the columns to figure out where this difference is coming from. The comparison between columns (4) and (5) shows that this difference is not coming from differential response rates: the average salary of the survey respondents is quite similar (just 7.5% higher) to the average salary of non-respondents. The comparison between columns (2) and (3) shows that, instead, the bulk of the difference in mean salary between the subject pool and the universe of employees is coming from the selection of employees to be invited to the survey. By construction, we did not send the survey invitation to employees in the highest paybands. The high salaries of some of the excluded employees, such as the CEO and senior vice-presidents, can account for the difference in average pay between the subject pool and the universe of employees.

We provide some suggestive evidence on the accuracy of perceptions about the promotion opportunities. The average subject thinks 3.65 promotions are needed to reach the managerial position they are asked about. To construct a benchmark, we leverage the fact that increases in pay grade typically, although not always, indicate a promotion. Thus, one reasonable proxy for the number of promotions required to reach the managerial position is the difference in pay grades between the employee and his or her manager. On average, employees were 4.32 pay grades away from the managerial position. This distance seems consistent with the subjects' perceived need of 3.65 promotions to reach the managerial position. Moreover, the perceived number of promotions needed to reach the managerial position is significantly correlated to the actual number of pay grades separating the employee from the managerial position (correlation coefficient of 0.403, p -value <0.001).

Last, we provide more details about the definition of the post-treatment outcomes. For the minority of employees who leave the company during the relevant time window, we use the average outcome between the survey date and the exit date. For outcomes that are based on monthly data, such as sales, the post-treatment period corresponds to the month when the survey was taken and the following two months. This specification can lead to an attenuation bias because individuals who respond to the survey on the first day of the month (who were exposed to the information for a full month) would be coded the same as

individuals responding on the last day of the month (who were exposed for one day).

C.4 Training Module on the Definition of Base Salary

In this section we provide evidence that individuals understood the definition of base salary that we were using when eliciting beliefs about salaries.

In the training module of the survey, after we provided details about the definition of base salary, we used an incentivized question to ask respondents to report their own base salaries. Figure C.4 compares the individual's guesses about their own salaries with respect to their actual salaries. More than 80% of respondents provide a guess of own salary that is within 5% of the truth – moreover, the typical respondent reported their salary exactly up to the last digit. This outcome confirms the anecdotal evidence that base salary is the most salient aspect of compensation. The remaining 20% of employees missed the mark, typically by a large margin. While it is possible that some misperceptions about own salary exist, these large differences are more consistent with a misunderstanding of the definition of salary. For example, some respondents seem to have reported their salary after taxes and other deductions while in fact we were asking about the salary before taxes. Some respondents provided a guess that is less than half the actual base salary. These employees probably reported an alternative definition of salary that is only relevant for tax purposes (the word in the native language happens to sound similar to base salary). To make sure that these misunderstandings do not extend to the rest of the survey, to the employees who did not guess accurately we showed the employees what their base salary is and then explained the definition of the base salary once again. After this second round of training, we asked these employees who had misreported their own salary if they agreed with our measure and 87% of them responded affirmatively.

C.5 Heterogeneity of Misperceptions and Willingness to Pay

In this section, we provide some heterogeneity analysis of the misperceptions and willingness to pay for salary information.

Table C.4 presents evidence on the heterogeneity of the average error, absolute error and willingness to pay for information. The first set of three columns correspond to perceptions about average manager salary, while the second set of three columns correspond to perceptions about average peer salary. The first row of Table C.4 shows the averages over the entire subject pool. The rest of the rows break down these averages by different subgroups: females vs. males, above and below 4 years of tenure, higher vs. lower paybands, sales vs. non-sales roles, and front office vs. back office roles. In each of these breakdowns, we report

the p-value of the test of the null hypothesis that the relevant average is the same across the two sub-groups.

The most important result from this table is that the patterns are qualitatively consistent across all different subgroups: the absolute error of manager salary is always large, and there is always a systematic bias that is negative and large; the absolute error of peer salary and its bias is also similar across subgroups; and the willingness to pay is always in the neighborhood of \$250. This evidence suggests that the results are not driven by any specific group of the population.

Due to the large sample size, we have enough statistical power to detect even small differences, and for that reason many of the differences in magnitude are statistically significant. However, a large majority of the differences are economically small. We mention below some of the exceptions. Regarding the error on manager salary, the most notable difference is between front (-10%) vs. back office (-24%) roles. Regarding the absolute error on manager salary, the most notable difference is again between front (25%) and back office (36%) employees, which arises mechanically from the differences in biases. Regarding the willingness to pay for information on manager salary, the most notable difference is by front office (\$201.90) versus back office (\$164.91) roles. Regarding the error on peer salary, the most notable exception is the difference by tenure: a systematic bias of 0% for employees with less than 4 years of tenure vs. 5% for employees with more than 4 years of tenure. Regarding the absolute error on peer salary, there are no notable differences. Regarding the willingness to pay for peer salary, the most notable difference is by tenure: \$218.33 for higher tenure vs. \$174.79 for lower tenure.

C.6 Robustness Checks for the Willingness to Pay Data

In this section, we discuss some limitations of the BDM elicitation method, and some evidence addressing those concerns.

The first concern is that our estimates of willingness to pay may be sensitive to the elicitation method – in particular, the lists of prices given in the hypothetical scenarios may act as a suggestion for what the employees “should” pay for the information. As a robustness check, we can take advantage of the fact that we measured willingness to pay in a followup study (Cullen and Perez-Truglia, 2018). In that study, we measured willingness to pay for a similar piece of information: the average salary among a random sample of five peers. Instead of using the price-list method, however, we used an open-ended variation (Andersen et al., 2006), in which the respondent bids against the computer. The rules are as follows. The respondent’s bid is compared to a price that is determined by a random number generator. If the respondent’s bid is lower than the price, then the respondent gets

a dollar amount equal to the price. If the bid is higher than the price, then the subject gets the information and no dollar amount. The rules of this mechanism makes it a dominant strategy for respondents to bid exactly their true valuation for the information. The results are presented in Figure C.5. The left panel corresponds to the willingness to pay as measured in this study, while the right panel replicates the left panel but using the data from Cullen and Perez-Truglia (2018). The distributions should not be expected to be identical, because they are based on non-overlapping samples of employees which differ in a number of features. Despite these differences, the distributions of willingness to pay are in the same order of magnitude across the two elicitation methods. Moreover, this finding is consistent with evidence from other studies showing that measures of willingness to pay are not equal but still largely similar across different elicitation methods (Brebner and Sonnemans, 2018).

Last, despite our efforts for making the willingness to pay elicitation easy to understand, measurement error may still exist. In other words, while subjects may seem quite heterogeneous in their bids for information, some of that heterogeneity may simply reflect factors such as misunderstanding how the elicitation mechanism works. However, our main objective is not to study the heterogeneity in information value, but mostly to assess the average value of information. As a result, the sources of measurement error are likely to be averaged out.

C.7 Determinants of Willingness to Pay for Salary Information

We find that some employees are willing to pay large amounts to acquire information about the average salary of their peers or managers. Our favorite interpretation for that finding is that the interest emerges not just due to curiosity, but primarily due to the instrumental value of the information. In this section, we provide some suggestive evidence in favor of this interpretation.

We test two hypotheses. The first hypothesis is related to the willingness to pay for information about the average peer salary. One of the instrumental reasons for wanting to buy information about the average peer salary is that this information can be useful for salary negotiations. The employee may want to use the information about peer salary as a bargaining chip – for example, they may plan to take a picture of this information and use it as an argument for a raise with their manager or Human Resources representative. Even if the employee was not planning on using the information in the meeting directly, he or she may still use it to decide whether it would be worthwhile to ask for a raise in the first place. According to this channel, we would expect employees who are under-paid to benefit the most from this information, and thus to be willing to pay the most for it, relative to employees who are over-paid.

The results for this first hypothesis are presented in columns (1)–(3) of Table C.5. Each

of the columns from this table corresponds to a separate interval regression, and all of these regressions control for tenure and past performance evaluations. In columns (1) through (3) the dependent variable is the willingness to pay for information on peer salary. In column (1), the key independent variable is the employee's perceived relative salary: i.e., the (log) difference between the employee's own salary and his or her perception of the average peer salary. The coefficient on relative salary is negative (-141.361) and statistically significant (p-value=0.070). This coefficient implies that, consistent with the hypothesis at hand, the employees who are under-paid relative to their peers are the ones most willing to buy information about the average peer salary. The magnitude of this effect is economically significant too: relative to an employee who earns 10% above the peer average, an employee who earns 10% below average would be willing to pay an additional \$31 ($= -141 \cdot (e^{0.2} - 1)$) for the information on the average peer salary.

Columns (2) and (3) of Table C.5 provide two robustness checks for the results presented in column (1). According to this channel, what should really matter for the willingness to pay for information is not what the employee's true relative salary is, but what the employee believes this relative salary to be. In other words, even if an employee was being over-paid relative to peers, if that employee thinks that he or she is under-paid, then he or she should still be willing to pay more for information about the peer salary. The results reported in columns (2) and (3) are consistent with this conjecture. The specification from column (2) is identical to that from column (1), except that it includes as a regressor the employee's true relative salary instead of the employee's perceived relative salary. The coefficient on the true relative salary reported in column (2) is close to zero (-16.166) and statistically insignificant (p-value=0.869). In turn, column (3) is identical to columns (1) and (2), except that both variables (the perceived and true relative salaries) are jointly included in the model. As expected, the coefficient on the perceived relative salary in column (3) is negative (-181.120), statistically significant (p-value=0.069) and, if anything, larger in magnitude relative to the corresponding coefficient from column (1); in turn, the coefficient on the true relative salary is statistically insignificant (p-value=0.462).

The second hypothesis is related to the willingness to pay for information about the manager salary. One of the instrumental reasons for wanting to buy information about the average manager salary is that this information can be helpful for career planning. For example, an employee may want to decide whether it is worth working harder to get promoted to the manager's position (or another position on the same level). Alternatively, an employee may be considering whether to seek an outside job offer, or may already be sitting on an outside offer. To decide whether it is worth staying at this firm or not, in addition to the current salary, the employee should consider the salary growth potential at the current firm.

According to this channel, the willingness to pay for information about a managerial position should depend on whether the employee can aspire to attain that position or not. If the manager's position is out of reach for the employee, then the salary at that position should be largely irrelevant for the employee's career planning. On the contrary, if the manager's position is within the employee's reach, then the information about the salaries in that position should be more useful for career planning and thus more valuable. To test this hypothesis, we can exploit the variation in the distance between the managerial position and the respondent's own position. For some respondents we elicited the willingness to pay for information about a managerial position that is a few promotions away (e.g., asking a junior analyst in investment banking about the salaries of the senior analysts), while for other employees we elicited the willingness to pay for information about the average salary in a managerial position that was well above the employee's current position.

The results for the second hypothesis are presented in columns (4) and (5) of Table C.5. In these two columns, the dependent variable is the willingness to pay for information about manager salary. In column (4), the key independent variable is the employee's perceived probability of being promoted within five years to the managerial position they are being offered the opportunity to buy information about. More precisely, *Perceived P(Promoted to Manager Position)* was constructed based on one of the questions included in the survey, and takes the values from 0.05 (0-10% probability) to 0.95 (90-95% probability), in steps of 0.05. Consistent with the hypothesis of instrumental value, the coefficient reported in column (4) is positive (160.140) and statistically significant (p-value<0.001). This coefficient is significant in magnitude too: relative to a managerial position that the employee could not aspire to reach (probability of promotion of 0%), the employee is willing to pay an additional \$160 to learn the average salary of a managerial position that the employee can aspire to reach (probability of 100%). The results from column (4) are of course subject to the caveat of potential omitted variable biases. In particular, one concern is that the variable on the probability of being promoted to the position may be just picking up the effects of the employee's own salary. To address this concern, column (5) includes the employee's own salary (in logs) in the regression jointly with the perceived probability of being promoted to the manager's position. Despite the addition of the new variable, the coefficient on the perceived probability of promotion remains similar both in terms of magnitude (160.140 in column (4) versus 160.403 in column (5)) as well as in terms of statistical significance (p-value<0.001 in columns (4) and (5) both).

Indeed, some simple back-of-the-envelope calculations indicate that the instrumental value of salary information can be quite significant (Stigler, 1962). For example, assume that an employee is considering acquiring information about peer salary for use in her salary

negotiations. If the expectation is that, with 50% probability, the information will help negotiate a one-year 10% raise, then the employee should be willing to pay up to two weeks of her salary for the information. And since employees may plan to use the information for multiple decisions (e.g., whether to switch jobs or positions), the value of the information can add up quite rapidly across the different margins. In addition to these back-of-the-envelope calculations, there are other approaches to estimating the value of information that also suggest this valuation can be significant. For example, Conlon et al. (2018) use a structural model to estimate the value of salary information for unemployed individuals. They estimate that the average U.S. college graduate looking for a job should be willing to pay \$817 to acquire full information about the distribution of future wage offers.

C.8 Further Analysis on Learning

In this section, we provide additional analysis and robustness checks regarding how individuals updated their beliefs based on the information provided to them.

In principle, it is possible for the signal of peer salary to affect beliefs about manager salary, or vice-versa. We explore this possibility in Figure C.6. Panels (a) and (b) of Figure C.6 just reproduce panels (a) and (b) from Figure 2. Panel (c) of Figure C.6 is similar to panel (b), only that the dependent variable is the revision about the manager salary: that is, this figure measures whether the signals about the average peer salary had a causal effect on the subsequent beliefs about the average manager salary. The slope is close to zero and statistically insignificant. In other words, the respondents used the signal about peer salary to update beliefs about peer salary but did not use signals about the average peer salary to update beliefs about the average manager salary. For the sake of completeness, we panel (d) of Figure C.6 is similar to panel (b), except the dependent variable is the revision about the peer salary: that is, it measures whether the signal about manager salary had a causal effect on the beliefs about peer salary. Note such updating would be practically impossible, because the posterior beliefs about peer salary were elicited before the provision of the signal about the manager salary. As expected, Figure C.6.d shows that the signal about the manager salary does not have a causal effect on the reported beliefs on peer salary. Due to the timing of the survey, this does not prove that individuals did not extrapolate from the signal of manager salary to the belief about peer salary. However, this seems unlikely given the above evidence that subjects did not extrapolate from peer salary to manager salary.

For the sake of completeness, Figure C.7 provides a more detailed view of the results from Figure C.6. Figures C.7.a and C.7.b break down the results from Figure C.6.c in treatment and control groups, and they include the raw scatterplots in addition to the binned scatterplots. Likewise, Figures C.7.c and C.7.d break down the results from Figure C.6.d with

the added scatterplots. The results shown in Figure C.7 are essentially the same as those presented in Figure C.6.

Next, we can assess whether certain subgroups of the population were more likely to incorporate the signals in their posterior beliefs. Table C.6 measures the heterogeneity in learning rates by subgroups of the population. Each column corresponds to a different subgroup: column (1) breaks down the sample by gender, column (2) by tenure, column (3) by payband, column (4) by sales role and column (5) by front versus back office. The top and bottom halves of each column correspond to the estimated learning rates for each subgroup – e.g., in column (1), the top half corresponds to the results for females and the bottom half corresponds to the results for males. The table also reports p-values from the test of the null hypothesis that the learning rates are equal between the two subgroups. The results from Table C.6 indicate that all the differences in learning rate are small in magnitude. For instance, the learning rate for peer salary is 0.536 for females and 0.456 for males, with a difference p-value of 0.413; while the learning rate for manager salary is 0.673 for females and 0.748 for males, with a difference p-value of 0.285. Moreover, only one of the ten differences (five for peer salary and five for manager salary) is statistically significant at conventional levels.

C.9 Additional Results on Information Diffusion

In this section, we present some additional results on information diffusion.

To aid in the interpretation of the results reported in Section 3.4, Table C.7 provides some basic descriptive statistics of all the main variables used for the analysis. Next, we provide a series of robustness checks for the experimental results discussed in Section 3.4.

In the baseline econometric model, given by equation (7), the variable I_i^M indicates if the individual received information about manager salary indirectly. In the baseline definition, this variable takes the value zero if the subject received the information directly. The rationale behind this specification is that if the individual already received the information directly, then whatever information her or she can receive indirectly through peers is pretty much redundant. Even though each individual receives a signal based on a different sample of five peers, the sampling variation in the signals is limited. Thus, the marginal informational value from observing a second signal is quite limited. As a robustness check, Table C.8 reproduces the results from Table 1 but allowing I_i^M to take non-zero values even when the respondent received the information directly. The results from the alternative specification of Table C.8 are almost identical to the results from the baseline specification of Table 1. Another aspect of the baseline specification of equation (7) that is worth discussing further

is that the dependent variable is based on posterior beliefs: $M_i^{abs} = \left| \frac{M_i^{post} - M_i^{true}}{M_i^{true}} \right|$. This specification focuses on posterior beliefs to be able to have the effects of direct information assignment as a benchmark. Due to the structure of the survey, the direct assignment can only affect posterior beliefs. However, if we only want to measure the effects of the indirect assignment, we can focus instead on the effects of prior beliefs. Table C.9 reproduces the results from Table 1 but using prior beliefs as dependent variable: $\left| \frac{M_i^{prior} - M_i^{true}}{M_i^{true}} \right|$. Moreover, Table C.9 uses the definition of I_i^M that can take non-zero values even when the respondent received the information directly (because the respondent would not see that direct information until after the elicitation of the prior beliefs). The results from the alternative specification of Table C.9 are almost identical to the results from the baseline specification of Table 1.

In the baseline results discussed in Section 3.4, the dependent variable is the degree of misperceptions, as measured by the absolute value of the difference between perceptions and reality: $M_i^{abs} = \left| \frac{M_i^{post} - M_i^{true}}{M_i^{true}} \right|$. An alternative outcome of interest is the degree of bias: $M_i^{bias} = \frac{M_i^{post} - M_i^{true}}{M_i^{true}}$. The average of this outcome measures the systematic tendency to either under-estimate or over-estimate average salaries. The results from this alternative dependent variable are presented in Table C.10, which is identical to Table 1 except that the dependent variable is the raw difference between perceptions and actual salaries instead of the absolute difference. In columns (1) through (5) from Table C.10, the dependent variable is the bias on the manager salary. The negative constant from column (1) is negative (-0.088) and statistically significant (p-value<0.001). This estimate indicates that, in the absence of feedback, posterior beliefs end up under-estimating the manager’s true average salary on average. The coefficient on *Direct* from column (1) is positive (0.105) and statistically significant (p-value<0.001). Moreover, the magnitude of the coefficient on *Direct* is similar (in absolute value) to the estimate for the constant. This result indicates that the direct feedback fully eliminates the tendency to underestimate manager salary. Most importantly, the coefficients on all of the different variables associated with indirect feedback in columns (2) through (5), such as *Closest Peer*, are close to zero, statistically insignificant and precisely estimated. Moreover, columns (6)–(10) of Table C.10 reproduce the same analysis as columns (1)–(5) but look at peer salary instead of manager salary. Again, we find robust evidence of an absence of information diffusion.

Intuitively, the results from the baseline specification in Section 3.4 suggest that posterior beliefs get more “compressed” around the truth whenever the subject receives feedback directly, but does not get more compressed around the truth when the subject’s peers receive feedback. By means of graphical analysis, we can inspect whether the feedback impacted the distribution of posterior beliefs in other ways. The results are presented in Figure C.8. Fig-

Figure C.8.a provides histograms of the percentage difference between posterior beliefs about the average manager salary and the true average, broken down by whether the subject received a signal of the manager salary (*Direct Feedback*, in red) or not (*No Direct Feedback*, in gray). Figure C.8.a shows that, consistent with the results presented in Section 3.4, the beliefs are more “compressed” around the truth for subjects who received a signal of the manager salary relative to those who did not. Indeed, according to the Kolmogorov-Smirnov test, we can reject the null hypothesis that those two distributions are the same (p-value<0.001). Figure C.8.b is similar to Figure C.8.a, but it corresponds to the subsample of subjects who did not receive feedback directly and breaks them down by whether their closest peer received feedback (*Indirect Feedback*) or not (*No Indirect Feedback*). This figure is also consistent with the results from Section 3.4 in that there is no evidence of information diffusion: we cannot reject the null hypothesis that the distribution of beliefs are equal between those who received feedback indirectly and those who did not (p-value=0.228). In turn, Figures C.8.c and C.8.d are similar to Figures C.8.a and C.8.b, except that they correspond to average peer salary instead of average manager salary. The results are again consistent with a lack of information diffusion: there is a significant difference in beliefs between subjects who received feedback directly versus those who did not (p-value<0.001, from Figure C.8.c); but there is no significant difference in beliefs between subjects who received feedback indirectly versus those who did not (p-value=0.340, from Figure C.8.d).

Last, we supplement the experimental test of information diffusion with a non-experimental test. The theoretical and empirical evidence in the literature indicates that, in presence of information diffusion, individuals who are most central in a network are the ones who are best informed (Alatas et al., 2016; Banerjee et al., 2013; Mobius and Rosenblat, 2014). We can test this hypothesis in our data by comparing the misperceptions between individuals who are more and less central in the network. To measure centrality, we use the directed network of emails sent by employees over the three months prior to the completion of the first survey. We exclude from this sample the emails directed outside of the institution and emails received from outsiders. These results are based on eigenvector centrality, but the findings are similar with alternative definitions of centrality.

The relationship between misperceptions (measured as mean absolute error) and network centrality is shown in Figures C.9.a and C.9.b. Figure C.9.a shows the raw data through a scatterplot. Since the raw data can be difficult to interpret, Figure C.9.b presents a binned scatterplot representation of that same data, along with a linear fit. Contrary to the hypothesis of information diffusion, we do not find that misperceptions about peer or manager salary decrease with network centrality. On the contrary, the slopes are slightly positive for manager salary (0.283, p-value=0.181) and peer salary (0.299, p-value=0.003). To illus-

trate these magnitudes, we can compare the misperceptions between employees who have below-median and above-median centrality. For perceived manager salary, the mean absolute error is 28.0% for individuals with below-median centrality and 28.2% for individuals with above-median centrality (difference p-value=0.910); we find similar results if we look at perceptions on peer salary. For peer salary, the misperceptions are 10.7% for employees with above-median centrality and 12.3% for employees with below-median centrality (difference p-value=0.002).

An alternative way of assessing information diffusion is by using self-reported data on whether employees communicate with each other. We use a survey proxy for whether employees engaged in information diffusion by asking directly how frequently they discussed salaries in the previous year. If accurate information flows in the network, individuals who reported to have discussed salaries with coworkers should have lower misperceptions (Alatas et al., 2016). The results are presented in binned scatterplot form in Figures C.9.c and C.9.d. Figure C.9.c shows the raw data through a scatterplot, while Figure C.9.d presents a binned scatterplot representation of that same data, along with a linear fit, for an easier interpretation. Contrary to the prediction of information diffusion, we find misperceptions to be statistically indistinguishable between employees who discussed salaries in the past and those who did not. The slopes between the misperceptions and frequency of communication are close to zero and statistically insignificant for both peer and manager salary. Again, to illustrate the magnitude of these differences we can compare misperceptions across employees who never discussed salaries and employees who discussed it once or more. The manager misperceptions are 27.8% for individuals who discussed salaries vs. 28.5% individuals who did not discuss salaries (difference p-value=0.49). For peer salary, the corresponding misperceptions are 11.5% for individuals who discussed salaries vs. 11.6% individuals who did not discuss salaries (difference p-value=0.88). This evidence suggests that, even if employees sometimes discuss salaries with coworkers, they may be sharing noisy or misleading information, or they must not be processing it properly.

C.10 Reduced Form and First Stage Results

Table 2 presents the main results from the IV estimator. For reference, Table C.11 presents the results from the corresponding reduced form and first stage regressions. The results from the reduced form regression are qualitatively similar to the results from the Two-Stage Least Squares regression. The only difference between the two sets of coefficients exists in terms of magnitudes. The first stage results show that individuals did not fully incorporate the feedback given to them. The IV estimates simply scales the reduced form estimates to correct for the incomplete reaction to the information.

C.11 Additional Specification Checks

In this section, we provide some additional specification checks to the results presented in Section 4.

To be able to capture proportional effects, the baseline specification used for Table 2 used the logarithmic transformation for the dependent variables. Next, we present results under a different approach that does not rely on using the logarithmic transformation. In this alternative specification, the dependent variable is defined as the ratio between the post-treatment outcome and the pre-treatment outcome. As such, this specification should be able to capture proportional effects without the need for the logarithmic transformation. The downside of the alternative specification is that, unlike the baseline specification, it cannot be readily interpreted as an elasticity.

The results are presented in Table C.12. Columns (1) through (3) of Table C.12 are a reproduction of the baseline results from the top panel of Table 2. The specifications in columns (4) through (6) of Table C.12 are identical to columns (1) through (3) except that, instead of using the logarithmic transformation, we use the alternative specification discussed above. Since the dependent variables are different, we cannot compare the magnitudes of the coefficients between the baseline specification (columns (1)–(3)) and the alternative specification (columns (4)–(6)). However, we should at least expect them to be qualitatively consistent (i.e., in terms of sign and statistical significance). In terms of the sign, the results are highly consistent between the two specifications: the effects of manager salary are always positive, while the effects of peer salary are always negative. In terms of statistical significance, the results are weaker under the alternative specification (3 out of the 6 coefficients are statistically significant at the 5% level) than under the baseline specification (5 out of the 6 coefficients are statistically significant at the 5% level).

The baseline specification used for Table 2 controls for some basic characteristics of the employee: salary, tenure, and five productivity rating dummies. Table C.13 assesses how robust the results are when we include some additional employee characteristics as control variables. Columns (1) through (3) of Table C.13 are a reproduction of the baseline results from the top panel of Table 2. The specifications in columns (4)–(6) of Table C.13 are identical to columns (1)–(3) except for the inclusion of an additional control variable: a binary variable indicating if the employee is female. The inclusion of this additional control variable makes almost no difference: the results reported in columns (4)–(6) are qualitatively and quantitatively similar to the baseline results reported in columns (1)–(3). The specifications in columns (7)–(9) are identical to columns (1)–(3) except for the inclusion of the following additional control variables: a set of 29 indicator variables corresponding to the location where the employee works. Again, the results are quantitatively and qualitatively robust

to the inclusion of the additional controls. Last, the specifications in columns (10)–(12) are identical to columns (1)–(3) except for the inclusion of the following additional control variables: a set of 180 indicator variables corresponding to the employee’s position title. Relative to the number of observations (602, 2,060 and 791 in columns (1), (2) and (3), respectively) 180 is a large number of control variables. Moreover, some of those categories include a single subject, and thus the inclusion of those indicator variables is numerically equivalent to dropping those observations. For these reasons, the results from this demanding specification must be taken with a grain of salt. In any case, the results under this demanding specification are consistent with the baseline specification, although weaker in terms of statistical significance: 3 out of the 6 coefficients are statistically significant at the 5% level in columns (10)–(12), while 5 out of the 6 coefficients are statistically significant at the 5% level in the baseline specification given by columns (1)–(3).

Next, we assess the robustness checks to a couple of different features of the baseline specification. First, in the baseline specification it is important to control for prior gaps in beliefs. One potential concern is that, due to non-linearities, failing to control for prior beliefs flexibly may bias the estimates. This robustness check is presented in Table C.14. Columns (1) through (3) of Table C.14 are a reproduction of the baseline results from the top panel of Table 2. In turn, the specifications in columns (4)–(6) of Table C.14 are identical to columns (1)–(3) except for the addition of flexible controls for the prior gaps in beliefs about the manager salary and the peer salary. More specifically, in addition to including the prior gaps linearly, we also include sets of dummies for the deciles of the prior gaps (nine dummies for manager salary and nine dummies for peer salary). The findings are robust: the results from columns (4)–(6) are similar to the baseline results in columns (1)–(3), both qualitatively and quantitatively.

The second robustness check presented in Table C.14 relates to the role of belief certainty. When individuals receive information, they may not only change the first moment of the posterior belief (i.e., the mean) but may also change other moments of the distribution (e.g., the dispersion). For example, when an individual receives a signal x he or she may not only shift the mean of the probabilistic belief towards x but also make it more compressed around x (i.e., become more certain). This creates a potential concern: our specification may be attributing all the effects on behavior to the shifts in the first moment of beliefs, when in reality some of the effects are due to higher moments (e.g., certainty). It is unlikely that effects on certainty would explain all the effects we document, however. If, for example, the effects on behavior were due to certainty, we would expect the information to affect behavior in the same direction regardless of whether individuals are updating their mean belief up or down. Instead, our evidence suggests that the effects on behavior go in different directions

depending on whether the individual is updating the mean belief up or down. It is still possible, however, that there is a certainty channel and is contaminating the estimates to some extent – and for that reason, we address this concern empirically below.

To address these concerns, we leverage the fact that we elicited not only the mean of the posterior belief but also its distribution. Thus, we can include in the IV regressions the higher moments of the posterior beliefs as additional control variables. If the higher moments are responsible for the effects on behavior, then the IV coefficients should get closer to zero, perhaps even all the way. If, as argued above, the higher order beliefs are not doing any of the heavy lifting, the IV coefficients should remain unchanged when we add these additional controls. The results are presented in columns (7)–(9) of Table C.14, which are identical to the baseline specification from columns (1)–(3), except for the addition of eight control variables related to the distribution of posterior beliefs (four variables for the manager salary and four variables for the peer salary). For the manager salary, the four variables are: the perceived probability that the true average manager salary falls between -10% and -2.5% of the posterior belief of the average manager salary; the perceived probability that the true salary falls between -2.5% and +2.5% of the posterior belief; the perceived probability that the true salary falls between +2.5% and +10% of the posterior belief; the perceived probability that the true salary falls above +10% of the posterior belief. And the four variables on peer salary are defined in an analogous way to the four variables on manager salary. Note that by including this set of variables, we are controlling for the higher moments in a flexible way – that is, we do not need to make any functional form assumptions to estimate a certainty parameter. The findings are robust: the results from columns (7)–(9) are similar to the baseline results in columns (1)–(3), both qualitatively and quantitatively.

C.12 Effects on Other Email-Based Outcomes

In this section, we provide further analysis of the effects of salary perceptions on the employee’s email activity.

In Section 4, we use the number of emails sent by the employee as dependent variable. Our email data allows us to construct alternative dependent variables. The results are presented in Table C.15. Column (1) corresponds to the results from the baseline outcome: the number of emails sent. By construction, these coefficients are identical to the coefficients from column (2) of Table 2. In column (2) of Table C.15, the dependent variable is the number of emails received. While the number of emails sent is more directly linked to the employee effort, there is still an indirect effect of effort on emails received. For instance, if an employee increases the number of emails sent, we would expect some of those additional emails to be replied and thus to translate into additional emails received. Additionally, if the employee

takes on more responsibility while working longer hours, that could also translate into more emails received. The results suggest that the two coefficients for emails received (column (2)) are qualitatively consistent with the corresponding coefficients for emails sent (column (1)), only that smaller in magnitude – this difference is statistically significant for the coefficient on manager salary (p-value=0.096), but statistically insignificant for the coefficient on peer salary (p-value=0.310).

The rest of the columns from Table C.15 break down the effects on the emails sent by the identity of the receivers. First, we break down the number of emails sent by emails sent to emails accounts inside the same firm (column (3)) and emails sent outside of the firm (column (4)). The effects on these two outcomes are qualitative consistent – moreover, we cannot reject the null hypothesis that the two effects are equal to each other (p-values of 0.268 and 0.650 for the coefficients on manager and peer salary, respectively). The last three columns of Table C.15 break down the emails sent to other employees by the rank of those receivers: employees who are in higher paybands (column (5)), employees who are in the same payband (column (6)), and employees who are in lower paybands (column (7)). The results suggest that while the effects of manager salary operate mainly through emails sent to same- and higher-ranked employees, the effects of peer salary operate mainly through emails sent to lower-ranked employees. However, these results have to be taken with a grain of salt, because we cannot reject the null hypothesis that the coefficients are the same across the three columns (p-values of 0.161 for manager salary and 0.464 for peer salary).

C.13 Heterogeneity of Effects of Manager and Peer Salary

Table C.16 presents results on the heterogeneity of the effects of manager and peer salary. In all the regressions, the dependent variable is the number of emails sent. We focus on this outcome because it is available for the entire subject pool and thus it provides the most precision to detect heterogeneity. The top and bottom halves of the table present the coefficients estimated for different subgroups of the population. The first five columns use the same splits employed in the other analyses of heterogeneity: female vs. male (column (1)), higher vs. lower tenure (column (2)), higher vs. lower paybands (column (3)), sales vs. non-sales roles (column (4)), and front-office vs. back-office roles (column (5)). The bottom of Table C.16 provides p-values for the test of the null hypothesis that the coefficients are equal across a given pair of subgroups. We do not find any statistically significant evidence of heterogeneity: none of the ten differences are significant at conventional levels. This evidence suggests that our results are not driven by any specific group of the population. However, due to the precision of the coefficients, we cannot reject moderate differences either. The last column (column (6)) presents results for an additional heterogeneity (by the perceived

productivity rank of the employee) that is discussed in the section below.

C.14 Linearity of the Effects on Behavior

In the baseline model, we make the functional form assumption that the relationship between salary perceptions and behavior is log-log linear. In this section, we discuss and relax that assumption.

The baseline specification is simplest one, and is the most common specification in the literature on relative income concerns (see e.g., Senik (2004); Luttmer (2005); Clark et al. (2008); Clark and Senik (2010)). To assess whether this is a reasonable approximation, we provide binned scatterplot versions of the IV regressions from Table 2. The results are presented in Figure C.10. The three panels on the left correspond to the coefficients on manager salary and the three panels on the right correspond to the peer coefficients on peer salary. Each row of two panels correspond to a different regression: the first row corresponds to hours worked, the second row corresponds to emails sent, and the third row corresponds to sales performance. The statistical power available to conduct this type of analysis is limited, so one should not conclude the effects are perfectly linear. However, these binned scatterplots suggests that the log-log linear model is a reasonable approximation. Moreover, these binned scatterplots show that the results do not seem to be driven by outliers, in that the linear regression is not driven by any single bin.

C.15 Symmetry of the Effects on Behavior

Another functional form assumption from the baseline model relates to the symmetry of the responses. Let us start with the main object of interest, the vertical comparisons. Our baseline specification assumes that the effects of updating beliefs upwards are the mirror image of updating beliefs downwards. In practice, finding out that the managers are paid more than initially thought may have stronger or weaker effects than finding out that the managers are paid less than initially thought. For example, employees may have more flexibility to adjust their effort upwards in response to good news about the manager pay (e.g. by working extra hours) than to adjust their effort downwards in response to bad news. To allow for this type of asymmetries, we can augment the baseline model given by equation (8) as follows:

$$\begin{aligned} \log(Y_i^{post}) = & \eta_0 + \eta_{mgr}^{up} \cdot \mathbb{1}_{M_i^{post} > M_i^{prior}} \cdot \log(M_i^{post}) + \eta_{mgr}^{down} \cdot \mathbb{1}_{M_i^{post} \leq M_i^{prior}} \cdot \log(M_i^{post}) + \\ & + \eta_{peer}^{up} \cdot \mathbb{1}_{P_i^{post} > P_i^{prior}} \cdot \log(P_i^{post}) + \eta_{peer}^{down} \cdot \mathbb{1}_{P_i^{post} \leq P_i^{prior}} \cdot \log(P_i^{post}) + \xi_i \end{aligned} \tag{C.4}$$

While the baseline model had one parameter to represent the effects of manager salary (η_{mgr}), this extension has two parameters: one parameter for the upward revisions (η_{mgr}^{up}) and another parameter for the downward revisions (η_{mgr}^{down}). Likewise, the effects of the horizontal comparisons are split in two distinct parameters. It is straightforward to adjust the IV regression given by equations (9)–(11) to accommodate this form of asymmetry, by splitting the endogenous and instrumental variables into pairs of variables (one for upward revisions and another for downward revisions).

Table C.17 presents the regression results from the asymmetric specification. The top panel correspond to the original (symmetric) specification, which are identical to the baseline results reported in Table 2. The bottom panel corresponds to the asymmetric specification discussed above. Under the null hypothesis of symmetric effects, the coefficients η_{mgr}^{up} and η_{mgr}^{down} should be equal (in absolute value) – the bottom of the table reports the p-value of this test.

Regarding the vertical comparisons, we do not find any significant evidence for this type of asymmetry. We cannot reject the null hypothesis of symmetric effects in any of the seven tests reported in this table. This evidence indicates that the symmetric specification used in the baseline model constitutes a reasonable approximation. However, due to power limitations, we cannot rule out small or moderate asymmetries either. More precisely, the coefficients on downward revisions are quite imprecisely estimated (e.g., in column (1), the standard errors for the downward revisions are six times larger than for the upward revisions). This difference in precision comes from the fact that only a minority of employees end up making downward revisions.

For the horizontal comparisons, the baseline model makes two assumptions about symmetry. The first type of asymmetry is equivalent to the one discussed above for the case of vertical comparisons: the effects of updating beliefs upwards are the mirror image of updating beliefs downwards. Regarding the horizontal comparisons, Table C.17 shows that we reject the null hypothesis of symmetric effects in only one of the seven tests (for sales performance, with a p-value of 0.002). For this outcome, it seems like employees are more responsive to bad news (i.e., peers earning more than previously thought) than to good news. However, given the lack of consistency in the direction of the asymmetry across outcomes, this result is most likely spurious. Again, these findings suggest that the symmetric assumption from the baseline model is a reasonable approximation. However, the statistical power available to conduct this type of analysis is limited, so one should not conclude the effects are perfectly symmetric.

The second form of asymmetry that may arise in horizontal comparisons is the following: the effect of being 1 percent below the peer average salary is the mirror image of the effect

of being 1 percent above the peer average.⁷¹ Indeed, there is evidence from related studies on this type of asymmetry: while retention goes down when individuals are paid less than the average peer, retention does not go up as much when individuals are paid more than the average peer (Card et al., 2012; Dube et al., 2019; Breza et al., 2018). To allow for this type of asymmetry in horizontal comparisons, we can augment the baseline model given by equation (8) as follows:

$$\log(Y_i^{post}) = \eta_0 + \eta_{peer}^{below} \cdot \mathbb{1}_{P_i^{post} \leq O_i} \cdot \log(P_i^{post}) + \eta_{peer}^{above} \cdot \mathbb{1}_{P_i^{post} > O_i} \cdot \log(P_i^{post}) + \xi_i \quad (\text{C.5})$$

The variable O_i represents i 's own salary. While the baseline model had one parameter to represent the effects of peer salary (η_{peer}), this extension has two parameters: one parameter for when the own salary is below the average peer salary (η_{peer}^{below}) and another parameter for when the own salary is above the average peer salary (η_{peer}^{above}). For example, $\{\eta_{peer}^{below} < 0, \eta_{peer}^{above} = 0\}$ would suggest that employees care about the average peer salary but, once their own salary surpasses the peer average, they no longer care about it. We can expand the IV regression given by equations (9)–(11) to accommodate this form of asymmetry, by splitting the endogenous and instrumental variables for peer salary into pairs (one for when the salary is below the peer average and another for when the salary is above the peer average).

The results for this second form of asymmetry are presented in Table C.18. The top panel presents the results from the original specification, which are identical to the results from Table 2. The bottom panel corresponds to the asymmetric specification. In the null hypothesis of symmetric effects, the coefficients η_{peer}^{below} and η_{peer}^{above} are equal to each other – the bottom of the table reports the p-value of this difference test.

We start by discussing the effects on the retention outcome, which is the the form of behavior for which there is evidence of asymmetric responses (Card et al., 2012; Dube et al., 2019; Breza et al., 2018). These results are presented in column (4). For this outcome, we find a strong asymmetry and in the same direction reported in prior studies (Card et al., 2012; Dube et al., 2019; Breza et al., 2018). The coefficient on peer salary large (0.489) and highly significant when the own salary is below the peer salary, but close to zero (-0.099) and statistically insignificant when the own salary is above the peer salary. Most important, the difference between these two coefficients is statistically significant (p-value=0.047).

On the other hand, we do not find robust evidence for this type of asymmetry in the other outcomes. If anything, the point estimates are skewed in the opposite direction, with two of the differences being statistically insignificant (p-values 0.518 and 0.974) and one

⁷¹Since the perceived manager salary is always above the own salary, this form of asymmetry does not apply to vertical comparisons.

being borderline statistically significant (p -value=0.078). Once again, the statistical power available to conduct this type of analysis is limited, so one should not conclude the effects are perfectly symmetric. However, the evidence does indicate that the symmetry assumption from the baseline model is a reasonable approximation – with the exemption of the retention outcome, for which the symmetric model masks substantial asymmetry.

C.16 Potential Mechanisms Behind Horizontal Comparisons

In Section 4.5 we presented suggestive evidence that social preferences play a role in horizontal comparisons. However, the evidence does not rule out other explanations. In this section, we discuss and provide tests for some alternative channels.

One possibility is that employees reacted to the average salary of their peers because they used that information to learn about other aspects of their jobs. Being paid less than your peers may be a sign that you are less productive than your peers, and being paid more than your peers may be a sign that you are being more productive. Employees may also use the peer salary to infer what their managers think of them. Being paid less than your peers may signal that your manager thinks less of you, while being paid more than your peers may signal that your manager thinks highly of you. Under the right set of assumptions, these inferences could explain why employees work less hard when they receive feedback about a higher-than-expected peer salary.

We have two survey questions that can help us probe those mechanisms. The first question elicits the employee’s self-perception about his or her own relative productivity, using an incentivized method. At the end of every year, each employee is given a productivity rating on a 5-point scale. We elicited the individual’s perception about the share of employees who received each rating during the last yearly review. We incentivized this question by rewarding individuals for accurate responses. With these perceived shares and the employee’s own rating, we can infer the employee’s perceived productivity rank. This outcome, which is based on an incentivized question, can take values from 0 (least productive) to 1 (most productive). A higher perceived rank may reflect that the employee thinks more highly of herself. Additionally, since the manager provides key input in the performance review, a higher perceived rank may reflect that the manager thinks more highly of the employee. The second question elicits the employee’s probability of being promoted to the manager’s position within the next five years. The employee should expect a higher probability if she thinks more highly of herself, or if she thinks her manager thinks highly of her.

The effects of perceived salaries on these two survey outcomes are presented in Table C.19. This table uses the same IV specification from Table 3, only that it focuses on two different survey outcomes. Column (1) corresponds to the perceived productivity rank. The average

of this outcome, 0.47, indicates that individual’s perceptions about their productivity rank are accurate on average. In other words, employees do not seem to be systematically over- or under-confident. The coefficient on peer salary is close to zero (0.044), precisely estimated, and statistically insignificant. A 10% increase in peer salary, if anything, has a slight positive effect on the perceived productivity rank of just 0.44 pp. In turn, column (2) of Table C.19 presents the effects on the perceived probability of promotion. This dependent variable ranges from 1 (0%–10%) to 10 (90%–100%). The coefficient on peer salary is close to zero (-0.140), precisely estimated, and statistically insignificant. A 10% increase in peer salary reduces the perceived probability of promotion but by just 0.14 pp. In summary, the evidence from Table C.19 goes against the hypothesis that the individuals reacted to the information on peer salary because they inferred something about their productivity or the opinion of their managers.⁷²

The effect of peer salary could be the product of employees using the peer salary information to form beliefs about the salary that they could earn working for another firm. This mechanism provides a straightforward explanation for the effects on employee retention. Regarding the effects on effort, however, the connection is less clear. On the one hand, employees may work less hard because they expect to change firms and thus no longer care about their internal reputation. On the other hand, there are reason why employees could want to work harder. For example, employees may want to work harder to be in a better position to ask for a raise, to ask the firm to respond to an outside offer, or to obtain a positive recommendation from their current employer.

Another potential explanation for the effects of peer salary is that individuals use those perceptions to form beliefs about the returns to effort. This channel predicts heterogeneous effects: when receiving a signal that their peers are being paid more than expected, the least productive individuals should infer that the returns on effort are higher and should therefore work harder; on the contrary, the most productive individuals should infer that the returns on effort are lower and should then work less hard. We can provide a test of this channel using the survey data on perceived productivity rank. The results are presented in the last column of Table C.16, where we break down the effects of peer salary by whether the employee’s perceived productivity rank is below or above the median. This channel predicts that the effects of higher peer salary should be positive for individuals with below-median perceived productivity and negative for the rest. We do not find any evidence of this form of

⁷²For the sake of completeness, column (1) of Table C.19 also reports the coefficient on manager salary. This coefficient is also close to zero (0.000), precisely estimated, and statistically insignificant. A 10% increase in manager salary increases perceived productivity rank by less than 0.01 pp. This constitutes evidence that employees did not react to the manager salary because they inferred something about their productivity or their manager’s opinion. In column (2) we cannot estimate the effect of manager salary because the information on manager salary was provided after the elicitation of the perceived productivity rank.

heterogeneity: the effects of peer salary in these two groups are close to each other and their difference is statistically insignificant (p -value=0.788). This constitutes evidence that the effects of horizontal comparisons do not operate through beliefs about the returns to effort.

Last, it is possible that employees do not care about how much they are paid relative to their peers but are averse to salary inequality. Our information interventions were designed to shift beliefs about the relative standing rather than beliefs on the dispersion of salaries. If employees process the information rationally, then a signal about the average peer salary should have a small or no effect on the belief about the dispersion of salaries within the peer group (Hoff, 2009). As a result, it is unlikely that the effects of average peer salary operate through perceptions of peer inequality. In other words, while inequality aversion may be important, we would need a different experiment to estimate it.

C.17 Determinants of Horizontal Salary Differences

In this section, we discuss some evidence regarding how horizontal salary differentials are determined and discuss how these findings relate to the interpretation of the experimental results.

We start by providing a bit more detail about the institutional context. Several factors can influence whether one employee gets paid more than a peer. In a nutshell, the salary of a given employee is determined by the Human Resources division in conjunction with the employee's manager. For the sake of simplicity, we start with the case of new hires. The job description determines the pay grade for the position, which is set as part of the bank's overall strategy before a particular individual is assigned the job. This pay grade sets some bounds on the maximum and minimum pay for the position, but those bounds are wide and even overlap across different pay grades. As a result, the pay grade leaves a lot of leeway for horizontal salary differences. Within a given pay grade, the HR division recommends a salary based on market benchmarks that consider not only the position title but also some of the employee's characteristics such as his or her experience. However, that recommendation from HR is not written in stone: the manager has the option to override the HR recommendation. For example, a manager can recommend a salary above the market benchmark if the employee has a competing job offer. For the incumbent employees, the same factors are at play. For example, during the annual review a manager can recommend a higher raise for some employees based on the last year's performance review. And HR has policies that influence the salary growth – e.g., three years of top performance ratings translate automatically into a salary increase based on a fixed formula.

One specific factor that is of particular importance is that of meritocracy. One potential interpretation for the demoralizing effects of horizontal comparisons is that employees want

everyone in the same position to get paid the same salary regardless of any differences in effort or productivity. Another interpretation is that employees do not want to tolerate horizontal salary differences due to non-meritocratic factors such as luck or office politics. The institutional factors described above could support reasons why employee may see the horizontal pay differences as meritocratic (e.g., the raises based on performance reviews) or non-meritocratic (e.g., favoritism due to the manager’s discretion in setting pay). To provide a more quantitative assessment, below we provide quantitative evidence that, indeed, non-meritocratic factors may play a significant role in this context.

We want to assess whether employees end up with higher relative salaries due to meritocratic reasons (e.g., they are more productive or work harder). Let S_i be employee i ’s salary and \bar{S}_{-i} be the average salary among all of i ’s peers. The difference between S_i and \bar{S}_{-i} corresponds to the horizontal salary differential at the time of the experiment. The regression of interest is the following:

$$\log(S_i) - \log(\bar{S}_{-i}) = X_i\beta + \epsilon_i \tag{C.6}$$

Where X_i be a vector of characteristics for employee i such as demographics or measures of past productivity. As a first measure of prior productivity, we use the latest annual performance review. As explained above, the employee’s managers have discretion in determining the employee’s performance rating and the employee’s annual raises. At the end of every year, each employee is given a productivity rating on a 5-point scale. Moreover, the performance reviews and the raises are evaluated around the same time of the year. For these reasons, the performance rating is perhaps the measure of productivity that we should expect to have the strongest association with relative salaries.

The regression results are presented in Table C.20. Column (1) corresponds to a regression where the relative performance review rating is the only explanatory variable. Since we want to explain horizontal salary differences, we calculate the within-peer-group rank in this performance review. Thus, *Performance Rating* takes values from 0 to 1, where 0 would mean that the employee had the lowest performance review rating in the peer group while 1 would mean that the employee had the highest rating. The coefficient on *Performance Rating* is positive (0.052) and highly statistically significant (p-value<0.001). This coefficient is significant in magnitude too: climbing from the bottom to the top of the performance rank in the peer group would be associated with a 5.2% increase in the relative salary (equivalent to 0.33 standard deviations of the dependent variable). While economically significant, this effect is far from explaining all of the horizontal salary differences: the R^2 from column (1) suggests that the past performance ratings can only account for 0.9% of the horizontal differences in salaries.

As complementary evidence, columns (2) through (4) of Table C.20 introduces the three measures of effort and performance used as the dependent variables in the field experiment: the hours worked, the number of emails sent, and the sales performance. Since we are interested in measuring past performance, we construct the averages of these outcomes in the three months prior to the start of the experiment for each individual. Additionally, we use within-peer-group ranks to facilitate the comparison to the results for the performance review reported in column (1). As a result, the variables *Hours Worked*, *Email Sent* and *Sales* take values from 0 to 1, with 0 corresponding to the bottom of the distribution and 1 to the top of the distribution.

The results from columns (2) through (4) of Table C.20 are roughly consistent with the results from column (1). In column (2), the coefficient on *Hours Worked* is positive (0.016) but smaller in magnitude than the coefficient from column (1) and statistically insignificant (p-value=0.225). In column (3), the coefficient on *Emails Sent* is positive (0.055), similar in magnitude to the coefficient from column (1) and statistically significant (p-value=0.012). In column (4), the coefficient on *Sales* is positive (0.064), similar in magnitude to the coefficient from column (1) and statistically significant (p-value=0.011). On average, the R^2 in columns (2) through (4) is also consistent with the corresponding value from column (1).

In column (5) of Table C.20 we use the employee’s demographic characteristics to explain horizontal salary differences. The coefficient on *Female* indicates that female employees are paid 1.3% less than male employees with the same demographic characteristics. These differences can be taken as suggestive evidence that some non-meritocratic factors, such as discrimination, are responsible for some of the horizontal differences – for example, Cullen and Perez-Truglia (2019) provides evidence in this regard using data from this same organization. The variable $\log(\textit{Tenure})$ is meant to proxy for the employee’s experience at this specific firm while *Age* is meant to proxy for overall working experience. Both coefficients are positive and statistically significant, indicating that more experienced employees tend to have higher relative salaries. In the case of tenure, the effect could even be mechanical: even if they are not promoted, employees who continue working at this firm would still get raises every year, typically known as cost of living adjustments. To the extent that more experienced employees may be more valuable to the firm, these two coefficients could be taken as evidence of meritocratic pay. The last three variables, *College*, *Business Major* and *Finance Major* are related to the educational credentials of the employees. There is no evidence that, within a peer group, relative salaries are associated with these educational traits.

The R^2 from column (5) of Table C.20 indicates that, taken together, the demographic factors can explain 7.3% of the horizontal salary differences. Moreover, column (6) shows that even when taking the relative performance rating jointly with the demographic char-

acteristics, we can only explain 7.6% of the horizontal salary differences. In sum, the vast majority of the horizontal salary differences remain largely unexplained. Based on anecdotal accounts, those differences are probably due to sheer luck.⁷³ For example, your initial salary may be higher or lower depending on the representative from Human Resources who was in charge of hiring you, or whether the company was urgently looking to fill a new position at the time they hired you. Moreover, since employees have large misperceptions about peer salary, some employees may end up with lower salaries just because they happen to have more pessimistic beliefs at the time of salary negotiation. Due to the significant role that luck plays, it is plausible that employees perceive the horizontal differences as largely unfair.

C.18 Preferences for Transparency

In this section, we describe two survey questions that elicit more directly whether employees favored or opposed higher transparency. We explained that salaries are currently confidential information at the firm, and asked employees to report whether how they felt about alternative disclosure policies. In the first scenario, we propose the creation of a website showing the same type of information that we provided in our field experiment: i.e., the average salaries by position and unit. Employees could report their support or opposition to this new policy using the following scale: “strongly in favor,” “in favor,” “I would not care,” “against,” or “strongly against”. The results are presented in Figure C.11.a. A majority (65.26%) favors the policy, while 14.22% feel indifferent and only a minority (20.52%) opposes the policy. This survey evidence is consistent with our revealed-preference evidence based on the willingness to pay for salary information, according to which some employees have a lot to gain from having access to more information about salaries. The results are different under the second scenario, in which we offer employees to replace the status quo by a website that shows itemized information about salaries. In other words, you can use this website to look up any specific employee and find out how much they get paid, and other employees can look up your salary. The results, presented in Figure C.11.b indicate that there is little support for non-anonymous disclosure of information: a strong majority (74.83%) opposes the disclosure policy, while 11.84% feel indifferent and only a minority (13.33%) supports it. One plausible interpretation is that while employees value the salary information a lot, they may value their privacy even more.

⁷³These remaining salary differences may still be due to meritocratic factors that we cannot measure. For example, some employees may obtain higher salaries because they are able to attract outside offers (Caldwell and Harmon, 2018) and use them to negotiate raises, which some employees could see as fair. However, we do not have data on outside offers to account for this channel.

C.19 Average Effects from Transparency

In this section, we discuss the average effects of disclosing salary information.

Recall from Section 4.1 that Y_i^{post} is the outcome of interest (e.g., number of emails sent) and T_i^M and T_i^P are the binary variables indicating whether we disclosed information on manager salary and peer salary, respectively. The regression of interest is the following:

$$\log\left(Y_i^{post}\right) = \gamma_0 + \gamma_{mgr} \cdot T_i^M + \gamma_{peer} \cdot T_i^P + X_i \gamma_X + \epsilon_i \quad (\text{C.7})$$

The parameter of interests are γ_{mgr} and γ_{peer} , corresponding to the average effects of disclosing information about the manager salary and about the peer salary, respectively. The vector of additional control variables (X_i) is included to reduce the variance of the error term and thus improve the precision of the estimates. This corresponds to the same set of controls used in the analysis from Section 4.1: the employee’s own salary (in logs), tenure (in logs), dummies for performance evaluations in the previous year, and, following the standard practice in field experiments (McKenzie, 2012), the pre-treatment outcomes.

Before showing what the average effects of disclosure are, however, it helps to first describe how they relate to the parameters estimated in Section 4. We start by reproducing equation (8), which represents the relationship between salary perceptions and behavior:

$$\log\left(Y_i^{post}\right) = \eta_0 + \eta_{mgr} \cdot \log\left(M_i^{post}\right) + \eta_{peer} \cdot \log\left(P_i^{post}\right) \quad (\text{C.8})$$

In turn, the Bayesian learning model tells us how the disclosure of information affects each of those posterior beliefs:

$$\log\left(M_i^{post}\right) = \log\left(M_i^{prior}\right) + T_i^M \cdot \alpha_{mgr} \cdot \left(\log\left(M_i^{signal}\right) - \log\left(M_i^{prior}\right)\right) \quad (\text{C.9})$$

$$\log\left(P_i^{post}\right) = \log\left(P_i^{prior}\right) + T_i^P \cdot \alpha_{peer} \cdot \left(\log\left(P_i^{signal}\right) - \log\left(P_i^{prior}\right)\right) \quad (\text{C.10})$$

We can combine equations (C.8)–(C.10) to make a prediction about the average treatment effect of disclosing information:

$$\begin{aligned} \log\left(Y_i^{post}\right) = & \eta_0 + \eta_{mgr} \cdot \left(\log\left(M_i^{prior}\right) + T_i^M \cdot \alpha_{mgr} \cdot \left(\log\left(M_i^{signal}\right) - \log\left(M_i^{prior}\right)\right)\right) \\ & + \eta_{peer} \cdot \left(\log\left(P_i^{prior}\right) + T_i^P \cdot \alpha_{peer} \cdot \left(\log\left(P_i^{signal}\right) - \log\left(P_i^{prior}\right)\right)\right) + \nu_i \end{aligned} \quad (\text{C.11})$$

We define $\Delta_{T_i^M} \log\left(Y_i^{post}\right)$ as the effect of disclosing information about the manager salary to individual i on his or her own behavior. That is, the value under $T_i^M = 1$ minus the value

under $T_i^M = 0$:

$$\Delta_{T_i^M} \log(Y_i^{post}) = \eta_{mgr} \cdot \alpha_{mgr} \cdot (\log(M_i^{signal}) - \log(M_i^{prior})) \quad (\text{C.12})$$

Intuitively, the direction of the effect will depend on whether individual i was under- or over-estimating the salary of the manager – and the effect should be zero if the individual’s prior belief was the same as the signal. We can average over individuals to predict the average treatment effect of disclosing the manager salary:

$$\frac{1}{N} \sum_{i=1}^N \Delta_{T_i^M} \log(Y_i^{post}) = \eta_{mgr} \cdot \alpha_{mgr} \cdot \frac{1}{N} \sum_{i=1}^N (\log(M_i^{signal}) - \log(M_i^{prior})) \quad (\text{C.13})$$

In other words, the average treatment effect is the combination of the average bias in prior beliefs multiplied by the degree of belief updating (α_{mgr}) and the degree to which beliefs affect behavior (η_{mgr}). And we can reproduce the same exercise for the horizontal transparency:

$$\frac{1}{N} \sum_{i=1}^N \Delta_{T_i^P} \log(Y_i^{post}) = \eta_{peer} \cdot \alpha_{peer} \cdot \frac{1}{N} \sum_{i=1}^N (\log(P_i^{signal}) - \log(P_i^{prior})) \quad (\text{C.14})$$

We can use the estimates from Section 4 to predict the average treatment effects. The most straightforward prediction is for horizontal transparency: since the prior beliefs are accurate on average, we expect the average treatment effect to be zero. For the vertical transparency, however, we expect a positive average treatment effect, because on average employees underestimate the salary of their managers. However, that average treatment effect should be rather small in magnitude. Take for example the number of hours worked. We estimated an average bias in the prior belief of 0.139 log points, $\hat{\eta}_{mgr} = 0.150$ (column (1) from Table 2) and $\hat{\alpha}_{mgr} = 0.69$ (the slope from Figure 2.b). Thus, we predict that disclosing the manager salary should on average increase hours worked by 1.4% ($= 0.139 \cdot 0.15 \cdot 0.69$). One potential concern with these predictions, however, is that they rely on a number of assumptions such as linearity and symmetry. We can get a more direct measure of the average treatment effects by estimating equation (C.7) directly. And, as a validation exercise, we can compare the results from this simple regression to the predictions discussed above.

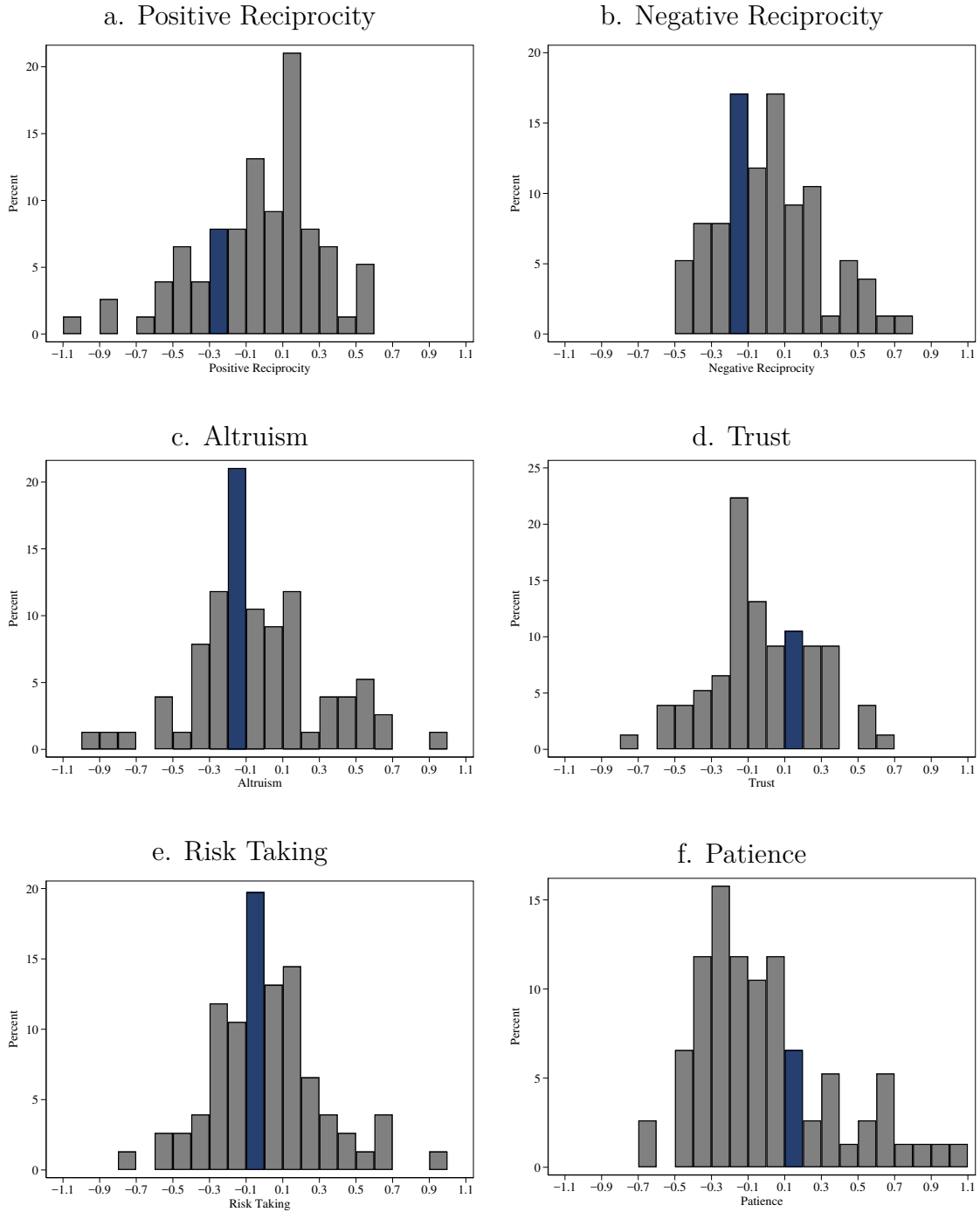
The estimates of equation (C.7) are presented in Table C.21. Different columns correspond to the different measures of effort and performance: the number of hours worked (column (1)), the number of emails sent (column (2)) and the sales performance (column (3)). And consistently with Table 2, in addition to the effects on the post-treatment outcomes, we present the (falsification) coefficients corresponding to the effects on the pre-treatment outcomes. For vertical transparency, columns (1)–(3) provide a consistent picture: on av-

erage, disclosing manager salary has small, positive effects on effort and performance. The effects are 4%, 3.3% and 2.6% for hours worked, number of emails and sales performance, respectively. However, due to lack of precision of the estimates, these coefficients are statistically insignificant. Most important, these estimates are consistent with the predictions discussed above. For example, we predicted that disclosure should increase hours worked by 1.4%, which is statistically indistinguishable from the corresponding effect reported above (4%, from column (1) of Table C.21). The results from Table C.21 are also consistent with the prediction regarding horizontal comparisons that, due to the unbiased beliefs, the average effects from disclosure should be null. Consistent with this prediction, across columns (1) through (3), we find that the average effects of disclosing peer salary are close to zero and statistically insignificant.

Additional References

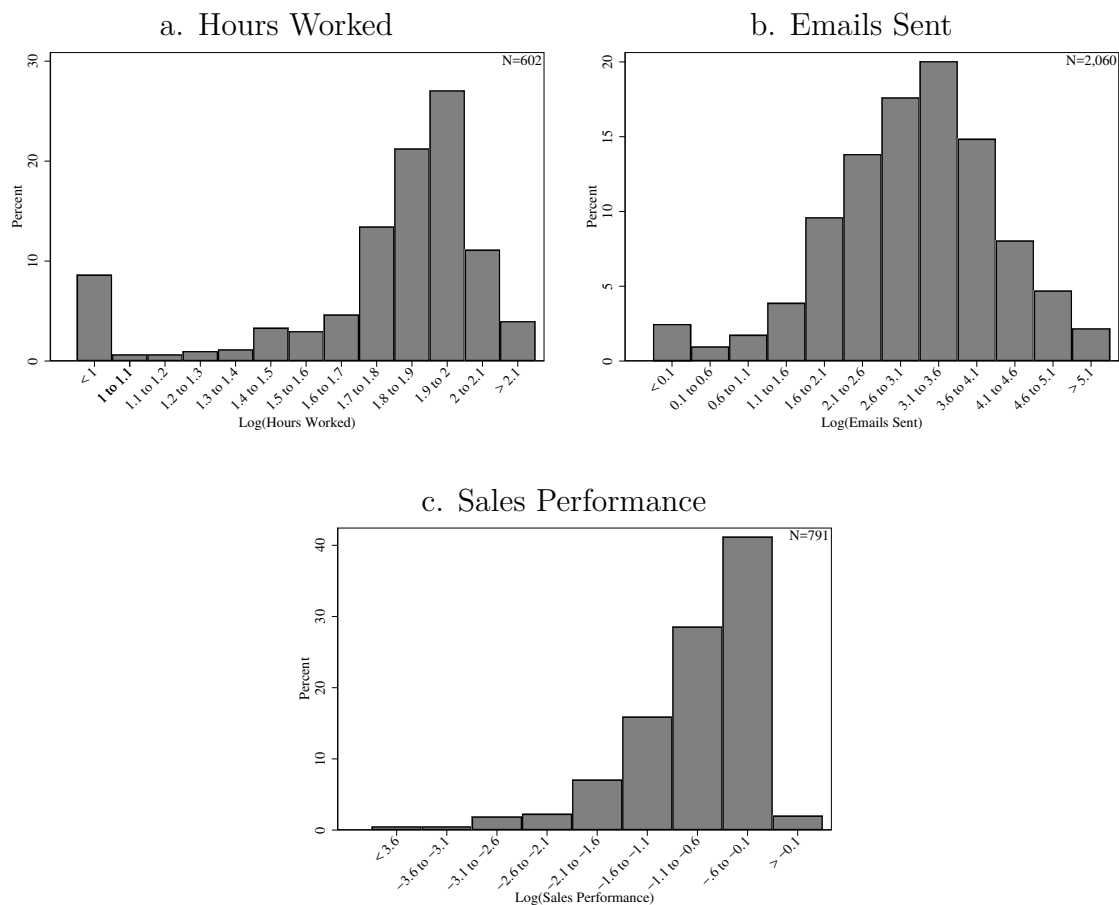
- Alatas, V., A. Banerjee, A. G. Chandrasekhar, R. Hanna, and B. A. Olken (2016). Network Structure and the Aggregation of Information: Theory and Evidence from Indonesia. *American Economic Review* 106 (7), 1663–1704.
- Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The Diffusion of Microfinance. *Science* 341 (6144).
- Brebner, S. and J. Sonnemans (2018). Does the elicitation method impact the WTA/WTP disparity? *Journal of Behavioral and Experimental Economics* 73(C), 40–45.
- Conlon, J., L. Pilossoph, M. Wiswall, and B. Zafar (2018). Labor Market Search With Imperfect Information and Learning. NBER Working Paper No. 24988.
- Luttmer, E. F. P. (2005). Neighbors as Negatives: Relative Earnings and Well-Being. *The Quarterly Journal of Economics* 120 (3), 963–1002.
- Mobius, M. and T. Rosenblat (2014). Social Learning in Economics. *Annual Review of Economics* 6 (1), 827–847.
- Stigler, G. J. (1962). Information in the Labor Market. *Journal of Political Economy* 70(5), 94–105.

Figure C.1: Country Context: Global Preferences Survey



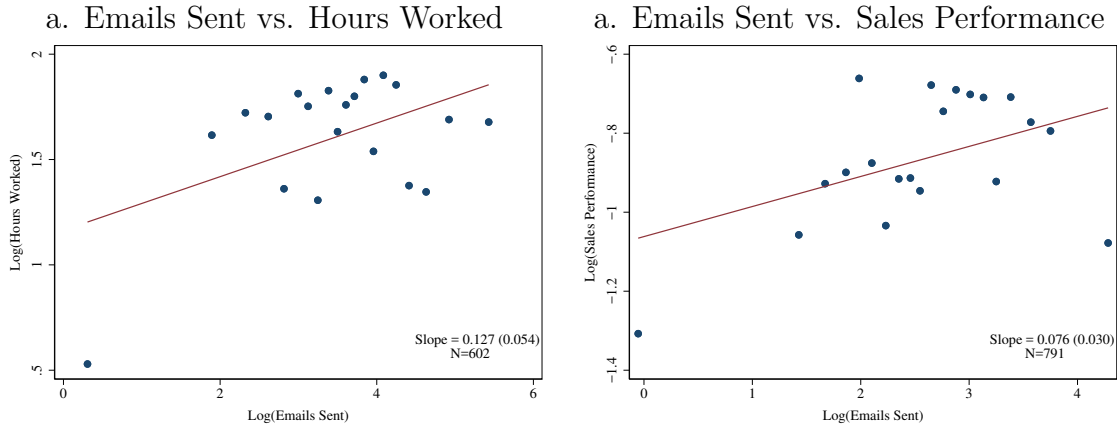
Notes: Distribution of the preference measures across 76 countries. The blue bar corresponds to the country where the firm is located. Data from the Global Preferences Survey (Falk et al., 2018).

Figure C.2: Distribution of Measure of Effort and Performance



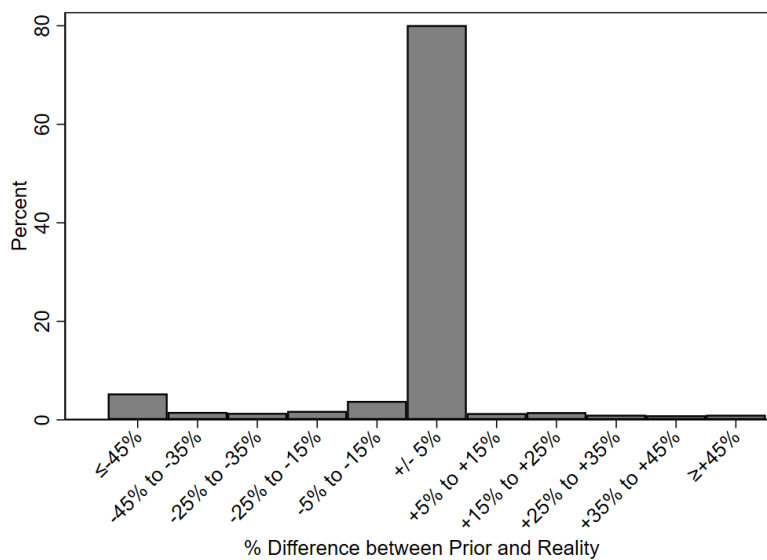
Notes: Histograms of the outcome variables used in columns (1) through (3) of Table 2. Panel (a) corresponds to the (log) mean number of hours worked per day over the 3 months post-treatment. Panel (b) corresponds to the (log) mean number of emails sent per day over the 3 months post-treatment. Panel (c) corresponds to the (log) mean sales performance index over the 3 months post-treatment.

Figure C.3: Association between Different Measures of Effort and Performance



Notes: Binned scatterplots of the association between the outcome variables used in columns (1) through (3) of Table 2. Panel (a) corresponds to the association between the (log) mean number of emails sent per day over the 3 months post-treatment and the (log) mean number of hours worked per day over the 3 months post-treatment. Panel (b) corresponds to the association between the (log) mean number of emails sent per day over the 3 months post-treatment and the (log) mean sales performance index over the 3 months post-treatment. We do not report the association between the hours worked and sales performance because there is almost no overlap between those two measures (see Table C.1). The slopes were estimated with linear regressions, with robust standard errors reported in parentheses.

Figure C.4: Training Module: Eliciting Own Salary



Notes: N=2,060. In the training module, we used an incentivized question to ask employees to report their own base salary. This graph shows the different between the employee's guess and the actual salary (according to the firm's administrative records), divided by the actual salary.

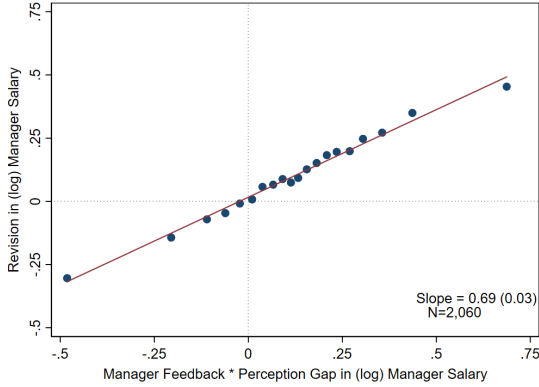
Figure C.5: Willingness to Pay for Salary Information: Comparison to Cullen and Perez-Truglia (2018)



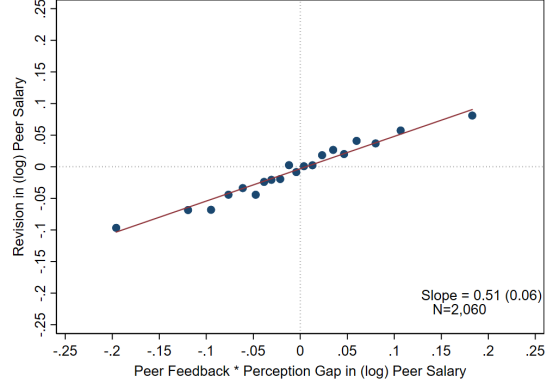
Notes: Panel (a) shows the distribution of the willingness to pay for information about the average salary among peers, using the multiple price-list menu method. The sample is restricted to the subset of respondents with consistent responses across the five price scenarios. Panel (b) is based on data from a follow-up study (Cullen and Perez-Truglia, 2018). It shows the distribution of the willingness to pay for information about the average salary among a sample of five peers, as measured by the respondent’s incentive-compatible bid using the open-ended method. Study participants are a non-overlapping representative sample from the same institution.

Figure C.6: Cross-Learning in the Information-Provision Experiment

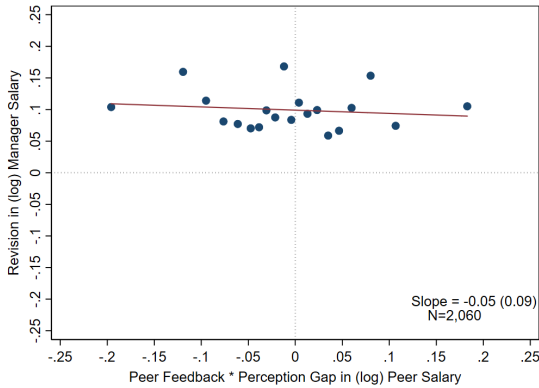
a. Effect of Manager Feedback on Belief About Managers



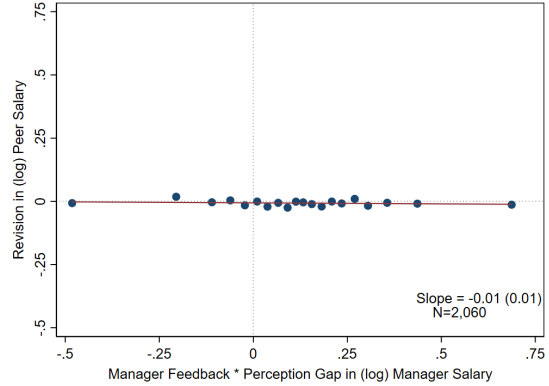
b. Effect of Peer Feedback on Belief About Peers



c. Effect of Peer Feedback on Belief About Managers

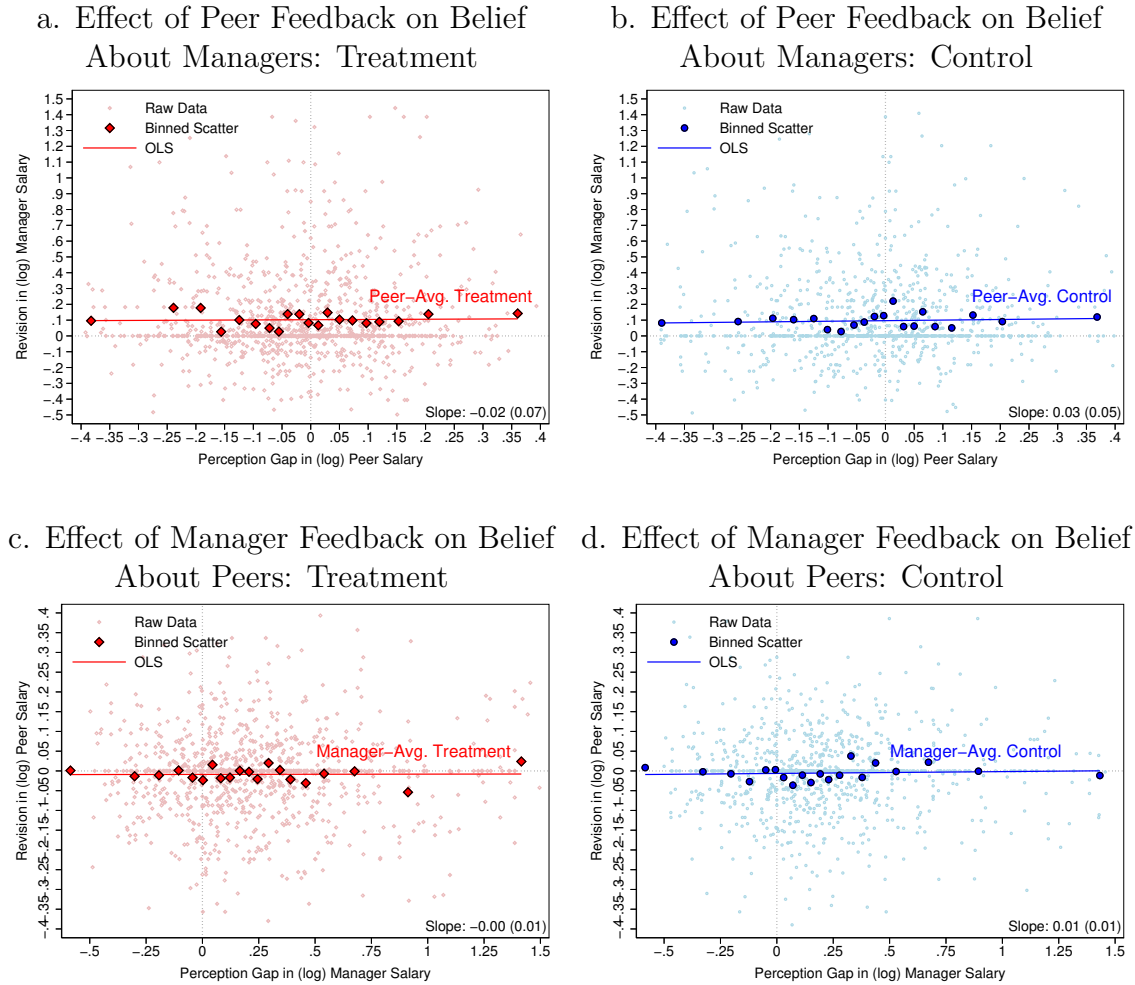


d. Effect of Manager Feedback on Belief About Peers



Notes: Panels (a) and (b) present partial regression binned scatterplot of the Bayesian learning equation (6) presented in Section 3.3. The y-axis corresponds to the respondent's update: i.e., the posterior belief minus the prior belief. The x-axis corresponds to the information treatment: the difference between the feedback chosen for the employee (e.g., the average salary among the random sample of 5 managers) and the employee's prior belief, multiplied by a binary variable for whether the information was randomly chosen to be shown to the respondent. The regression controls for the difference between the feedback chosen for the employee and the employee's prior belief; also, it controls for the prior belief and position title dummies. The slope was estimated with a linear regression, with standard errors (clustered at position level) reported in parentheses. Panel (a) shows how the feedback about manager salary affects posterior beliefs about manager salary. Panel (b) shows how the feedback about peer salary affects posterior beliefs about peer salary. Panel (c) shows how the feedback about peer salary affects beliefs about manager salary: it is identical to panel (b) except that the dependent variable is the update about manager salary. Panel (d) shows how the feedback about manager salary affects beliefs about peer salary: it is identical to panel (a) except that the dependent variable is the update about peer salary.

Figure C.7: Cross-Learning in the Information-Provision Experiment: Raw Scatterplots



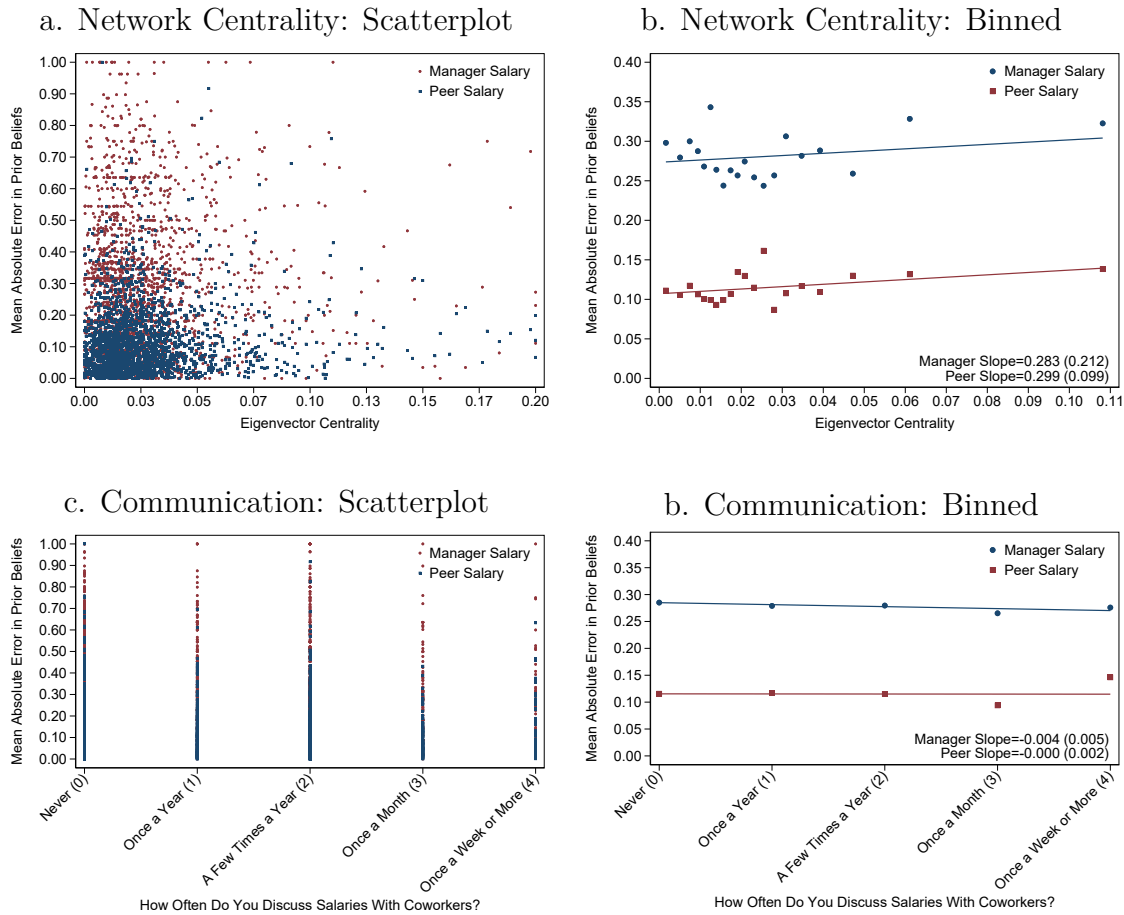
Notes: Panels (a) and (b) correspond to the effect of the information on peer salary on the beliefs about manager salary. The y-axis corresponds to the respondent’s update: i.e., the posterior belief about manager salary minus the corresponding prior belief. The x-axis corresponds to the gap between the feedback chosen for the employee (i.e., the average salary among the random sample of 5 managers) and the employee’s corresponding prior belief. Panel (a) shows the results for subjects in the treatment group (i.e., who received information about the peer salary) while panel (b) shows the results for subjects in the control group (i.e., who did not receive information about the peer salary). The raw data corresponds to a regular scatterplot, where each circle/diamond corresponds to a different respondent (for ease of exposition, we do not plot datapoints outside the range $[-0.5, 1.5]$). The larger circles/diamonds labeled binned scatter correspond to a binned scatterplot based on the same data. The slopes correspond to a linear regression, with standard errors clustered at the position level and presented in parentheses. Panels (c) and (d) correspond to the effect of the information on manager salary on the beliefs about peer salary. The y-axis corresponds to the respondent’s update: i.e., the posterior belief about peer salary minus the corresponding prior belief. The x-axis corresponds to the gap between the feedback chosen for the employee (i.e., the average salary among the random sample of 5 peers) and the employee’s corresponding prior belief. Panel (a) shows the results for subjects in the treatment group (i.e., who received information about the manager salary) while panel (b) shows the results for subjects in the control group (i.e., who did not receive information about the manager salary).

Figure C.8: Salary Misperceptions After Direct and Indirect Feedback



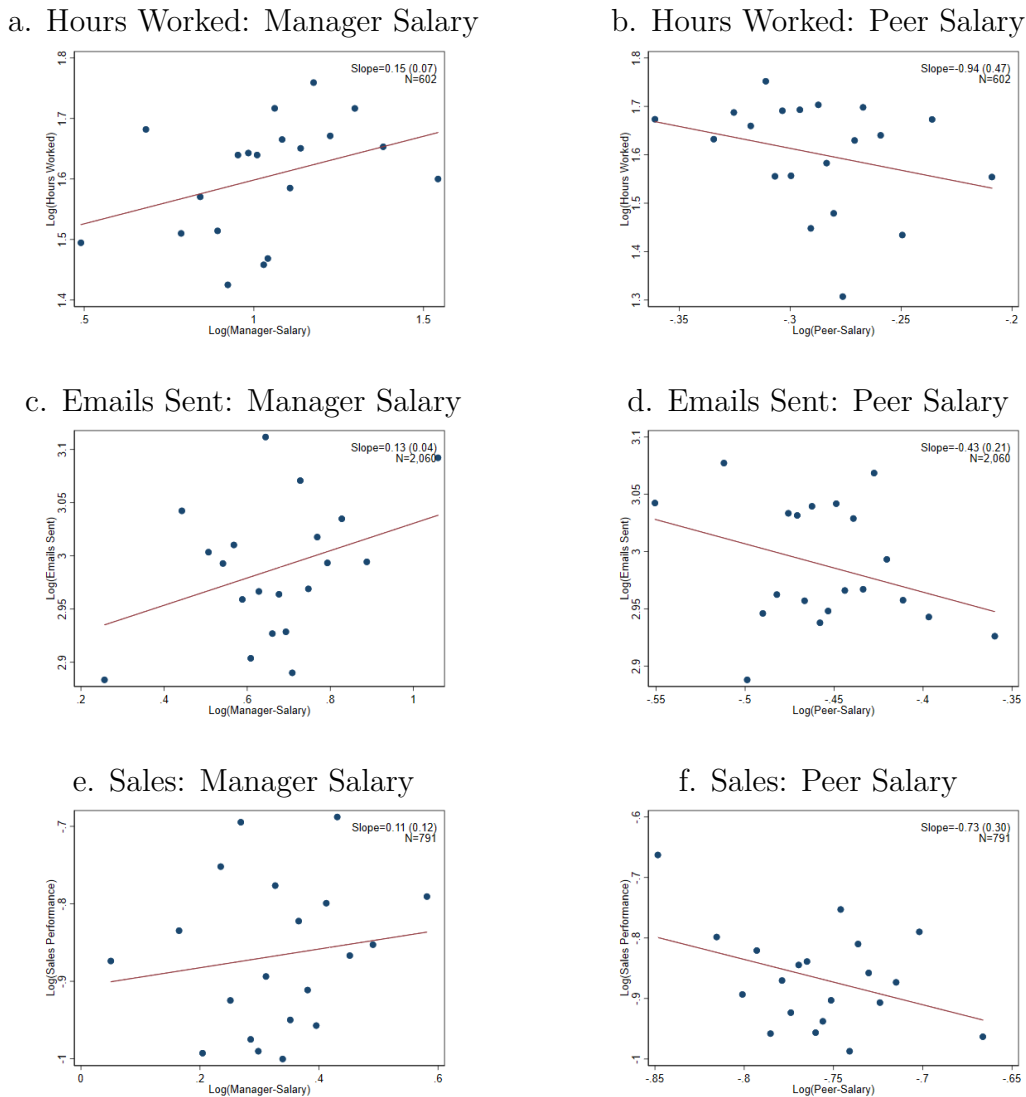
Notes: Panels (a) and (b) show the distribution of the difference between the employee’s posterior belief of the average salary (according to an incentivized survey question) and the actual average (according to the firm’s administrative records), divided by the actual salary. Panel (a) breaks down the respondents by whether they were randomly assigned to receive a signal about the average manager salary (*Direct Feedback*) or not (*No Direct Feedback*). Panel (b) is based on the subsample of respondents who did not receive direct feedback about the manager salary, and breaks them down by whether their closest peer received a signal about the average manager salary (*Indirect Feedback*) or not (*No Indirect Feedback*). Panels (c) and (d) are equivalent to panels (a) and (b) but about peer salary instead of manager salary. Panel (c) breaks down the respondents by whether they were randomly assigned to receive a signal about the average peer salary (*Direct Feedback*) or not (*No Direct Feedback*). Panel (d) is based on the subsample of respondents who did not receive direct feedback about the peer salary, and breaks them down by whether their closest peer received a signal about the average peer salary (*Indirect Feedback*) or not (*No Indirect Feedback*). In each panels, the p-value corresponds to the two-sample Kolmogorov-Smirnov test for equality of the two distributions shown in the corresponding panel.

Figure C.9: Non-Experimental Tests of Information Diffusion



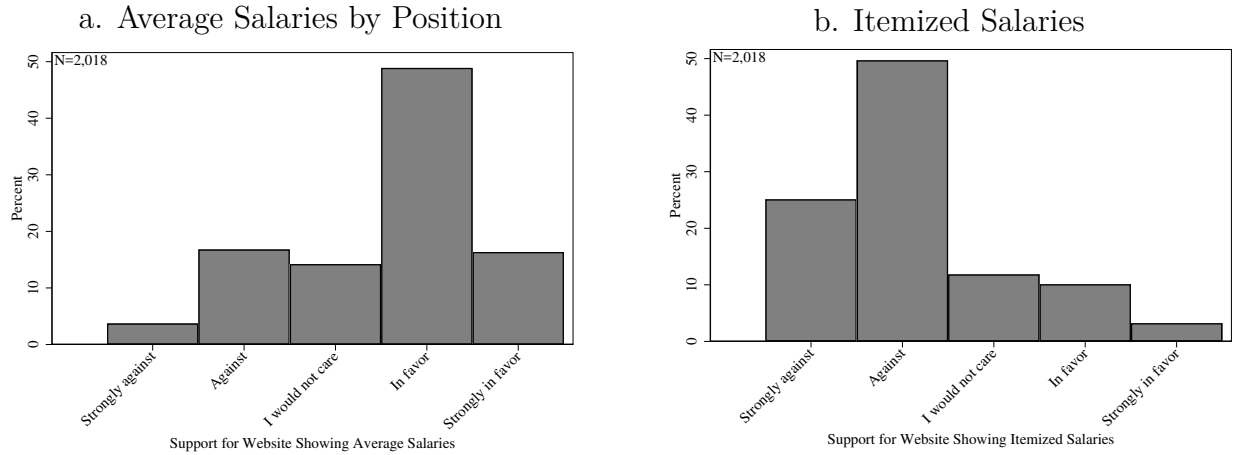
Notes: This figure presents the relationships between misperceptions and network centrality (panels (a) and (b)) and misperceptions and self-reported communication (panels (c) and (d)). In all panels, the y-axis corresponds to the respondent’s misperceptions: the prior belief of average (manager/peer) salary minus the true average, divided by the true average. In panels (a) and (b), the x-axis measures the employee’s network centrality, defined as their eigenvector centrality in the directed network of emails over the three months prior to the completion of the first survey and excluding emails directed outside of the institution and emails received from outsiders. In panels (c) and (d), the x-axis corresponds to the answer to the survey question “How often do you talk about salaries with coworkers?” Panels (a) and (c) present the raw scatterplots, where each point corresponds to a different subject. Panels (b) and (d) present binned scatterplots with linear fits, where the slopes are estimated with standard errors (clustered at position level) in parentheses.

Figure C.10: Binned Scatterplots of the Effects of Salary Perceptions on Effort and Performance



Notes: This figure presents binned scatterplots corresponding to the regressions reported in Table 2. Each row corresponds to a different IV regression. The dependent variable is the average behavior in the 90 days after the completion of the survey: *Hours* is the daily number of hours worked; *Emails* is the daily number of emails sent; and *Sales* is the sales performance index. The independent variables are: *Manager-Salary* (the posterior belief about the average manager salary) in the first column and *Peer-Salary* (the posterior belief about the average peer salary) in the second column. Each panel reports the corresponding coefficient from the IV regression, with standard errors (clustered at position level) in parentheses. For more details about the regression specification, see notes to Table 2.

Figure C.11: Preferences over Hypothetical Disclosure Policies



Notes: These histograms show the results of two questions included in our survey. Employees are reminded that salaries are confidential information at the firm. In panel (a), employees are asked whether they would support the replacement of the current transparency policy by a website showing the average salary by position/unit for all positions within the bank. In panel (b), employees are asked whether they would support the replacement of the current transparency policy by a website showing the list of names and salaries of all the employees, including your names and salaries.

Table C.1: Availability of Data on Emails Sent, Hours Worked and Sales Performance

Hours Data	Sales Data					
	No		Yes		Total	
	No.	%-Col.	No.	%-Col.	No.	%-Col.
No	721	56.8	737	93.2	1,458	70.8
Yes	548	43.2	54	6.8	602	29.2
Total	1,269	100.0	791	100.0	2,060	100.0

Notes: All 2,060 employees have data available on the number of emails sent. Among these, this table cross-tabulates the availability of data for hours worked and sales performance. The columns (*Sales Data*) indicates whether we observe a sales outcome during the 3 months post-treatment. The rows (*Hours Data*) indicates whether we observe an outcome for hours worked during the 3 months post-treatment.

Table C.2: Descriptive Statistics and Randomization Balance Test

	All	Treatment Group				P-value
	(1)	Manager (2)	Peer (3)	Both (4)	None (5)	
Female	0.73 (0.01)	0.74 (0.02)	0.71 (0.02)	0.74 (0.02)	0.74 (0.02)	0.32
Age	29.20 (0.11)	29.35 (0.22)	29.35 (0.21)	28.92 (0.19)	29.19 (0.22)	0.99
College (or Higher)	0.86 (0.01)	0.84 (0.02)	0.87 (0.01)	0.86 (0.01)	0.86 (0.02)	0.14
Tenure (Years)	4.99 (0.08)	5.14 (0.16)	5.08 (0.15)	4.92 (0.14)	4.79 (0.16)	0.81
Own Salary (Masked)	0.72 (0.01)	0.72 (0.02)	0.72 (0.02)	0.72 (0.03)	0.72 (0.02)	0.93
Avg. Manager Salary (Masked)	2.84 (0.05)	2.80 (0.10)	2.89 (0.10)	2.86 (0.10)	2.80 (0.11)	0.54
Avg. Peer Salary (Masked)	0.72 (0.01)	0.72 (0.02)	0.72 (0.02)	0.73 (0.02)	0.73 (0.02)	0.91
Observations	2,060	510	528	559	463	

Notes: Average pre-treatment characteristics of the employees, with standard errors in parentheses. *Female* takes the value 1 if the employee is female and 0 otherwise. *Age* is the employee's age (in years) as of March 2017. *College* takes the value 1 if the employee finished College or a higher degree, and 0 otherwise. *Tenure* is the number of years from the date when the employee joined the company until March 2017. *Own Salary* is the employee base monthly salary as of March 2017. *Avg. Manager Salary* and *Avg. Peer Salary* are the true average salaries among the manager and peer groups, respectively. Due to the sensitive nature of the data, we do not reveal the unit of measurement for salary variables. Column (1) corresponds to the entire subject pool, while columns (2) through (5) correspond to the four treatment groups that subjects were randomly assigned to: receiving information about the average manager salary only (column (2)); receiving information about the average peer salary only (column (3)); receiving information about both manager and peer salary (column (4)); and receiving no salary information (column (5)).

Table C.3: Average Characteristics in Subject Pool vs. Universe of Employees

	All	Invited		Responded	
	(1)	No (2)	Yes (3)	No (4)	Yes (5)
Female	0.71 (0.01)	0.65 (0.01)	0.76 (0.01)	0.79 (0.01)	0.73 (0.01)
Age	30.14 (0.06)	31.33 (0.10)	29.13 (0.08)	29.04 (0.12)	29.20 (0.11)
College (or Higher)	0.87 (0.00)	0.92 (0.00)	0.83 (0.01)	0.81 (0.01)	0.86 (0.01)
Tenure (Years)	5.09 (0.04)	5.32 (0.07)	4.90 (0.05)	4.80 (0.08)	4.99 (0.08)
Own Salary (Masked)	1.00 (0.01)	1.39 (0.03)	0.66 (0.01)	0.60 (0.01)	0.72 (0.01)
Avg. Manager Salary (Masked)	3.38 (0.04)	4.61 (0.08)	2.52 (0.03)	2.15 (0.04)	2.84 (0.05)
Avg. Peer Salary (Masked)	1.00 (0.01)	1.39 (0.03)	0.67 (0.01)	0.60 (0.01)	0.72 (0.01)
Observations	(Masked)	(Masked)	3,841	1,781	2,060

Notes: This table presents average pre-treatment characteristics of the employees, with standard errors in parentheses. *Female* takes the value 1 if the employee is female and 0 otherwise. *Age* is the employee's age (in years) as of March 2017. *College* takes the value 1 if the employee finished College or a higher degree, and 0 otherwise. *Tenure* is the number of years from the date when the employee joined the company until March 2017. *Own Salary* is the employee base monthly salary as of March 2017. *Avg. Manager Salary* and *Avg. Peer Salary* are the true average salaries among the manager and peer groups, respectively. Column (1) corresponds to the entire subject pool. Columns (2) and (3) split the universe of employees by whether they were invited (or not) to participate in the survey. Columns (4) and (5) split the sample of employees invited to the survey by whether they responded to the survey or not. Due to the sensitive nature of the data, we do not report the unit of measurement for the salary variables or the total number of employees in the organization.

Table C.4: Heterogeneity in Misperceptions and Willingness to Pay for Information

	Average Manager Salary			Average Peer Salary			Observations
	Error	Abs. Error	WTP	Error	Abs. Error	WTP	
All	-0.14 (0.008)	0.28 (0.005)	190.74 (6.652)	0.03 (0.003)	0.12 (0.002)	197.01 (6.506)	2,060
By Gender:							
Male	-0.14 (0.016)	0.29 (0.012)	208.03 (12.775)	0.02 (0.007)	0.12 (0.005)	216.40 (12.481)	554
Female	-0.14 (0.008)	0.28 (0.006)	183.77 (7.778)	0.03 (0.004)	0.11 (0.003)	189.22 (7.609)	1,506
Diff p-value	0.64	0.14	0.09	0.32	0.16	0.05	
By Tenure:							
> 4ys	-0.16 (0.010)	0.28 (0.007)	192.01 (9.349)	0.05 (0.005)	0.12 (0.004)	218.33 (9.478)	1,054
≤ 4ys	-0.12 (0.011)	0.28 (0.008)	189.40 (9.471)	0.00 (0.005)	0.11 (0.003)	174.79 (8.828)	1,006
Diff p-value	<0.01	0.86	0.84	<0.01	<0.01	<0.01	
By Payband:							
<i>HigherPayband</i>	-0.16 (0.011)	0.27 (0.008)	187.37 (9.873)	0.03 (0.006)	0.12 (0.004)	206.07 (9.788)	898
<i>LowerPayband</i>	-0.13 (0.011)	0.29 (0.007)	193.44 (9.003)	0.02 (0.004)	0.11 (0.003)	189.66 (8.704)	1,162
Diff p-value	0.01	0.25	0.65	0.38	<0.01	0.20	
By Sales Role:							
<i>Sales</i>	-0.09 (0.011)	0.25 (0.007)	203.12 (9.999)	0.04 (0.005)	0.11 (0.004)	190.83 (9.483)	972
<i>Non – Sales</i>	-0.19 (0.011)	0.31 (0.007)	180.06 (8.884)	0.02 (0.005)	0.12 (0.003)	202.31 (8.941)	1,088
Diff p-value	<0.01	<0.01	0.08	<0.01	0.75	0.37	
By Role:							
Front Office	-0.10 (0.008)	0.25 (0.006)	201.90 (8.139)	0.03 (0.004)	0.11 (0.003)	196.81 (7.780)	1,454
Back Office	-0.24 (0.015)	0.36 (0.011)	164.91 (11.381)	0.02 (0.007)	0.12 (0.005)	197.48 (11.875)	606
Diff p-value	<0.01	<0.01	0.01	0.22	0.02	0.96	

Notes: This table presents average error, absolute error, and WTP for various groups with standard errors in parentheses (clustered at the position level). *Error* is the gap between the perceived (manager/peer) average salary and the true average salary, divided by the true average salary. *Abs. Error* is the absolute value of *Error*. WTP is the willingness to pay for (manager/peer) information, calculated from interval data using the most conservative approach that focuses on the lower bound of each interval (see Section 3.2 for more details). The row “All” corresponds to the full sample and the other rows correspond to different subsamples. P-values correspond to the test of the null hypothesis that the average is equal across the two subgroups.

Table C.5: Determinants of Willingness to Pay for Salary Information

	WTP Peer Info			WTP Mgr Info	
	(1)	(2)	(3)	(4)	(5)
Log(Own Salary) - Log(Perceived Peer Avg. Salary)	-141.361*		-181.120*		
	(77.889)		(99.761)		
Log(Own Salary) - Log(True Peer Avg. Salary)		-16.166	94.952		
		(98.138)	(129.090)		
Perceived P(Promoted to Manager Position)				160.140***	160.403***
				(34.146)	(33.988)
Log(Own Salary)					12.542
					(32.003)
Observations	1,748	1,748	1,748	1,637	1,637

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column presents results for a separate interval regression. All regressions control for tenure (in logs) and dummies for performance evaluations in the previous year. In columns (1)–(3) the dependent variable is the willingness to pay for information on peer salary (restricted to the sample of respondents with consistent responses across the five scenarios). In columns (4)–(5) the dependent variable is the willingness to pay for information on manager salary (restricted to the sample of respondents with consistent responses across the five scenarios). *Log(Own Salary)* is the logarithm of the subject’s own salary. *Log(Perceived Peer Avg. Salary)* is the logarithm of the average peer salary according to the subject’s prior belief reported in the survey. *Log(True Peer Avg. Salary)* is the logarithm of the true average peer salary according to the administrative records. *Perceived P(Promoted to Manager Position)* is the subject’s perceived probability of being promoted to the managerial position according to survey data and can take values from 0.5 to 0.95 in increments of 0.05.

Table C.6: Heterogeneity in Learning Rates

	(1)	(2)	(3)	(4)	(5)
	Female	> 4ys	High-Band	Sales	Front-Office
α_{manager}	0.673*** (0.040)	0.630*** (0.056)	0.714*** (0.045)	0.633*** (0.030)	0.682*** (0.030)
α_{peer}	0.536*** (0.080)	0.516*** (0.093)	0.537*** (0.110)	0.568*** (0.069)	0.550*** (0.059)
Observations	1,506	1,054	898	972	1,454
	Male	$\leq 4ys$	Low-Band	Non-Sales	Back-Office
α_{manager}	0.748*** (0.058)	0.791*** (0.036)	0.677*** (0.049)	0.713*** (0.047)	0.697*** (0.057)
α_{peer}	0.456*** (0.054)	0.508*** (0.049)	0.478*** (0.033)	0.458*** (0.083)	0.415*** (0.128)
Observations	554	1,006	1,162	1,088	606
P-value Diff.:					
Manager	0.285	0.016	0.583	0.156	0.811
Peer	0.413	0.940	0.610	0.306	0.338

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Estimation of Bayesian learning equation (6) presented in Section 3.3 – for more details about the regression specification, such as outcomes, independent variables and data definitions, see the notes to Table 2. α_{mgr} and α_{peer} correspond to the learning rates for manager and peer beliefs (i.e., the weight that the individual assigns to the signal relative to the weight assigned to the prior belief), which are estimated from separate regressions. Each column corresponds to a different split of the sample (e.g., females vs. males), and the p-values correspond to the test of the null hypothesis that the learning rates are equal across the two subgroups.

Table C.7: Descriptive Statistics about Information Diffusion Analysis

	Manager Salary			Peer Salary		
	All	Received Own?		All	Received Own?	
		Yes	No		Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)
Misperceptions (Posterior Belief)	0.189 (0.201)	0.113 (0.139)	0.271 (0.225)	0.090 (0.111)	0.069 (0.102)	0.114 (0.115)
Information Assignment						
Direct	0.519 (0.500)	1.000 (0.000)	0.000 (0.000)	0.528 (0.499)	1.000 (0.000)	0.000 (0.000)
Indirect: Closest Peer	0.082 (0.274)	0.000 (0.000)	0.171 (0.376)	0.075 (0.263)	0.000 (0.000)	0.158 (0.365)
Indirect: No. Peers	1.346 (2.651)	0.000 (0.000)	2.797 (3.249)	1.313 (2.471)	0.000 (0.000)	2.779 (2.976)
Indirect: (No. Peers > 0)	0.340 (0.474)	0.000 (0.000)	0.707 (0.455)	0.352 (0.478)	0.000 (0.000)	0.745 (0.436)
Indirect: Share of Peers	0.060 (0.109)	0.000 (0.000)	0.125 (0.128)	0.061 (0.108)	0.000 (0.000)	0.129 (0.125)
Observations	2,060	1,069	991	2,060	1,087	973

Notes: This table presents average values of the key variables used for the information diffusion analysis, with standard deviations in parentheses. Columns (1)-(3) refers to the manager salary. *Misperceptions (Posterior Belief)* is the absolute value of the difference between the posterior belief about average manager salary and the true average, divided by the true average. *Direct* is a binary variable indicating whether the subject received the signal on manager salary directly. *Closest Peer*, *No. Peers*, *(No. Peers > 0)* and *Share of Peers* measure if the subject peer's received the signal on manager salary before the individual started his or her own survey. By definition, these variables take the value 0 if the employee received the information directly in his or her own survey. *Closest Peer* is a binary variable taking the value 1 if the individual's closest peer (defined as the peer with whom the employee exchanges the most number of emails in Jan-Mar 2017) received the information. *No. Peers* denotes the number of employees in the employee's peer group who received the information. *(No. Peers > 0)* is a binary variable indicating if *No. Peers* is positive. *Share of Peers* denotes the share of employees in the employee's peer group who received the information. Columns (4)-(6) show the corresponding results from columns (1)-(3) but for peer salary instead of manager salary.

Table C.8: Information Diffusion Analysis: Alternative Specification for Indirect Assignment

	Misperceptions on Manager Salary					Misperceptions on Peer Salary				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Information Assignment										
Direct	-0.160*** (0.011)	-0.160*** (0.011)	-0.160*** (0.011)	-0.160*** (0.011)	-0.160*** (0.011)	-0.045*** (0.005)	-0.045*** (0.005)	-0.045*** (0.005)	-0.045*** (0.005)	-0.045*** (0.005)
Indirect: Closest Peer		0.004 (0.013)					0.010 (0.008)			
Indirect: No. Peers			0.000 (0.002)					0.001 (0.001)		
Indirect: (No. Peers > 0)				0.012 (0.013)					0.008 (0.007)	
Indirect: Share of Peers					0.024 (0.048)					0.008 (0.020)
Constant	0.276*** (0.019)	0.276*** (0.019)	0.277*** (0.021)	0.275*** (0.019)	0.276*** (0.019)	0.111*** (0.011)	0.111*** (0.011)	0.114*** (0.011)	0.110*** (0.011)	0.111*** (0.011)
Observations	2,060	2,060	2,060	2,060	2,060	2,060	2,060	2,060	2,060	2,060

Notes: N= 2,060. Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the peer group level. This Table reproduces Table 1, only that the variables *Closest Peer*, *No. Peers*, *(No. Peers > 0)* and *Share of Peers* are not forced to take the value 0 if the employee received the information directly in his or her own survey. See the notes to Table 1 for more details.

Table C.9: Information Diffusion Analysis: Alternative Specification based on Prior Beliefs

	(Prior) Misperceptions on Manager Salary					(Prior) Misperceptions on Peer Salary				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Information Assignment										
Indirect: Closest Peer		0.006 (0.018)					0.005 (0.007)			
Indirect: No. Peers			-0.000 (0.003)					0.002* (0.001)		
Indirect: (No. Peers > 0)				0.015 (0.017)					0.002 (0.006)	
Indirect: Share of Peers					0.078 (0.061)					0.014 (0.021)
Constant	0.293*** (0.027)	0.293*** (0.027)	0.292*** (0.030)	0.291*** (0.027)	0.292*** (0.027)	0.090*** (0.009)	0.090*** (0.009)	0.093*** (0.009)	0.090*** (0.009)	0.089*** (0.009)
Observations	2,060	2,060	2,060	2,060	2,060	2,060	2,060	2,060	2,060	2,060

Notes: N= 2,060. Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the peer group level. This Table reproduces Table C.8, but using as dependent variable the misperceptions in prior beliefs instead of the misperceptions in posterior beliefs. we do not include the *Direct* binary variable because that information is always provided after the elicitation of prior beliefs. Also, the variables *Closest Peer*, *No. Peers*, *(No. Peers > 0)* and *Share of Peers* are not forced to take the value 0 if the employee received the information directly in his or her own survey. See the notes to Table 1 for more details.

Table C.10: Information Diffusion Analysis: Specification with Alternative Dependent Variables

	Bias on Manager Salary					Bias on Peer Salary				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Information Assignment										
Direct	0.105*** (0.014)	0.106*** (0.015)	0.108*** (0.018)	0.123*** (0.026)	0.108*** (0.019)	-0.004 (0.007)	-0.007 (0.007)	0.001 (0.009)	-0.005 (0.013)	-0.002 (0.009)
Indirect: Closest Peer		0.006 (0.026)					-0.020 (0.016)			
Indirect: No. Peers			0.001 (0.004)					0.002 (0.002)		
Indirect: (No. Peers > 0)				0.025 (0.030)					-0.001 (0.013)	
Indirect: Share of Peers					0.022 (0.104)					0.017 (0.044)
Constant	-0.088*** (0.033)	-0.088*** (0.033)	-0.088*** (0.033)	-0.098*** (0.035)	-0.089*** (0.034)	0.013 (0.014)	0.015 (0.014)	0.012 (0.014)	0.014 (0.015)	0.012 (0.014)
Observations	2,060	2,060	2,060	2,060	2,060	2,060	2,060	2,060	2,060	2,060

Notes: N= 2,060. Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the peer group level. This table is identical to Table 1 except that the dependent variables are not constructed in absolute values: e.g., in columns (1) through (5) the dependent variable is the difference between the posterior belief of the average manager salary and the true average, divided by the true average. See the notes to Table 1 for more details.

Table C.11: Reduced Form and First Stage Results from the Instrumental Variables Estimator

	Effort and Performance		
	(1) log(<i>Hours</i>)	(2) log(<i>Emails</i>)	(3) log(<i>Sales</i>)
2SLS:			
Log(Manager-Salary)	0.150** (0.074)	0.130*** (0.041)	0.106 (0.122)
Log(Peer-Salary)	-0.943** (0.472)	-0.431** (0.210)	-0.731** (0.297)
Reduced Form:			
(Log(Manager Signal)-Log(Manager Prior))*I(Manager Treatment)	0.103* (0.053)	0.094*** (0.029)	0.094 (0.066)
(Log(Peer Signal)-Log(Peer Prior))*I(Peer Treatment)	-0.379** (0.160)	-0.211* (0.112)	-0.381* (0.195)
First Stage Log(Manager-Salary):			
(Log(Manager Signal)-Log(Manager Prior))*I(Manager Treatment)	0.705*** (0.066)	0.692*** (0.035)	0.628*** (0.029)
(Log(Peer Signal)-Log(Peer Prior))*I(Peer Treatment)	0.018 (0.133)	0.073 (0.050)	0.125** (0.054)
First Stage Log(Peer-Salary):			
(Log(Manager Signal)-Log(Manager Prior))*I(Manager Treatment)	0.002 (0.011)	-0.008 (0.008)	-0.038*** (0.009)
(Log(Peer Signal)-Log(Peer Prior))*I(Peer Treatment)	0.404*** (0.115)	0.511*** (0.059)	0.538*** (0.043)
Observations	602	2,060	791

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. The panel *2SLS* reproduces the regression results from Table 2 – see its note for details about the variables and the specification. The remaining three panels reproduce the reduced form and first stage results from those Two-Stage Least Squares (2SLS) regressions. *Log(Manager-Salary)* and *Log(Peer-Salary)* correspond to the posterior beliefs about manager and peer average salaries, respectively. *Log(Manager Prior)* and *Log(Peer Prior)* correspond to the prior beliefs about manager and peer average salaries, respectively. *Log(Manager Signal)* and *Log(Peer Signal)* correspond to the signals about manager and peer average salaries, respectively. *I(Manager Treatment)* is a binary variable that takes the value 1 if the signal about manager salary was shown to the respondent and 0 otherwise. *I(Peer Treatment)* is a binary variable that takes the value 1 if the signal about peer salary was shown to the respondent and 0 otherwise.

Table C.12: Effects of Salary Perceptions on Effort and Performance: Alternative Specification

	Baseline Specification (logs)			Alternative Specification (%- Δ)		
	Hours (1)	Emails (2)	Sales (3)	Hours (4)	Emails (5)	Sales (6)
Post-Treatment (3-Months):						
<i>Log (Manager-Salary)⁽ⁱ⁾</i>	0.150** (0.074)	0.130*** (0.041)	0.106 (0.122)	0.523 (0.368)	0.465 (0.299)	0.552*** (0.210)
<i>Log (Peer-Salary)⁽ⁱⁱ⁾</i>	-0.943** (0.472)	-0.431** (0.210)	-0.731** (0.297)	-3.210 (2.256)	-1.278** (0.591)	-1.992*** (0.395)
P-value H_0 : (i)=(ii)	0.026	0.007	0.000	0.141	0.016	0.000
Cragg-Donald F-Stat.	29.8	204.0	98.2	29.3	203.7	98.1
Observations	602	2,060	791	602	2,060	791

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column corresponds to a different IV regression. Columns (1) through (3) are a reproduction of the baseline results from the top panel of Table 2 – for more details about the regression specification, such as outcomes, independent variables and data definitions, see the notes to that table. The specifications in columns (4) through (6) are identical to columns (1) through (3) except that, instead of using the logarithm transformation, the dependent variable are defined as the ratio between the post-treatment outcome and the pre-treatment outcome.

Table C.13: Effects of Salary Perceptions on Effort and Performance: Additional Controls

	log(<i>Hours</i>) (1)	log(<i>Emails</i>) (2)	log(<i>Sales</i>) (3)	log(<i>Hours</i>) (4)	log(<i>Emails</i>) (5)	log(<i>Sales</i>) (6)	log(<i>Hours</i>) (7)	log(<i>Emails</i>) (8)	log(<i>Sales</i>) (9)	log(<i>Hours</i>) (10)	log(<i>Emails</i>) (11)	log(<i>Sales</i>) (12)
Post-Treatment (3-Months):												
<i>Log</i> (Manager-Salary) ⁽ⁱ⁾	0.150** (0.074)	0.130*** (0.041)	0.106 (0.122)	0.152** (0.075)	0.130*** (0.041)	0.106 (0.126)	0.153** (0.073)	0.119*** (0.040)	0.116 (0.137)	0.139 (0.089)	0.146*** (0.047)	0.067 (0.133)
<i>Log</i> (Peer-Salary) ⁽ⁱⁱ⁾	-0.943** (0.472)	-0.431** (0.210)	-0.731** (0.297)	-1.010** (0.508)	-0.434** (0.210)	-0.720** (0.301)	-0.894* (0.536)	-0.414* (0.247)	-0.495* (0.286)	-0.329 (0.453)	-0.456** (0.213)	-0.872*** (0.289)
Gender Dummies				✓	✓	✓						
Location Dummies							✓	✓	✓			
Position Dummies										✓	✓	✓
P-value H_0 : (i)=(ii)	0.026	0.007	0.000	0.028	0.007	0.000	0.057	0.028	0.001	0.338	0.003	0.000
Cragg-Donald F-Stat.	29.8	204.0	98.2	29.6	204.0	98.3	24.4	198.2	94.5	41.7	253.2	92.8
Observations	602	2,060	791	602	2,060	791	602	2,060	791	602	2,060	791

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column corresponds to a different IV regression. Columns (1) through (3) are a reproduction of the baseline results from the top panel of Table 2 – for more details about the regression specification, such as outcomes, independent variables and data definitions, see the notes to that table. The specifications in columns (4)–(6) are identical to columns (1)–(3) except for the inclusion of an additional control variable: a binary variable indicating if the employee is female. Columns (7)–(9) are identical to columns (1)–(3) except for the inclusion of the following additional control variables: a set of 29 indicator variables corresponding to the location where the employee works. Columns (10)–(12) are identical to columns (1)–(3) except for the inclusion of the following additional control variables: a set of 180 indicator variables corresponding to the employee’s position title.

Table C.14: Effects of Salary Perceptions on Effort and Performance: Additional Controls

	$\log(\text{Hours})$	$\log(\text{Emails})$	$\log(\text{Sales})$	$\log(\text{Hours})$	$\log(\text{Emails})$	$\log(\text{Sales})$	$\log(\text{Hours})$	$\log(\text{Emails})$	$\log(\text{Sales})$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post-Treatment (3-Months):									
<i>Log (Manager-Salary)⁽ⁱ⁾</i>	0.150** (0.074)	0.130*** (0.041)	0.106 (0.122)	0.152** (0.077)	0.129*** (0.040)	0.086 (0.095)	0.137* (0.074)	0.123*** (0.042)	0.128 (0.150)
<i>Log (Peer-Salary)⁽ⁱⁱ⁾</i>	-0.943** (0.472)	-0.431** (0.210)	-0.731** (0.297)	-1.171** (0.496)	-0.499** (0.201)	-0.555* (0.306)	-1.105** (0.558)	-0.434** (0.214)	-0.729** (0.330)
Prior Gaps Dummies				✓	✓	✓			
Posterior Certainty							✓	✓	✓
P-value H_0 : (i)=(ii)	0.026	0.007	0.000	0.010	0.002	0.013	0.031	0.008	0.001
Cragg-Donald F-Stat.	29.8	204.0	98.2	31.0	196.0	95.4	26.7	198.7	97.8
Observations	602	2,060	791	602	2,060	791	597	2,035	782

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column corresponds to a different IV regression. Columns (1) through (3) are a reproduction of the baseline results from the top panel of Table 2 – for more details about the regression specification, such as outcomes, independent variables and data definitions, see the notes to that table. The specifications in columns (4)–(6) are identical to columns (1)–(3) except for the addition of the following control variables: sets of dummies indicating the decile of the prior gap in beliefs (nine dummies for manager salary and nine dummies for peer salary). Columns (7)–(9) are identical to columns (1)–(3) except for the addition of eight control variables (four variables for the manager salary and four variables for the peer salary). For the manager salary, the four variables are: the perceived probability that the true average manager salary falls between -10% and -2.5% of the posterior belief of the average manager salary; the perceived probability that the true salary falls between -2.5% and +2.5% of the posterior belief; the perceived probability that the true salary falls between +2.5% and +10% of the posterior belief; the perceived probability that the true salary falls above +10% of the posterior belief. And the four variables on peer salary are defined in an analogous way to the four variables on manager salary.

Table C.15: Effects of Perceived Manager and Peer Salary on Various Email Outcomes

	By Direction		Sent to		Sent to Pay Band		
	Sent (1)	Received (2)	Inside (3)	Outside (4)	Higher (5)	Same (6)	Lower (7)
<i>Log</i> (Manager-Salary)	0.130*** (0.041)	0.041 (0.035)	0.139*** (0.042)	0.047 (0.072)	0.249*** (0.086)	0.123* (0.065)	-0.047 (0.145)
<i>Log</i> (Peer-Salary)	-0.431** (0.210)	-0.198** (0.092)	-0.400* (0.207)	-0.602 (0.394)	0.046 (0.320)	-0.061 (0.276)	-1.076 (0.666)
P-value Diff.							
Manager		0.096		0.268		0.161	
Peer		0.310		0.650		0.464	
Observations	2,060	2,060	2,060	2,060	2,060	2,060	2,060

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column corresponds to a different IV regression. We always use the baseline specification from Table 2 – for more details about this specification see the notes to that table. Each column uses a different dependent variable: the average number of emails sent in the 90 days following the survey (column (1)), the number of emails received (column (2)), the number of emails sent to email accounts inside the firm (column (3)), emails sent outside the firm (column (4)), emails sent to employees with a higher (column (5)), same (column (6)) and lower (column (7)) paybands.

Table C.16: Heterogeneity in Effects of Perceptions on Number of Emails Sent

	(1)	(2)	(3)	(4)	(5)	(6)
	Female	> 4ys	High-Band	Sales	Front-Office	High Rank
$\text{Log}(\text{Manager-Salary})^{(i)}$	0.155*** (0.052)	0.075 (0.073)	0.129* (0.069)	0.196** (0.083)	0.153** (0.077)	0.118 (0.088)
$\text{Log}(\text{Peer-Salary})^{(ii)}$	-0.448** (0.224)	-0.215 (0.312)	-0.380 (0.277)	-0.548* (0.283)	-0.422* (0.231)	-0.321 (0.366)
Observations	1,506	1,054	898	972	1,454	750
	Male	$\leq 4ys$	Low-Band	Non-Sales	Back-Office	Low Rank
$\text{Log}(\text{Manager-Salary})^{(iii)}$	0.066 (0.092)	0.205*** (0.074)	0.127* (0.071)	0.130** (0.053)	0.094 (0.065)	0.149** (0.058)
$\text{Log}(\text{Peer-Salary})^{(iv)}$	-0.724 (0.707)	-0.616 (0.474)	-0.463 (0.342)	-0.135 (0.387)	-0.558 (0.613)	-0.443 (0.270)
Observations	554	1,006	1,162	1,088	606	1,310
P-value $H_0 : (i)=(iii)$	0.710	0.479	0.851	0.389	0.835	0.788
P-value $H_0 : (ii)=(iv)$	0.394	0.213	0.991	0.505	0.559	0.774

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column corresponds to a different IV regression. We always use the baseline specification from Table 2 – for more details about this specification see the notes to that table. Each column corresponds to a different split of the sample (e.g., females vs. males), and the p-values correspond to the test of the null hypothesis that the learning rates are equal across the two subgroups. In column (6), there is missing data on perceived productivity rank for 52 observations, which we impute using the employee’s true position in the productivity distribution.

Table C.17: Asymmetry of Effects of Perceptions about Manager and Peer Salary: Upwards vs. Downwards Revisions

	Effort and Performance			Career Moves			
	log(<i>Hours</i>) (1)	log(<i>Emails</i>) (2)	log(<i>Sales</i>) (3)	P(Left) (4)	P(Transfer) (5)	log(<i>Salary</i>) (6)	<i>P</i> (Δ <i>Title</i>) (7)
Symmetric Model:							
<i>Log</i> (Manager-Salary)	0.150** (0.074)	0.130*** (0.041)	0.106 (0.122)	-0.016 (0.023)	-0.003 (0.030)	0.002 (0.011)	0.012 (0.029)
<i>Log</i> (Peer-Salary)	-0.943** (0.472)	-0.431** (0.210)	-0.731** (0.297)	0.232** (0.106)	0.093 (0.106)	0.004 (0.052)	0.114 (0.123)
Asymmetric model:							
<i>Log</i> (Manager-Salary) Upwards ⁽ⁱ⁾	0.095 (0.078)	0.154*** (0.054)	0.141 (0.120)	-0.026 (0.025)	-0.025 (0.036)	-0.006 (0.014)	0.018 (0.029)
Downwards ⁽ⁱⁱ⁾	0.805 (0.542)	0.008 (0.255)	0.016 (0.227)	0.057 (0.075)	0.138 (0.101)	0.065 (0.048)	-0.042 (0.107)
<i>Log</i> (Peer-Salary) Upwards ⁽ⁱⁱⁱ⁾	0.051 (0.887)	-0.969** (0.480)	0.081 (0.257)	0.254* (0.154)	0.316 (0.245)	-0.064 (0.086)	0.194 (0.230)
Downwards ^(iv)	-1.930** (0.798)	0.094 (0.373)	-1.426*** (0.420)	0.228 (0.189)	-0.107 (0.131)	0.092 (0.061)	0.030 (0.081)
P-value Test: H_0 : (i)=(ii)	0.144	0.153	0.002	0.921	0.179	0.164	0.482
P-value Test: H_0 : (iii)=(iv)	0.204	0.612	0.621	0.313	0.152	0.186	0.580
Mean Outcome	5.98	35.57	0.48	0.05	0.09	0.92	0.10
Std. Dev. Outcome	1.88	44.93	0.23	0.21	0.28	0.70	0.30
Observations	602	2,060	791	2,060	2,060	2,060	2,060

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column corresponds to a different pair of IV regressions. In “Symmetric Model” we use the baseline specification from Table 2 – for more details about the regression specification, such as outcomes, independent variables and data definitions, see the notes to that table. In “Asymmetric Model”, we use the specification described in Appendix C.15, which differs from the baseline specification in that it allows the effects of *Peer Salary* and *Manager Salary* to be different depending on whether the individual revised her prior beliefs upwards or downwards.

Table C.18: Asymmetry of Effects of Perceptions about Peer Salary: Above vs. Below Own Salary

	Effort and Performance			Career Moves			
	$\log(\text{Hours})$ (1)	$\log(\text{Emails})$ (2)	$\log(\text{Sales})$ (3)	P(Left) (4)	P(Transfer) (5)	$\log(\text{Salary})$ (6)	$P(\Delta\text{Title})$ (7)
Symmetric Model:							
<i>Log</i> (Peer-Salary)	-0.943** (0.472)	-0.431** (0.210)	-0.731** (0.297)	0.232** (0.106)	0.093 (0.106)	0.004 (0.052)	0.114 (0.123)
Asymmetric Model:							
<i>Log</i> (Peer-Salary)							
Above Own-Salary ⁽ⁱ⁾	-0.448 (0.600)	-0.402 (0.341)	-0.304 (0.276)	0.489*** (0.184)	-0.062 (0.162)	0.011 (0.077)	0.083 (0.140)
Below Own-Salary ⁽ⁱⁱ⁾	-2.357 (2.582)	-0.421 (0.392)	-1.114*** (0.388)	-0.099 (0.179)	0.291 (0.290)	0.023 (0.068)	0.151 (0.191)
P-value Test: $H_0: (i)=(ii)$	0.518	0.974	0.078	0.047	0.383	0.911	0.750
Mean Outcome	5.98	35.57	0.48	0.05	0.09	0.92	0.10
Std. Dev. Outcome	1.88	44.93	0.23	0.21	0.28	0.70	0.30
Observations	602	2,060	791	2,060	2,060	2,060	2,060

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column corresponds to a different pair of IV regressions. In “Symmetric Model” we use the baseline specification from Table 2 – for more details about the regression specification, such as outcomes, independent variables and data definitions, see the notes to that table. In “Asymmetric Model”, we use the specification described in Appendix C.15, which differs from the baseline specification in that it allows the effect of *Peer Salary* to be different depending on whether *Peer Salary* is below or above *Own-Salary*.

Table C.19: Effects of Salary Perceptions on Additional Survey Outcomes

	(1)	(2)
	Prod.	P(Prom.)
<i>Log</i> (Manager-Salary) ⁽ⁱ⁾	0.000 (0.015)	
<i>Log</i> (Peer-Salary) ⁽ⁱⁱ⁾	0.044 (0.040)	-0.014 (0.118)
P-Value (i)=(ii)	0.280	
Cragg-Donald F-Stat.	250.5	512.1
Mean Dep. Var.	0.47	0.56
Std. Dev. Dep. Var.	0.22	0.29
Observations	1,999	2,051

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column presents results for a different IV regressions, following the same specification from Table 3 – see the corresponding notes for more details. *Manager-Salary* is the posterior belief about manager salary, and *Peer-Salary* is the posterior belief about the average peer salary. *Prod.* denotes the individual self-perceived position in the distribution of performance ratings in the firm and was elicited after the elicitation of the posterior beliefs. *P(Prom.)* is the perceived probability of being promoted to the managerial position in the next five years. Since this belief was elicited prior to the experimental provision of information about manager salary, we do not include *Manager-Salary* as independent variable.

Table C.20: Determinants of Horizontal Variation in Salaries

	Dep. Var.: $\log(\text{Salary}) - \log(\text{Average Peer Salary})$					
	(1)	(2)	(3)	(4)	(5)	(6)
Within-Peer-Group Rank in:						
Performance Rating	0.052*** (0.007)					0.035*** (0.010)
Hours Worked		0.016 (0.013)				
Emails Sent			0.055** (0.022)			
Sales				0.064** (0.023)		
Demographics:						
Female					-0.013** (0.006)	-0.013** (0.006)
Age					0.005*** (0.001)	0.005*** (0.001)
$\log(\text{Tenure})$					0.024*** (0.007)	0.023*** (0.007)
College					-0.001 (0.007)	-0.003 (0.007)
Business Major					0.002 (0.008)	0.003 (0.008)
Finance Major					0.008 (0.007)	0.009 (0.006)
Mean Outcome	-0.011	-0.009	-0.013	-0.010	-0.011	-0.011
Std. Dev. Outcome	0.157	0.141	0.159	0.144	0.157	0.157
R^2	0.009	0.001	0.011	0.017	0.073	0.076
Observations	(Masked)	1,765	4,587	1,597	(Masked)	(Masked)

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column presents results for a separate OLS regression. The dependent variable is the log-difference between the individual's own base salary and the average salary in the peer group. *Performance Rating*, *Hours Worked*, *Emails Sent* and *Sales* correspond to the individual's rankings in their peer groups and thus range from 0 (bottom of the group) to 1 (top of the group). *Performance Rating* corresponds to the ranking of performance review rating in 2016. *Hours worked* correspond to the ranking in the number of hours worked in the three months prior to the experiment. *Emails Sent* correspond to the ranking in the number of emails sent in the three months prior to the experiment. *Sales* correspond to the ranking in the sales performance index during the three months prior to the experiment. *Female* is an indicator variable that takes the value 1 if the employee is female. *Age* is the age in years. $\log(\text{Tenure})$ is the (log) of the number of months since the employee joined the firm. *College* takes the value 1 if the employee has a College degree. *Business Major* takes the value 1 if the employee's College degree is a business major. *Finance Major* takes the value 1 if the employee's College degree is a finance major. Columns (1), (5) and (6) includes all employees working at the firm at the start of the experiment. Columns (2) through (4) correspond to the subsample of employees with non-missing data in *Hours Worked*, *Emails Sent* and *Sales*, respectively.

Table C.21: Average Effects of Information Disclosure

	Effort and Performance		
	(1) log(<i>Hours</i>)	(2) log(<i>Emails</i>)	(3) log(<i>Sales</i>)
Post-Treatment:			
Shown Manager Info ⁽ⁱ⁾	0.040 (0.039)	0.033 (0.026)	0.026 (0.032)
Shown Peer Info ⁽ⁱⁱ⁾	0.044 (0.056)	-0.020 (0.036)	-0.009 (0.021)
Pre-Treatment (Falsification):			
Shown Manager Info	-0.012 (0.043)	0.014 (0.024)	0.032 (0.039)
Shown Peer Info	0.010 (0.045)	-0.006 (0.030)	-0.029 (0.032)
P-value H_0 : (i)=(ii)	0.959	0.231	0.426
Mean Outcome	5.98	35.57	0.48
Std. Dev. Outcome	1.88	44.93	0.23
Observations	602	2,060	791

Notes: Significant at *10%, **5%, ***1%. Standard errors in parentheses clustered at the position level. Each column presents results for two sets of OLS regressions, each based on equation (C.7) from Appendix C.19: in *Post-Treatment*, the dependent variable is the average behavior 90 days after the completion of the survey; in *Pre-Treatment (Falsification)*, the dependent variable is the average behavior before the completion of the survey. *Shown Manager Info* is a binary variable indicating whether the subject was randomly chosen to receive information about the average manager salary. *Shown Peer Info* is a binary variable indicating whether the subject was randomly chosen to receive information about the average peer salary. The regressions control for three monthly lags of the dependent variable, (log) own salary, (log) tenure, and five productivity rating dummies. *Hours* is the daily number of hours worked. *Emails* is the daily number of emails sent. *Sales* is the sales performance index. The mean and std. dev. reported in the bottom rows correspond to the values prior to taking the logarithm function. Columns (1) corresponds to the subsample of employees in the headquarter offices and column (3) to the subsample of employees with sales roles.