Does employee happiness have an impact on productivity?

Monograph:

Bellet, Clement, De Neve, Jan-Emmanuel and Ward, George (2020) Does employee happiness have an impact on productivity? CEP Discussion Papers. Centre for Economic Performance, LSE, London, UK.

ISSN 2042-2695

CENTRE *for* ECONOMIC PERFORMANCE

CEP Discussion Paper No 1655

Revised February 2020
(Replaced October 2019 version)

Does Employee Happiness Have an Impact on Productivity?

Clément S. Bellet
Jan-Emmanuel De Neve
George Ward

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

ESRC ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
This article provides quasi-experimental evidence on the relationship between employee happiness and productivity in the field. We study the universe of call center sales workers at British Telecom (BT), one of the United Kingdom's largest private employers. We measure their happiness over a 6 month period using a novel weekly survey instrument, and link these reports with highly detailed administrative data on workplace behaviors and various measures of employee performance. Exploiting exogenous variation in employee happiness arising from weather shocks local to each of the 11 call centers, we document a strong causal effect of worker happiness on sales. This is driven by employees working more effectively on the intensive margin by making more calls per hour, adhering more closely to their workflow schedule, and converting more calls into sales when they are happier. In our restrictive setting, we find no effects on the extensive margin of happiness on various measures of high-frequency labor supply such as attendance and break-taking.

Key words: happiness, productivity
JEL Codes: D03; J24; M5; I31

This paper was produced as part of the Centre's Wellbeing Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

We thank Alex Bryson, Ed Diener, Alex Edmans, Paul Frijters, Sergei Guriev, John Helliwell, Erin Kelly, Tom Kochan, Richard Layard, Armando Meier, Robert Metcalfe, Michael Norton, Paul Osterman, Alexandra Roulet, Mark Stabile, Andrew Stephen, John Van Reenen, and Ashley Whillans for helpful comments on manuscript drafts. We also thank seminar participants at Oxford, MIT, London School of Economics, Erasmus University Rotterdam and Paris School of Economics for helpful discussions. We are grateful to British Telecom for the opportunity to design and run this study and for providing access to their administrative data. We also thank Butterfly AI for their technical support in implementing the employee survey

Clément S. Bellet, Erasmus University and Centre for Economic Performance, London School of Economics. Jan-Emmanuel De Neve, University of Oxford and Centre for Economic Performance, London School of Economics. George Ward, MIT and Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

What explains the large differences in labor productivity that are typically observed between firms and individuals? Although answers to this question have historically focused in economics on monetary aspects of motivation (e.g. Lazear, 2000), a recent stream of research has begun to broaden the focus to include non-monetary aspects of work like intrinsic motivation, social preferences, fairness concerns, meaningfulness of tasks, and social comparisons (e.g. Benabou and Tirole, 2003; Cassar and Meier, 2018; DellaVigna and Pope, 2017; Gneezy and Rustichini, 2000; Ichniowski and Shaw, 2003). In this paper, we add to this growing literature by studying the productivity effects of a typically-overlooked aspect of workers' lives: their day-to-day happiness.

A recent trend has seen a growing number of employers at least claiming to care about the happiness of their employees, and beginning to invest in management practices and services aimed at creating and maintaining a happy workforce. At least one reason for this is the expectation that happier workers will be more productive in their jobs.[1] However, it remains unclear whether this belief is based on sound empirical evidence or is, rather, a case of "management mythology".

We build on a long history of work, largely outside of economics, on employee well-being and performance. A large initial stream of research examined the association between measures of self-reported job satisfaction and worker performance, typically in field settings, and has generally found a modest positive (partial) correlation between the two (see Judge et al., 2001, for a review). However, there is a general lack of causal research on the issue, and the best existing evidence relies principally on within-worker estimates. A more recent stream of research—arising largely out of cognitive psychology—has moved away from evaluative assessments of employee satisfaction, in order to study employees' affective or emotional states in the workplace (Brief and Weiss, 2002; Tenney et al., 2016). One benefit of studying emotional states is that they are more readily manipulable in laboratory settings, allowing for more clearly causal research designs. A recent pioneering paper in economics, for example, influences happiness in a laboratory setting to show a robust causal effect on a stylized, piece-rate productivity task (Oswald et al., 2015).[2] However, the extent to which these effects translate into real-world employment settings remains an open question.

In this paper, we present evidence of a causal effect of workers' week-to-week happiness on their productivity, in a field context at one of the United Kingdom's largest private employers. We study the universe of around 1,800 sales workers at British Telecom (BT), a large telecommunications firm in the UK. Employees are distributed across 11 call centers, where their job is predominantly to take incoming calls from new and existing customers, and seek to sell them various products such as broadband internet contracts, cell and landline phone deals, and television packages. We observe the happiness of these workers on a weekly basis over a six month period, and match this psychological data with detailed administrative data on a number of

[1] Other motivations may also include a desire to attract and retain high quality workers, the enhancement of firms' reputation among customers, as well as concerns to do with corporate social responsibility.

[2] For earlier work in psychology showing the effects of positive mood on task performance in a non-incentivized setting, see Erez and Isen (2002).

workplace behaviors and performance outcomes.

Whereas much of the prior literature has been forced to rely on subjective outcomes—such as managerial performance evaluations—the benefit of studying a call center population is that we are able to study a range of objective, quantitative performance metrics. Even so, despite this abundance of data, it is often unclear what constitutes a positive outcome for the firm. For example, one leading contender is the number of calls per hour (or, analogously, the average duration of calls). But it is far from clear whether making many faster calls is any better than making fewer longer calls. As our principal measure of performance we focus on sales, which is an unambiguously positive performance outcome for the firm and is a measure that combines call quantity with call quality.[3] In order to investigate channels through which any effect of happiness may translate into sales, we then also study high-frequency measures of labor supply like attendance, sickness absence and break-taking, as well as fine-grained measures of labor productivity such as minutes per call, the percentage of calls converted to sales and the extent to which workers "adhere" to their scheduled workflow.

The use of high-frequency panel data allows us to estimate productivity equations with individual fixed effects, thus purging from our estimates any unobserved heterogeneity between workers. Nevertheless, there remain reasons to believe our within-worker estimates could still be biased by time-varying third factors, reverse causality, and/or measurement error. Given this, we study adverse local weather shocks in order to set up a natural experiment wherein we are able to estimate the causal effect of happiness on productivity. We instrument for an employee's weekly happiness using adverse weather conditions close to each call center. Poor weather has a strong ("first-stage") negative effect on worker happiness, which we show ultimately leads to fewer weekly sales in two-stage IV models. Depending on the specification, we find that a one standard deviation increase in reported happiness within-workers leads to around an 18-24% increase in sales, which is equal to around 4 to 6 additional sales per week. We provide a detailed discussion on the magnitude of the effect. In particular, we show that our causal estimate is consistent with prior evidence from laboratory experiments (Oswald et al., 2015), and with the measured productivity impact of management policies that simultaneously impacted employee happiness and productivity in randomized field experiments (e.g. Bloom et al., 2014).

Our key identifying assumption is that weather patterns have an effect on sales only through variation in worker happiness. We explore in detail at least three principal threats to this assumption. First, there is a possibility that weather will also affect product demand, either directly or indirectly through the mood of customers. It is worth noting here that demand is national, and that calls are allocated to call centers based on operator availability and the type of enquiry the call is based on.[4] We estimate restrictive models that include both worker and week fixed effects, such that our key piece of identifying variation in the data comes from differences across call centers (which are very widely dispersed geographically across 3 countries, from the

[3] One could even argue it is the only relevant outcome, as assuming customer satisfaction remains constant (which we can control for in our analysis), the ultimate goal of a sales worker is the number of sales she makes. In that sense, call duration is only a means to an end.

[4] This is not an entirely random process, since if a customer wants to buy, say, an internet package rather than a TV package, she may be directed to a different call center. However, despite not being entirely random, the process should nevertheless be orthogonal to either weather patterns local to each call center and to individual operator happiness.

south coast of England to the north of Scotland) within any given week, and not from movements in national weather from week-to-week that could be correlated with demand. Second, poor weather could influence commuting patterns, or induce sickness among workers causing them to attend less. Finally, weather patterns could affect the opportunity cost of leisure compared to work (see Connolly, 2008), and have effects on attendance. One of the key benefits of studying workers in a call center setting is that employees enjoy very little discretion over their day-to-day work hours, making it an ideal population in which to cleanly measure labor productivity effects (holding labor supply constant).

In line with the restrictiveness of the setting, when examining channels we find null effects at the extensive margin on measures such as attendance, sickness leave, overtime, or length of paid and unpaid breaks.[5] Instead, leveraging detailed data on labor productivity, we show that the effects of happiness on sales are driven principally at the intensive margin by employees both working more efficiently (i.e. making more calls per hour and adhering more closely to their workflow) and also converting more of their calls into sales during weeks when they feel happier. The latter channel suggests that the effects of happiness are likely to be strongest in service industry settings, where work is directly customer-facing. Happiness may in this sense be beneficial to people's social skills, the importance of which have been shown to be increasing in the labor market (Deming, 2017). Underscoring the importance of happiness as an aid to social skills, we show that the effects are greater for upgrade and re-contracting sales—where there is more leeway for inter-personal connection, persuasion, and negotiation—than they are for more routine order-taking. We also discuss the contribution of alternative psychological mechanisms such as intrinsic motivation and cognitive processing, and provide some suggestive evidence on each.

The findings contribute to various strands of research across economics, psychology, sociology, and related disciplines. First, as noted above, the analysis relates directly to a large and long-standing literature in economics on the determinants of productivity (Syverson, 2011). Second, the results contribute to a large body of work with a long history of examining the relationship between employee well-being and performance. In studying well-being in a field setting, our work is closely related to research on job satisfaction and performance. But in studying a measure of positive affect, our findings are also closely related to a largely more laboratory-based literature on emotional states in determining work-related outcomes (see Walsh et al., 2018, and citations within for an extensive review of the evidence arising from both streams of work). In a working paper close in nature to our own, Coviello et al. (2018) study the psychological state of predominantly non-sales call center workers in the USA and find a negative relationship with performance.[6] Our study differs from this in a number of ways. Rather than happiness, the authors study a 1-to-5 mood scale corresponding to feeling 'unstoppable', 'good', 'so so', 'exhausted', and 'frustrated'. In addition, the outcome measure studied is very different to our focus on sales, in that the authors show a negative relationship between mood and the

[5]We do not interpret this as strong evidence that happiness could not have an effect on labor supply elsewhere, rather that in our setting very few things could affect working hours.

[6]Relatedly, Rothbard and Wilk (2011) study the mood and performance of a sample of 29 call center workers in an insurance firm, and present within-worker estimates of a positive relationship between affect and task performance.

speed of work (using the number of calls per hour and the average duration of calls as their main outcome measures). Although the authors also exploit variation in local weather, the lack of geographical dispersion in the call centers they study does not allow them to leverage weather variation within any given time period.

Third, our findings provide micro-foundation for work showing that firms listed in the "100 Best Companies To Work For In America" outperform industry benchmarks in terms of long-run stock returns (Edmans, 2011), as well as for research suggesting a positive link between firm-level employee satisfaction and financial performance (Böckerman and Ilmakunnas, 2012; Bryson et al., 2017; Krekel et al., 2019). Fourth, our findings are in line with an experimental literature at the individual level, largely in psychology, showing in non-incentivized laboratory settings that induced happiness has effects on processes like motivation, cognitive flexibility, negotiation, and problem solving skills (see Isen, 2001, for a review).

Fifth, the results contribute to work in economics that implies the existence of mood effects on workplace behaviors, but is unable to measure them directly. Heyes and Saberian (2019) show, for example, that temperature has significant effects on the decision-making of judges in high-stakes court cases, while work by Eren and Mocan (2018) suggests that judges' decisions can be swayed by college football results. Equally, behavior in stock markets is affected by sunshine and daylight hours (Hirshleifer and Shumway, 2003; Kamstra et al., 2003) and national soccer results (Edmans et al., 2007). While none of these papers is able to measure mood directly, all explicitly tell a story in which emotional states causally affect workplace behaviors. Our analysis contributes to this literature by more directly measuring the effects of happiness and unhappiness on performance.

Finally, our study also ties in to an emerging field-experimental literature showing that the improvement of management practices can have simultaneously positive impacts on i) employee happiness and satisfaction as well ii) as productivity. Gosnell et al. (2020) show, for example, that the introduction of various management practices improved both the performance and well-being of airline captains. Equally, Bloom et al. (2014) show that allowing call center employees to work from home improves their productivity while also enhancing happiness and satisfaction, and an intervention by Breza et al. (2017) shows that pay inequality has negative effects on morale while also depressing productivity.[7] Each of these papers suggests that employee happiness may be one channel through which workplace organization feeds through to productivity. In this paper, we isolate this channel more clearly in order to show a positive causal relationship between happiness and employee performance.

2 Data and Institutional Setting

We collect data from British Telecom (BT), a large telecommunications company based in the United Kingdom. We focus our analysis on the firm's 11 call centers located across England, Scotland and Wales that have a large number of workers concentrating predominantly on sales. We use administrative data obtained from the firm on worker characteristics, work schedules

[7] Similarly, Moen et al. (2017) show that improving worker autonomy has a positive effect on well-being and also decreases turnover (see also Moen et al., 2016). The authors are unable, however, to observe productivity in their IT firm setting.

and productivity, and combine these data with an original survey instrument we implemented to measure workers' psychological well-being over a six month period.

2.1 Psychological Well-Being Survey

The use of "happiness" data in economics has grown substantially over the past decade (see, e.g, Aghion et al., 2016; Benjamin et al., 2012; Haushofer and Shapiro, 2016; Kahneman et al., 2004; Luttmer, 2005). Subjective well-being (SWB) is typically divided into two separate components that measure i) how people feel as well as ii) how they think about their lives (Kahneman and Krueger, 2006).[8] *Evaluative* measures of psychological well-being typically comprise answers to global, cognitive judgements like the extent to which people are satisfied with their job or their life overall. *Affective* measures of well-being – which are often also referred to as 'hedonic' or 'experienced' measures – refer, on the other hand, to the extent to which people experience positive and negative emotional states (like happiness, enjoyment, stress and worry) in the course of their day-to-day lives.

Whereas much of the existing literature on well-being in the workplace focuses on evaluative measures of job satisfaction, which are typically slow-moving overall cognitive assessments that lend themselves to between-worker comparisons, we focus instead on employees' experienced well-being as it varies week-to-week. In doing so, we build on a small experimental literature that manipulates positive and negative affect in the laboratory in order to examine effects on economic outcomes like productivity (Oswald et al., 2015), time preferences (Ifcher and Zarghamee, 2011), and behavior in ultimatum and trust games (Capra, 2004). For ease of exposition, we use the terms 'happiness', 'positive mood', 'affect', 'affective state', 'emotional well-being' and 'emotional state' interchangeably throughout the paper.[9]

The survey instrument was designed with the OECD'S guidelines on the measurement of SWB in mind (see OECD, 2013).[10] Employees were asked "*Overall, how happy did you feel this week?*".[11] Following Kunin (1955) and decades of subsequent work in the psychological literature, we offer five response categories as a Faces Scale that ranges from very sad to very happy androgynous faces.[12] The use of faces in this way is both intuitive to respondents, and is also known to strongly pick up the affective component of well-being questions (Fisher, 2000).

[8] A third component of SWB is increasingly also identified, which comprises meaningfulness or "eudaemonia" (see, e.g., Ariely et al., 2008).

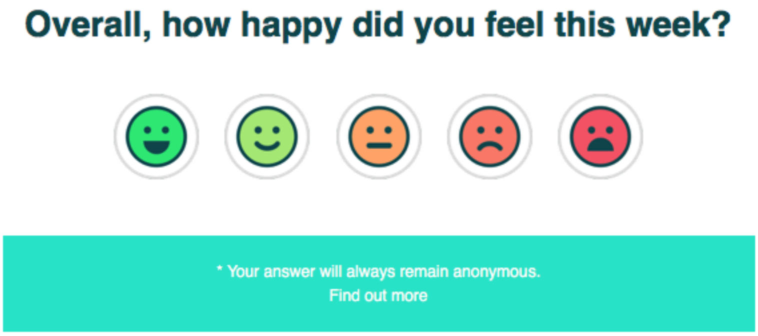
[9] Feeling states are typically categorized into moods and emotions. While emotions are discrete short-run reactions, usually to a specific stimulus, moods are more general positive or negative feeling states (Frijda, 1986). A related literature considers the effects of discrete emotions, rather than mood, on economic behavior (see Loewenstein, 2000).

[10] For a full discussion of the issues surrounding the measurement of subjective well-being, as well as a detailed account of the various ways in which the validity and reliability of such measures have been tested, see OECD (2013). See also Krueger and Stone (2014) and Krueger and Schkade (2008). The OECD reports that well-being measures are now being regularly collected in over 20 countries, many of which are also using them systematically in the policymaking process (Durand, 2018).

[11] This is (by design) very close in wording to the standard happiness question asked by the UK government's Office for National Statistics (ONS) in all of its main household and labor force surveys. The ONS question asks: "On a 0-10 scale where 0 is not at all happy and 10 is completely happy, overall, how happy did you feel yesterday?" (see Dolan et al., 2011).

[12] One notable issue with the use of faces in this way is that people – men and women, in particular – have been shown to interpret the neutral face in different ways (see, e.g., Elfering and Grebner, 2010). Since our main specifications use individual fixed effects and look within-workers over time, such issues of response style are much less of a concern here.

Figure 1: Happiness Survey Email



Notes: Screenshot of the happiness survey, which was sent weekly over a six month period to all workers. Respondents had to click a face within the email for their response to be registered. See text for more details.

The question was asked on a weekly basis for six months, beginning in July 2017. The survey was sent by email on Thursday afternoon, and is shown in Figure 1. It is important to note that we measured happiness toward the end of the week, and asked employees about their happiness during that particular week. This is key, since it allows us to link happiness reports and work outcomes over the same corresponding time period. An alternative option would be to ask employees at the start of the week, which would give a clearer temporal ordering between happiness and subsequent outcomes. However, this would come at the expense of not knowing workers' happiness during the period where they are actually working on those sales. Our use of happiness and sales measures that are temporally concurrent has significant benefits for the analysis; however, the contemporaneous nature of our happiness and sales measures underlines the need for an instrumental variables strategy in order to deal with potential issues of endogeneity.

In order to reduce the onerousness of the data collection on employees, and in doing so ensure as high a response rate as possible, the survey was a single-item question that could be answered simply by clicking an answer from within the email. Specifically, there was no need to click on a link through to a further webpage (as is typically the case with emailed surveys) in order to register a response. Workers were assured that their individual happiness responses would not be shared with the firm's management. Workers were also offered the opportunity to opt-out of the study at any time, via a simple email click-through.

The responses provide us with a weekly, ordinal measure of affective well-being. The distribution of responses is shown in panel (a) of Figure 2. As can be seen in Figure 2, the modal response is the least happy state. If we use the scale in a continuous manner, assigning numerical values 1 to 5 to the categorical responses as is typically done in the literature, the mean response is 2.6 with a within-person standard deviation of 0.95 (a full set of descriptive statistics are shown in Table 1). Importantly for our identification using individual fixed effect models, panel (c) of Figure 2 shows that the responses vary significantly within-workers over time during the study, suggesting that we are not simply picking up a more static overall measure of evaluative job satisfaction.

Table 1: Summary Statistics

	N	Mean	Standard Deviation			Min	Max
			Overall	Between	Within		
Happiness	12,282	2.60	1.38	1.02	0.95	1	5
Sales	12,282	25.25	19.16	14.89	12.41	0	101
Selling Time	12,282	19.99	8.18	5.69	6.38	0	52.43
Internal Shrinkage	12,282	10.52	13.30	10.11	10.74	0	100
Customer Satisfaction	10,120	7.90	2.26	1.20	2.06	0	10
Adherence	12,174	91.99	5.48	3.58	4.34	0	100
Minutes per Call	12,282	12.45	3.34	2.86	1.77	0	37.47
Conversion Rate	11,850	26.63	17.98	15.77	10.59	0	100
Sick Leave	12,279	1.58	7.73	3.37	7.26	0	100
Attendance	12,279	92.58	14.14	6.55	13.15	0	100
Paid Breaks (mins)	12,282	12.49	23.05	22.29	7.94	0	240
Unpaid Breaks (mins)	12,282	213.60	93.92	77.27	63.58	0	540
Overtime (hrs)	12,282	0.15	0.98	0.48	0.89	0	15.58
Paid Time Off (hrs)	12,282	1.07	2.99	1.36	2.78	0	38
Bad Weather Index	12,282	4.06	1.78	1.18	1.36	0	10
Age	1,157	33.81		10.41		17.5	67.5
Female	1,157	0.41		0.50		0	1
Tenure	1,157	4.99		7.17		0.09	43.79
Left Firm During Study	1,157	0.04		0.26		0	1

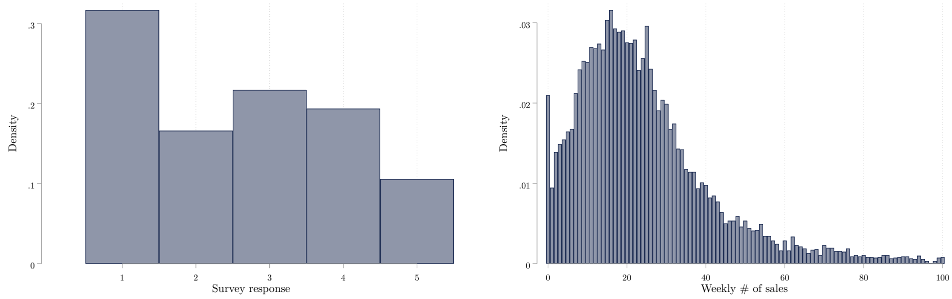
2.2 Administrative firm data

We combine the survey responses with detailed individual-level administrative data from the firm. We focus our attention on sales workers across the 11 call centers, whose job it is to sell a variety of BT products. These predominantly consist of landline and cellphone contracts, broadband internet contracts, and television subscription bundles. The vast majority of the work (91% of time and 82% of tasks) carried out by the employees in the sample are incoming calls from potential or existing customers, with the remainder consisting of outgoing calls (4% of time and 12% of tasks) and "other" activities (which includes tasks such as dealing with inbound and outbound letters, online customer chats, and incoming and outgoing SMS messaging).

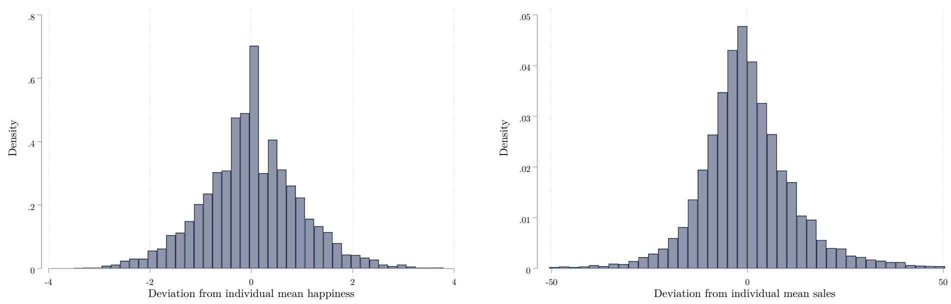
Workers are observed in the data on a daily basis and are identified by their unique personal ID as well as a time-varying team ID. Workers sit individually at desks with a computer terminal and a telephone headset, and are clustered physically in the workspace by their team. Although workers are organized into teams, this largely relates to the sharing of a line manager and not to any interdependence of workflow, since the job of selling in this context is almost exclusively an independent task with very little to no teamwork involved.

At the worker-day level we observe the number of sales, which include new sales, whether to a new or existing customer, as well as instances where the worker is able to retain a customer through re-contracting. The distribution of sales in our sample can be seen in panel (b) of Figure 2. As is typically the case with sales data, the distribution of this count variable is right-skewed. We also look more closely at mechanisms through which any relationship between happiness and sales may take place, and in doing so consider other work outcomes like call duration, adherence to workflow scheduling, call-to-sale conversion, and customer satisfaction.

Figure 2: Distribution of Happiness and Sales



(a) Happiness (b) Sales



(c) Happiness (within-worker) (d) Sales (within-worker)

Note: Panels (a) and (b) show the overall distribution of happiness and sales. Each observation is a worker-week. Panels (c) and (d) show the extent to which these two variables vary within-workers over time. These latter graphs show the residuals from OLS regressions of each variable on individual fixed effects. An observation is an individual-day residual from each regression.

In addition, we observe daily measures of workers' scheduling. Within this, we are able to track the number of hours the employee is scheduled to work, their attendance, the number and duration of paid and unpaid breaks they take, as well as the length of time they spend working on sales (which is factored for call waiting time). Finally, using data from the Human Resources department of the firm, we observe a limited number of time-invariant characteristics of workers such as their gender, age, and tenure.

2.3 Weather data

We link our data with measure drawn from the National Oceanic and Atmospheric Administration's (NOAA) Global Surface Summary of the Day database. We use the address of each call center to determine its latitude and longitude, and match each center to the closest weather station in the data, which is on average 14km away. Since not all station*days have non-missing data available, we link each call center with the closest 5 stations, and where the closest station has missing values we take the second closest, and if missing take the third closest, and so on.[13] Each station reports on a daily basis a separate indicator variable for whether there has been fog, snow and rain on that day.

We construct for each call center location a Bad Weather Index, corresponding to the total

[13] The 5th closest station, which is only very rarely used, is on average 51km away. See `https://data.noaa.gov/dataset/dataset/global-surface-summary-of-the-day-gsod` for more details of the data.

number of daily incidences of fog, rain and snow during the week for which individuals reported their happiness. If it were to rain, snow and be foggy every day of the week, the index would be 15, if it rains for 2 days and is foggy for 1 day, the index would be 3, and so on.

2.4 Sample construction and characteristics

Since our psychological well-being data is reported by worker-week, we aggregate all of the administrative data to the Monday-to-Friday week. In an appendix, we also make use of the daily nature of much of the productivity and scheduling data by constructing a full worker-by-day dataset and assuming happiness to be constant throughout the days of that week (given that the question asks them specifically how happy they have felt overall during the week).

We observe 1,793 sales workers, distributed across 11 different call centers (for a map of the spatial distribution of these call centers, see Figure 4). All of these workers were invited to take part in the study and were sent weekly psychological well-being surveys. Participation was very high. Indeed, of these employees, 1,438 (around 80%) participated by answering at least one survey over the subsequent 6 months. All of this cohort of workers were sent a weekly email, unless they had since left the organization (in which case their email survey is recorded as "bounced"). We do not follow any workers who subsequently joined the firm after the first week of the study.

Conditional on participating in the study, workers responded to a mean of 10.3 waves (with an SD of 7.1). The weekly response rate of workers who participated was on average around 37%. In the supplementary material, we assess the extent to which observable demographic characteristics and workplace performance are able to predict i) participation in the study, on the extensive margin (Table S1) and ii) the number of survey waves answered if the worker did participate, on the intensive margin (Table S2). Importantly, neither participation in the study nor frequency of response are significantly related to the average weekly number of sales made by workers during the course of the study. Mean hours of selling time is positively predictive of the number of response waves – an extra hour of mean daily sales time is associated with around a 2% increase in the number of waves responded.

We drop any participants who responded to only one survey wave, since we rely principally on within-worker variation over time in our main analysis. Also dropped from the analysis are observations that lead to statistical separation in our main individual and week fixed-effect Poisson models. This leaves us with a final sample of 1,157 employees. Summary statistics for this final sample are shown in Table 1. Around 59% are male, and the modal age category is 26-30 (with over 60% of the sample being between 21 and 35 years old).[14] Mean tenure in the firm is about 5 years, with a large standard deviation of 7 years. Half of the workforce in our sample has been in this position for less than 2 years, and around 7% of the sample experienced turnover during the 6 months – either leaving of their own accord or having their employment terminated.

[14] We take mid-points of our age scale, which was measured in 5 year bins. All results are similar when including age FEs and continuous age.

2.5 Non-Response

Using our final sample of workers, we do not observe a fully balanced worker-week panel since we are restricted by non-response to the happiness survey instrument. There is a significant concern that non-response to the survey is unlikely to occur randomly, and may indeed relate to our main variables of interest in ways likely to bias our estimates. For example, it may be that a worker does not respond in a given week because she is either too happy or miserable to spend time reading the email, or alternatively because she is too busy making sales.

In Table S3 we regress a dummy for having responded to the survey in a given week on a number of time-varying observables like sales, selling time, and team average happiness (as well as a set of individual and week fixed effects). Reassuringly, neither weekly sales performance nor team average happiness (minus the focal worker) is significantly related to non-response within-individuals over time. It is, however, positively related to the number of hours worked during the week, suggesting that workers are less likely to respond during weeks in which they are scheduled to work less. Importantly, response is also unaffected by local weather patterns as they vary week to week.[15]

3 Empirical Strategy

We are interested in whether workers' self-reported psychological well-being has any causal impact on their weekly performance at work. We focus on sales workers so that our preferred performance measure is, at least initially, the worker's total number of weekly sales. We estimate a within-worker productivity equation, such that

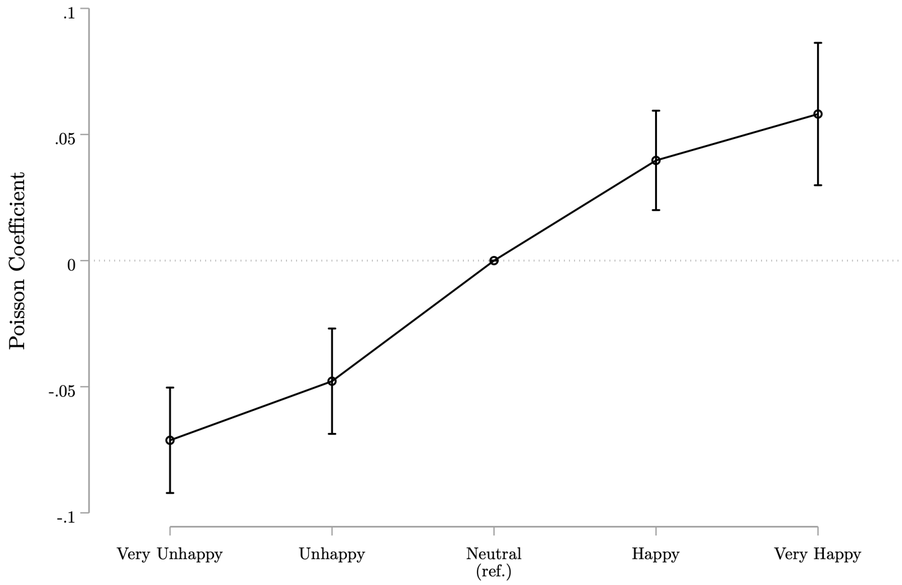
$$E[S_{ijt}|H_{ijt}, X_{ijt}] = exp\{\beta H_{ijt} + \gamma X_{ijt} + \nu_i + \tau_t\} \tag{1}$$

where S_{ijt} corresponds to the number of weekly sales for worker i in call center j during week t and H_{ijt} is her reported happiness during that same period t. Worker fixed effects ν_i capture any individual-specific characteristic that does not change over time, and τ_t is a time fixed effect partialing out any shocks that may affect both well-being and sales. Finally, we include a vector of controls X_{ijt} for two major work schedule variables that may vary over time and across workers, namely the (log of the) total number of selling hours during week t, and the fraction of time spent at work in the week on internal training. We adjust the error term to account for clustering on individuals. We estimate equation (1) with a Poisson quasi-maximum likelihood model. The Poisson model is particularly relevant for count data, and also makes the interpretation of β intuitive, as it can be interpreted in terms of a percent change in number of sales.[16]

[15] One approach to dealing with non-response to the survey would be to make an assumption as to the reason behind the lack of response, for example by imputing any missing values as, say, the lowest or the highest category. However, since the reasons for non-response could be many and are not observed, we choose not to do so. Indeed, the fact that response is related neither to weather nor to team happiness provides suggestive evidence that response behavior is not systematically related to individuals' happiness.

[16] Making comparisons between groups in terms of self-reported happiness can be problematic. As recently argued by Bond and Lang (2019), the comparison of group happiness requires strong identifying conditions when subjective well-being is measured on a discrete scale (namely assumptions on the latent distribution of happiness and similar reported functions between groups). Because in this study we use reported happiness as

Figure 3: Within-worker association of happiness and sales



Note: Coefficients and 95% confidence intervals shown from a Poisson model in which the number of sales are regressed on a series of happiness dummies, a full set of individual and time fixed effects, as well as scheduling controls.

Although we assign a numerical value to each of the responses in H_{ijt}, ranging from 1 (least happy) to 5 (most happy), we are naturally hesitant to take the scale at face value as a continuous measure. We thus estimate equation (1) to begin with using a set of indicator variables for responses (leaving aside the middle response as the omitted category). Figure 3 reports the coefficients from this exercise. We note three things about this model. First, in line with the existing literature on well-being and performance, there is a clear (conditional) correlation between happiness and productivity. Comparing weeks when workers report being very unhappy with weeks where they are very happy, the difference in sales is around 13%. Second, we interpret the pattern of coefficients as suggestive evidence for being able to use the happiness survey in a continuous manner in our instrumented analyses (meaning that we "only" require one valid instrument rather than one for each categorical response).

Finally, the coefficients on happiness estimated using equation (1) are highly likely to be biased, possibly strongly so. One initial reason for this is measurement error in the survey, which will bias the coefficients downward. A significant further empirical concern is that, even within-workers over time, a change in SWB may be endogenous to performance. This is of particular concern since our survey instrument asks about happiness during the whole week, at the end of week. In particular, we see two (opposing) major ways in which reverse causality may bias our coefficients. First, more productive workers can get compensated for their higher performance through financial incentives, or non-monetary rewards, for example from their colleagues or managers or it could simply be enjoyment of successfully completing tasks. This alone could explain their higher reported well-being, in which case the coefficient will be biased

an independent variable and our preferred specification is within-workers, we are not concerned by the issue of group happiness or possible heterogeneity in reporting functions between workers.

Figure 4: Spatial Distribution of Call Centers

Note: Map shows the location of the 11 call centers in the study, as well as the number of sales workers in each.

upward. Second, and conversely, higher productivity can mean being over-worked. This can lead to higher levels of stress and anxiety, which are likely to be strongly negatively correlated with worker happiness. Equally, doing more work may simply be more un-enjoyable. If this is the case, the coefficient will be biased downward.

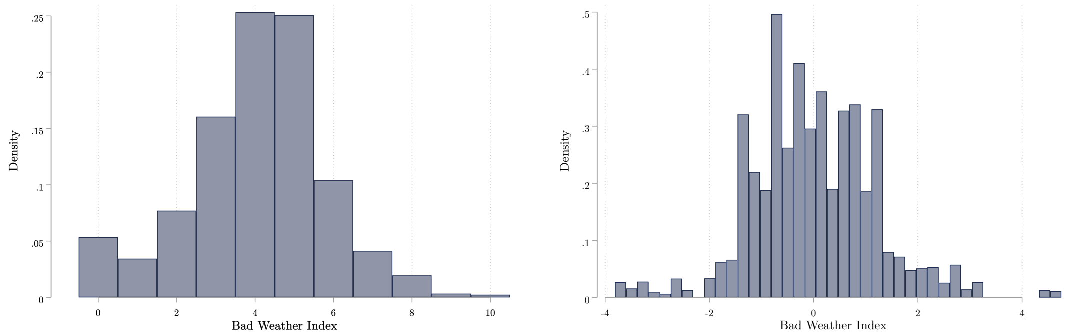
In a setting such as ours, where both extrinsic and intrinsic sources of motivation can be thought of as being relatively low, the downward bias is likely to be strong and dominate any upward bias. Furthermore, in call centers in particular, the pressure to deal with a constant flow of incoming calls, often initiated by unhappy customers, can be particularly distressful and has been shown to lead to burnouts or emotional exhaustion.[17] We use the data to provide some direct evidence on the direction of the bias. In Table S9, in the supplementary materials, we regress weekly happiness on the weekly number of calls received and a full set of individual and week fixed effects as well as the usual controls. We document a strong negative impact on workers' happiness of total number of weekly calls answered. A worker who answers twice as many calls in a given week relative to another week reports a fall in happiness of nearly 0.2 points, which corresponds to a happiness drop of about 20% of a standard deviation within workers. Because there is also a positive link between the total number of calls a worker answers and the number of sales she makes, our initial estimates from equation (1) are thus likely to be biased downward.[18]

Our main strategy to deal with the endogeneity of H_{ijt} is to rely on instrumental variables Z_{ijt} that are correlated with the latter, but are independent of sales. In our case, the first stage

[17] For evidence on this, see prior research in the health and management literatures, (e.g. Castanheira and Chambel, 2010; Zapf et al., 2001).

[18] Note that despite the positive relationship between total number of calls and sales, we still find that happiness improves sales using equation (1). This is consistent with sales calls generating less emotional distress than non-sales calls, or with happiness affecting performance mostly through the conversion of calls into sales, a possibility we later explore in the paper.

Figure 5: Variation in Weather



(a) Raw distribution of index

(b) Distribution conditional on call center and week fixed effects

Note: Panel (a) shows the distribution of weekly bad weather across call centers. The index is constructed by summing the daily incidences between Monday and Friday of fog, snow, and rain (source: NOAA). Panel (b) shows the residuals from an OLS regression of the weather index on a set of call center and week fixed effects.

is:

$$H_{ijt} = \omega Z_{jt} + \Gamma X_{ijt} + \nu_i + \tau_t + \eta_{ijt} \tag{2}$$

where our preferred instrument, Z_{ijt}, uses variation in local weather conditions. The inclusion of τ_t ensures that our key piece of identifying variation, in terms of Z_{jt}, is weather shocks across call centers j, within any given week t.

For identification, given the inclusion of time fixed effects, it is vital that we have a sufficient number of call centers that are spread widely in terms of geography. Moreover, it is important to have a geographical setting where weather does indeed vary significantly across space within any given time period. Great Britain happens to be a place with very variable weather, both across time and space, and, even within-weeks, locations vary significantly in the weather patterns they experience. As can be seen in Figure 4, our call centers are scattered across three countries, and in Figure 5 we show the distribution of the weather variable, and show significant variation even when conditioning on call center and week fixed effects in Panel (b).

The standard two-stage least square (2SLS) estimator is appropriate in the case of linear panel models. However, when it comes to non-linear models with individual fixed effects, a fully robust approach is to rely on control function methods – that is, estimate the first stage for H_{ijt} and add the first-stage residuals as a control in equation (1) (see Papke and Wooldridge, 2008, for another application to non-linear models). We bootstrap the standard errors in order to adjust for the first-stage estimation, re-sampling across individuals 1000 times.

4 Main Results

We first show the simple within-worker relationship between weekly happiness and sales. In column (1) of Table 2 we show a strong positive (partial) correlation between the two.[19]

[19] We look for potential sources of heterogeneity across gender, age, average level of happiness, or length of tenure but find no robust evidence of heterogeneity along these dimensions (see Table S17 in the online appendix).

Table 2: Impact of Happiness on Sales Performance

	Sales (Poisson-FE)		Happiness (OLS-FE)	Sales (Poisson-FE)
	(1) Non-IV	(2) Red.-Form	(3) 1st Stage IV	(4) 2nd Stage IV
Happiness (1-5)	0.0352***			0.2465**
	(0.0034)			(0.1023)
Bad Weather Index		-0.0059**	-0.0247***	
		(0.0026)	(0.0076)	
Observations	12,282	12,282	12,282	12,282
1st Stage F-Stat			10.45	

Notes: Robust standard errors in parentheses, clustered on individuals in models (1) to (3). Bootstrapped standard errors reported for model (4), re-sampling across individuals to account for clustering. All models include individual and week fixed effects, work schedule controls, and dummies for day of week of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

In column (2) we estimate the reduced form effects of local weather on sales performance. Poor weather has a strong negative effect on sales within-workers over time. Given that we can measure both weather and performance (but not happiness) at the daily level, we show this also using daily data. In table S4 we regress daily sales on daily weather, together with a full set of individual and date fixed effects, and our standard set of daily work schedule controls (equivalent to above). We find that poor weather in the geographic region of the call center has a negative effect on sales performance. One concern with our findings using the happiness survey is the threat of bias arising from non-response, which we discussed above. Here we are able to make a further check, and show that the effect of weather on sales is very similar for worker-days on which we have non-missing and missing happiness data.

Moving on to the first stage of our instrumented analysis, column (3) shows the impact of weather on happiness. The impact of weather on short-run emotional states is well documented (Baylis et al., 2018; Connolly, 2013; Feddersen et al., 2016), and we confirm this finding. As can be seen, within-workers over time there is a clear negative relationship between adverse weather conditions and their happiness. The F-statistic from this linear first stage, which is around 10.5, suggests the instrument is sufficiently powerful to be valid. This is not an especially powerful first stage, but is sufficiently so. It is worth re-iterating here how much we are asking of the data, given that we include time fixed effects in the equation: the coefficient here identifies the effect of weather on happiness across call centers within weeks.

The second stage regression, shown in column (4) of Table 2, suggests a strong causal effect of happiness on sales performance. A one point increase in the 1-5 happiness, which amounts to slightly more than one standard deviation increase in within-worker happiness, leads to around a 24.5% increase in weekly sales. This estimate has a bootstrapped 95% confidence interval lying between .046 and .447, rejecting the within-worker estimate presented in column (1). Indeed, the IV estimate is significantly higher than in the equivalent non-instrumented analysis, which is in line with the direct evidence presented above on the relationship between work volume and happiness (as well as with the presence of measurement error in our main right-hand side variable). We return below to a more in-depth discussion on magnitudes.

5 Robustness & Further Analysis

5.1 Threats to Validity of IV Analysis

The weather instrument will only be valid if it has a sufficiently strong impact on happiness, and under the condition that weather has no direct impact on sales. We discuss a number of main threats to these assumptions.

First Stage. The first assumption is testable, and the F-statistic of 10.5 suggests that although our instrument is not hugely strong it is sufficiently so in order to be valid. One concern is the functional form of the first stage. In Figure S1, we show a graphical representation of our first stage. As can be seen, while the relationship looks roughly linear, there is some visual evidence of non-linearity in that heavily bad weather seems to have a particularly strong (negative) effect on happiness. In order to investigate this more fully, we instead estimate it non-parametrically – here, we introduce the weather index into the first stage equation as a series of dummies (leaving out the "zero" category), and show the resultant coefficients and confidence intervals in Figure S2. As can be seen, moderately poor weather does have a significant effect on happiness (compared to good-weather weeks); however, very bad weather has an even stronger effect.[20]

Given this suggestive evidence of non-linearity in the first stage, we do two things (see Table S6). First, we instead use a dummy treatment variable for bad weather (equal to 1 if the index is seven or above). Second, we use the squared value of the weather index. In each case, the first stage is slightly stronger, giving F-statistics of around 13 (compared with 10.5 in the initial case). The resultant second-stage coefficients are within the confidence interval of that initially estimated above.[21]

Product Demand. One major concern is that weather may also have a direct impact on customer demand, or, equally, have an indirect effect on customer demand by affecting customers' mood. Thus in order to identify the causal effect in this set-up, it is vital to include the time fixed effects.[22] Our 11 call centers are distributed across the whole of Great Britain, from the south coast of England to the north of Scotland (see Figure 4 for a map). For identification we rely on variation in weather across call centers within any given week, rather than on movements in national weather conditions from week-to-week. Semi-structured interviews with management of the firm show that calls come from all parts of the country, and are then directed to call centers based on operator availability as well as call type – for example, some centers deal more with TV contracts and some more with broadband internet contracts. Key for our identification is that call centers do not field calls simply because they originate locally. Thus *local* weather in the vicinity of the focal call center should be independent from customer demand. In addition, we provide direct evidence that customer demand at a particular call

[20] Such non-linearities may also be at least partly explained by the discrete nature of our happiness survey instrument. In order to identify any impact of bad weather on happiness, our instrument must be strong enough to generate a one-point change in reported happiness (which is roughly equivalent to one within-worker standard deviation), at least for those workers whose latent happiness level corresponds exactly to their cut-off value. This is more likely to happen during extremely bad weather weeks.

[21] It is also worth noting that, in both cases, the 95% confidence intervals around the second-stage estimate include the non-instrumented estimate presented in column (1) of Table 2.

[22] A paper similar in nature to ours, Coviello et al. (2018), also uses weather to instrument for worker mood in a call center setting, but does not include time fixed effects.

center is not affected by local weather. Looking more directly in the data (collapsed into a panel of call-center-days), we show in Table S7 that the number of incoming calls per worker is not affected by daily weather that is local to that particular call center.

Labor Supply: Opportunity Cost of Leisure. An additional concern is that the presence of (in)clement weather outside may change the opportunity cost of leisure time (see Connolly, 2008), leading workers to adjust their labor supply decisions. Relatedly, bad weather could have direct impacts on employee's ability to attend, or arrive on time – for example, if rain affects their ability to commute to work.

First, it is worth noting that, in our setting, workers have very little discretion over labor supply. Work hours are scheduled by management, and once workers are at their terminal during their scheduled hours they face calls as they come in. Second, we note that all of our analyses are robust to the inclusion of a control for the number of working hours done by the employee during the week (as well as other scheduling controls for internal training time). That is, our effects are robust when conditioning on labor supply decisions, i.e. when looking solely at the intensive margin. In addition to this, we also show directly in Table S7 that daily attendance is not associated with local weather conditions. We further discuss labor supply effects below when discussion channels, and find null effects of instrumented happiness on various fine-grained measures of labor supply.

Labor Supply: Sickness. A concern related to this is the possibility that adverse weather conditions may cause sickness among workers, and impair their ability to attend work. We use the fine-grained detail of our dataset to show that local weather conditions do not significantly predict sickness leave (see Table S7). A related problem, however, is that our identifying variation has the potential to affect on-the-job health at a level lower than affecting sickness leave-taking. Unfortunately, we are not able to provide any direct evidence on this. One point to note here is that the issue is likely to be much less problematic in the UK context compared with, say, the USA (and many other places) in that workers have the right to self-certified paid sick leave (requiring a doctor's note only if this exceeds 7 days in a row).[23]

Treatment Heterogeneity. In the presence of treatment response heterogeneity, a valid instrument identifies a local average treatment effect (LATE), that is the effect driven by those who comply with the treatment. In our case, this may be an issue if only certain types of workers care about weather. Assuming these workers also tend to be those whose productivity reacts more strongly to happiness, this effect alone could explain the larger coefficient found in the IV case. We investigate the possibility of treatment response heterogeneity directly by looking at whether our bad weather index affects workers' mood differently across a number of important characteristics. We look at basic demographics (gender, age and workers' tenure), the total number of weekly sales and whether the worker reports an above median average happiness level during the entire period. Table S8 shows the first stage of our IV strategy where we interact our bad weather index with each of these five main characteristics. We find no evidence of heterogeneity across any of these dimensions.

[23]Thus sick workers are not likely to show up for work, and would be detected in our empirical test of the relationship between weather and sick-leave taking.

5.2 Additional Robustness

Sharper Analysis Using Daily Data. Although our happiness data is measured at the weekly level, our productivity data is reported largely at the daily level. In our main analysis, we aggregated performance data to the weekly level, but it is also possible to produce a sharper, more fine-grained analysis in order to produce daily estimates. We showed above that there is a strong daily reduced-form relationship between weather and sales.

We can go beyond this in order to also show instrumented estimates at the daily level. Here we assume that responses to the happiness question, which is asked on Thursday and refers specifically to "this week", apply equally to each of the weekdays. We again estimate Poisson regressions predicting the number of daily sales, and include a set of individual and date (that is, day*week*year) fixed effects as well as daily controls for selling time and internal training time. As can be seen in Table S5 in the appendix, results are very similar when we carry out this exercise – both in terms of our within-worker estimates as well as our instrumented estimates using local weather shocks in order to set up a natural experiment.

Production function approach. Our main approach used a Poisson model predicting a worker's number of weekly sales, using her level of happiness that week. An alternative approach, is to estimate a standard Cobb-Douglas worker-specific production function of weekly sales,[24] augmented by our measure of worker happiness, such that

$$ln(S_{ijt}) = \beta ln(H_{ijt}) + \gamma ln(X_{ijt}) + \nu_i + \tau_t + \epsilon_{ijt}.$$

We show in Table S14 that our main within-worker estimates are similar when doing so. The Poisson coefficient on log happiness is .077 (.008), suggesting that a 10% increase in our happiness scale increases weekly sales by a bit less than 1%. When instrumenting log happiness with our adverse weather index, we show in Table S14 the robustness of our findings to this alternative approach. In this specification, the first stage is slightly stronger than when using weekly data, with an F-statistic of 12. The resultant second-stage estimate is consistent with our weekly estimates.

Effects on Customer Satisfaction. Even if happier workers are able to extract more sales from customers, this may come at the cost of lower customer satisfaction (and potentially fewer customers in the long-run). Assuming customers are rational, they should only accept offers whose marginal cost does not exceed their marginal utility. Taking customer satisfaction as a proxy for customer surplus, a higher surplus (hence satisfaction) may occur if more customers are buying.

To investigate this issue further, the firm provided us with data on customer satisfaction, since after each call customers are asked by text or phone to report the extent to which they would recommend BT to others, on a 0 to 10 scale. This measure is somewhat noisy, since each worker on average receives very few feedback ratings per week.[25] Using the weekly average response for each worker, we show in Table S13 that, within-workers over time, employee

[24] See Black and Lynch (1996, 2001) or Attanasio et al. (2015) for a similar estimation on the impact of human capital on productivity.

[25] We take only instances where the worker receives two or more in a given week. We are of course also well aware that such measures suffer from a number of further limitations, including the possibility of selection bias in whether happy or unhappy customers are more or less likely to answer.

happiness increases customer satisfaction. Instrumenting for happiness using our weather data, however, we find little effect of employee happiness on customer satisfaction.

5.3 Evidence on Channels

Having shown a robust causal effect of happiness on overall sales performance, we move on in this section to what behaviors drive this effect. We consider two possible broad categories of channel: labor supply and labor productivity.

5.3.1 Extensive Margin: Effects on Labor Supply

In our setting we observe detailed, high frequency data on workers' labor supply decisions. At the extensive margin, we observe a percentage measure of weekly attendance, which has a mean of around 93%. Here we code whether the employee recorded perfect attendance to her scheduled hours during the week.[26] In the second-stage of a 2SLS equation in which we instrument for happiness using adverse weather, shown in column (1) of Table 3, we find no robust evidence of any happiness effects on attendance. As discussed earlier, this rules out the possibility that happier workers would chose to attend work more often, for instance due to higher levels of motivation. Equally, we find negative but non-significant coefficients for overtime working and paid vacation. If anything, increased happiness makes people more likely to take paid sick leave from the firm, though the standard error means our estimate is only weakly significant.

We also observe the number of length of paid and unpaid breaks taken by workers. Here we code the overall number of minutes taken during the week. The vast majority of breaks are unpaid (around 70% of observations for paid breaks are zero). In both cases, the coefficient is negative, suggesting that happiness increases workers' labor supply. However, in neither case is the estimate precise. This rules out the possibility of happier workers taking shorter breaks to work extra hours, or longer breaks to discuss with their co-workers (the "sociability hypothesis"). Overall, taking each of our labor supply measures together, we find very little evidence of any robust happiness effects on labor supply decisions. This is in line with what would be expected in the context of a call center, where employees work alone on independent tasks and have little autonomy or freedom to decide how much they work, once they arrive and are sat at their terminal. We thus do not interpret this evidence as strongly suggesting that happiness does not affect labor supply decisions in general.

As we noted above in relation to the validity of our weather instrument, the very limited labor supply flexibility in our call center field site provides us with an ideal setting in which to test for the pure productivity effects of happiness in a real-world workplace setting.

5.3.2 Intensive Margin: Effects on Labor Productivity

If happier workers do not work *more*, it may well be that they work *better* or *faster*. Here we consider three main measures of labor productivity. First, workers attend and have their day's workflow scheduled for them and displayed on their terminal screen. For example, they may

[26] Similar findings are found when using the continuous measure of attendance.

Table 3: Extensive Margin Effects: Happiness and High-Frequency Labor Supply

	Attendance (100% =1)	Overtime (Any = 1)	Paid Vacation (Any = 1)	Paid Breaks (Minutes)	Unpaid Breaks (Minutes)
	(1)	(2)	(3)	(4)	(5)
Happiness (1-5)	-0.0661	-0.0860	-0.1393	-4.7327	-19.1536
	(0.1575)	(0.0737)	(0.1293)	(3.1060)	(21.1061)
Observations	12,279	12,282	12,282	12,282	12,282

Notes: Second-stage 2SLS models reported, using adverse weather index as an IV for happiness. Robust standard errors in parentheses, clustered on individuals. All models include include individual and week fixed effects and dummies for day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Intensive Margin Effects: Happiness and High-Frequency Labor Productivity

	Adherence (Met Target=1)		Minutes Per Call (Log)		Conversion Rate (Log)	
	(1)	(2)	(3)	(4)	(5)	(6)
Happiness (1-5)	0.2748**	0.2763**	-0.1916**	-0.1972**	0.5428***	0.6419***
	(0.1374)	(0.1390)	(0.0787)	(0.0826)	(0.2071)	(0.2167)
# Sales		-0.0011		0.0015*		
		(0.0011)		(0.0008)		
Minutes per call (ln)						0.6204***
						(0.1177)
Observations	12,169	12,169	12,100	12,100	11,720	11,720

Notes: Second-stage 2SLS models reported, using adverse weather index as an IV for happiness. Robust standard errors in parentheses, clustered on individuals. All models include individual and week fixed effects, and controls working hours, internal training time, and day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

have the first hour scheduled as selling TV bundles, the second selling internet connections, a 15 minute break, and then an hour selling something else. The firm routinely records the extent to which employees "adhere" to this scheduled workflow. Occasional deviance from this workflow may be beneficial (if the worker has to stay on a call to complete a sale, for example), and the firm sets a loose target of 92% adherence each week. We code our outcome variable here as 1 if this target is met. We find in the second stage of a 2SLS regression (column (1) of Table 4), in which happiness is instrumented for using local weather, that happier workers adhere more closely to the workflow that has been set out for them.

Second, we observe on a daily basis the total number of minutes spent on incoming calls as well as the number of calls taken. We code the average length of each call during the week. This "speed" measure is what would typically be used as a labor productivity metric. However, in a call center setting it is not at all clear that taking more, shorter calls will be beneficial, when the goal is selling (hence the use of overall sales as our main performance metric). First, faster calls may displease customers and make them less likely to buy if the operator is too blunt or quick with them. Second, sales calls are likely to be mechanically longer, due to the time it takes to complete an order, take payment details, and so on. Nevertheless, we show in column (3) of Table 4 that in happier weeks workers do work faster. This remains true when controlling for the number of sales made, which is itself positively correlated with the length of calls.

Finally, we observe the ratio of sales to incoming calls. Here, we show in column (5) of Table 4 that in happier weeks, workers convert more of their calls to sales. Interestingly, we find in column (6) that the average length of calls is positively, not negatively, correlated with the conversion rate, though this does not affect the coefficient on happiness. Although in happier weeks workers work faster by making more calls per hour on the phone, this does not seem to come at a quality cost in terms of affecting their ability to convert more calls to sales.

6 Discussion

A long-running literature has sought to explain heterogeneity in productivity across individuals as well as firms (e.g. Ichniowski et al., 1997; Lazear, 2000; Syverson, 2011). We contribute to this growing body of work by showing the causal effect of a typically-overlooked aspect of workers' lives: how happy they are. We show that being in a positive mood has a significant impact on the number of sales made by employees. We find that when workers are happier, they work faster by making more calls per hour worked and, importantly, manage to convert more of these calls to sales (while maintaining customer satisfaction).

6.1 Magnitudes & Benchmarking

In our main specification, we document a 95% bootstrapped confidence interval for an effect of a one unit increase in happiness on sales of between .046 and .447 log points. Depending on specification, we find a point estimate somewhere between .18 to .24 log points. There are a number of different ways to think about the size of this effect. First , we can think in terms of the absolute number of sales our estimates correspond to. The estimate of a 17 to 23 percent increase (for a one standard deviation change in happiness) ultimately corresponds to around 4 to 6 extra sales a week (above a mean of 25). Or, in other words, the estimates translate into around one extra sale per day.

Second, while an estimate of around 24% seems initially high in magnitude, it is worth noting that the within-person standard deviation of happiness is 0.95 on the 1 to 5 scale. Thus a one point increase in the happiness scale corresponds to a very large change, and one that is not likely to be achieved with small changes in workplace organization and culture. For instance, in a field experiment with IT workers, Moen et al. (2017) find the introduction of flexibility and support program improved job satisfaction by 0.27 standard deviations (and reduced stress and distress by 0.17 and 0.16 standard deviations respectively).[27] In a very ambitious large-scale experiment on the performance impact of working from home, Bloom et al. (2014) were able to move workers' positive and negative emotions by around half a standard deviation.[28]

Third, we can compare our instrumented estimate with the non-instrumented one. The

[27] These effect sizes are taken from Table 2, Panel A. The standardized effects are calculated using the "usual-practice" control group standard deviations of each measure.

[28] Positive and negative emotions scales were adapted from the well-known Positive and Negative Affect Schedule (PANAS) scale. We use the replication data provided by the authors to re-run regressions of Table 7, replacing the logged dependent variables with standardized dependent variables. The experiment reduced negative emotions by .44 SDs and increased positive emotions by .55 SDs. Similar effect sizes are found for evaluative satisfaction measures. Keeping the logged dependent variables and inferring standardized effect sizes using the control group SDs provides similar estimates.

point estimates for our causal effects are between 5 and 7 times larger than our initial within-person partial correlations, which is undoubtedly a large difference. However, for the various reasons noted (and tested empirically) above, the within-person estimate is likely to be strongly downward biased. In addition, while our non-instrumented estimate is relatively precisely estimated, our IV estimates are much less so, which implies there is still the possibility that it could be smaller.[29]

Fourth, we can benchmark our estimates against other common predictors of sales in the data. We run a pooled cross-sectional regression of sales on a number of demographic characteristics of workers (as well as time fixed effects and the usual set of time-varying controls, as in our main specification). We show that workers who have at least 1 year of experience on the job (compared with those with less than 1 year of tenure) generally make around 24% more sales per week. Equally, those who are paid in a more strongly incentivized way (a point we return to below) also make around 23% more sales per week. Males workers and workers under 40 in general make around 10% more sales per week on average.

Fifth, we can compare our estimates to the findings of controlled laboratory experiments. Here we focus on the pioneering paper of Oswald et al. (2015), who manipulate happiness in the laboratory using comedy clips (as well as other techniques) and subsequently have subjects do a piece-rate productivity task. While this is clearly a very different context, we note that call center work is not entirely dissimilar, in that it involves a series of individual tasks with very little teamwork. In this experimental set-up, a short-run one standard deviation increase in happiness (induced by viewing a comedy as opposed to a placebo video clip) causes participants to correctly do around 29 to 35 percent more incentivized additions.[30]

Finally, we can look at the result of field experiments in which the improvement of management practices simultaneously impacted employee happiness and productivity. This is much less ideal than the laboratory experiments noted above, since in these contexts we cannot of course treat the relationship between happiness and productivity as a causal parameter. Even so, we can assess the extent to which our estimates are consistent with the observed patterns. Bloom et al. (2014), for example, run a field experiment that also takes place within a call center setting. Though they are primarily interested in the impact of working-from-home on the performance outcomes of non-sales workers (typically the total number of calls), they also assess how the new policy impacted workers' affective states. As noted above, the policy change led to a 0.55 standard deviation increase in positive emotions (and a 0.44 SD fall in negative emotions), which, using our IV estimate, is consistent with the 13% increase in productivity they observe.[31]

[29]However, the lower bound IV estimate of .046 remains higher than the within-person point estimate.

[30]This figure is implied by the results reported from Experiment 2 in Oswald et al. (2015). The authors measure happiness before and after viewing the comedy (or placebo) video clip. The difference in happiness between the two groups following the videos is 0.67 on their 1 to 7 scale. The standard deviation of this scale is 0.86 among the control group when measured prior to the clip. The treated group do 4.15 more additions than the control group in the raw data, over a base of 18.1 in the control group. This implies that a one unit increase in happiness causes a $4.15/0.67 = 6.194$ increase in correct additions. A one standard deviation increase in happiness causes a $6.194 * 0.86 = 5.327$ increase in additions. This is equivalent to a 29.4% increase in productivity. The implied difference in additions between treatment and control is larger (5.01) when accounting for various covariates in a regression analysis. This would imply that a one standard deviation increase in happiness causes a 35.5% increase in productivity.

[31]The coefficient reported in column (4) of Table 2 suggests that treated employees made 13% more phone

6.2 Psychological mechanisms

While the main focus of the paper has been to assess the extent to which there is a main, causal effect of happiness on performance, our data and setting nevertheless allow us to offer some suggestive evidence on mechanisms. In this section, we draw on experimental work—largely in psychology—on the effects of induced affect for clues as to potential mechanisms. We look at three broad psychological pathways through which happiness may feed through into higher sales: intrinsic motivation, cognitive processing and emotional or social skills.

6.2.1 Intrinsic Motivation

First, feelings of happiness have been shown to increase intrinsic motivation (Erez and Isen, 2002).[32] This effect could be particularly important in more routine work settings like ours, or in the absence of strong enough financial incentives to do good work. Because we cannot measure intrinsic motivation directly, we instead compare workers whose performance level may be more (versus less) *reactive* to intrinsic motivation. We rely on the presence of financial incentives to do good work – that is, on the nature of our sales workers' pay-for-performance (PFP) schemes. In high performance pay settings, productivity may be less reactive to intrinsic motivation motives relative to the extrinsic incentives to sell (see, e.g. Frey and Jegen, 2001). In other words, if happiness impacts sales through intrinsic motivation, an exogenous happiness shock may have a *smaller* effect on those workers who are working under strong enough financial incentives.

In our sample of sales workers, 70% work under performance pay: they receive a bonus amounting to 30% of their base salary if they meet their target (note that this is not an explicit piece-rate but rather a slow-moving bonus scheme). The 30% remaining workers do not benefit from this large performance pay scheme. In Table S10, we show that the performance of sales workers who do not benefit from the high performance pay scheme also seems more reactive to random unhappiness shocks, as captured by our bad weather index.[33] The sign of the interaction is positive for un-instrumented happiness (but not significant).

This suggests that strong enough extrinsic motivations may limit the scope of unhappiness on performance, but that when workers perceive these incentives to be too low (or when these are simply missing), the performance costs of unhappiness can be very large. Of course, these results are only suggestive in nature. In order to make a more definite claim, one would need to rely on exogenous changes in performance pay within-workers, which goes beyond the scope of this paper.

6.2.2 Cognitive Abilities

A further possibility is that better mood states may improve cognitive functioning and workers' abilities to remain focused on their daily tasks. Indeed, subjects induced into positive mood

calls per week.

[32]In this sense, our work relates to a recent literature in economics that has begun to concern itself with non-monetary forms of motivation (see, e.g. Benabou and Tirole, 2003; Gneezy and Rustichini, 2000).

[33]As we want to account for the endogeneity of our happiness measure when we discuss heterogeneous effects, we also look at interactions using our bad weather index.

Table 5: Effects of Happiness by Sales Type

	# Sales		
	(1) Phone + Internet Lines	(2) TV + Cell Phone Contracts	(3) Upgrades + Re-Contracting
Happiness (1-5)	0.0004	0.1549	0.3155*
	(0.2050)	(0.1767)	(0.1661)
Observations	12,241	12,268	12,264

Notes: Second-stage Poisson-IV models reported. Bootstrapped standard errors in parentheses, re-sampling across individuals. All models include controls for the log of working hours, internal training time, and day of week dummies for response to survey. Happiness is instrumented using local weather conditions in the first stage. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

states have been shown to become better at cognitively processing information (Estrada et al., 1997) as well as better at creative problem solving tasks (Isen et al., 1987). Relatedly, it has been shown that the thoughts of unhappier people are more likely to "wander" (Killingsworth and Gilbert, 2010), a mechanism that has been formalized into an economic model in which happiness reduces the amount of time spent worrying about negative aspects of people's lives, and thus drives productivity (Oswald et al., 2015).

Though we cannot measure cognitive abilities directly, we already saw that a worker's adherence, a proxy arguably capturing the capacity to stay focused on a given task, increases with happiness (see Table 4). We can further investigate this point by looking at weeks during which workers have to do other type of tasks on top of their usual incoming calls activity (which represent 91% of their average weekly hours). Multi-tasking generally requires stronger cognitive abilities as it becomes harder to stay focused on a single task. We should therefore expect that workers who have to multi-task more in a given week should also be less productive in terms of sales. However, if happiness improves cognitive processing, the sales performance of happy multi-tasking workers should be less negative than the performance of unhappy multi-tasking workers.

In Table S11 we interact happiness with the fraction of weekly working hours during which sales workers were doing other tasks than responding to incoming calls. We find that weeks during which workers are asked to allocate more time to other tasks or to outgoing calls in addition to their regular activity tend to have lower sales performance. However, we do not find any evidence of this varying at different levels of happiness.

6.2.3 Social Skills and Emotional Labor

Finally, happiness could also augment non-cognitive ability – that is, social and/or emotional skills. For example, happy workers have been shown to demonstrate higher powers of self-control and abilities to manage their emotions, which is particularly useful in customer-facing industries and occupations such call center work (Goldberg and Grandey, 2007; Tice et al., 2001). One might thus be able to think of happiness as an aid to social skills, allowing workers to better negotiate with customers and find appropriate, mutually agreeable solutions.

To test for this, we first examine the effect of happiness on different types of sales. Although we do not observe the amount of time spent (or number of calls) selling different products, we do observe the breakdown of realized sales. We split the sales into three categories: i) regular sales of phone lines and internet connections, which typically involve relatively simple order-taking; ii) TV and cell-phone contracts, which tend to be more complex, variable, and involve a greater amount of selling and negotiation skill; and iii) re-contracting and upgrades, which usually involve the most amount of social and negotiating skills (and also typically involve disgruntled customers).

In Table 5 we find that happiness does not have a significant effect on regular order-taking. This is consistent with the main mechanism being call-to-sales conversion rather than working faster, since line sales are largely mechanical order-taking. A positive effect is evident for TV and cell-phone contracts (though the estimate is somewhat imprecise), where there is more leeway for persuasion and negotiation. A stronger positive effect is evident for the re-contracting sales, suggesting happiness mostly impacts tasks involving the use of social and negotiating skills.

The sociological literature on *emotional labor* (see, e.g. Hochschild, 1983) has long argued that in tasks involving interactions with customers, it becomes particularly costly for unhappy employees to leverage their social skills and manage their emotions, as they need to "fake" happiness. This is particularly the case when dealing with unsatisfied customers, as employees' ability to empathize and "tame" customers' negative emotions then comes at a larger psychological cost.

To further test that hypothesis, we first look at the interaction between workers' happiness (and our bad weather index) and BT's overall customer satisfaction across all of our call centers in a given week. In weeks where national customer satisfaction is low, the impact of being in a good (bad) mood should have a much stronger positive (negative) effect on sales. We find evidence for this effect in Table S12. The data is thus also suggestively supportive of the emotional labor hypothesis since it implies that happiness mostly improves sales when workers are interacting with disgruntled customers.

There is a growing interest in the role of social skills in the labor market, and it is now well-accepted that the number of jobs requiring workers to interact socially is increasing rapidly (Deming, 2017). Our analysis, which points to a strong role for happiness in improving re-contracting and upgrade sales, in particular when workers have to deal with unsatisfied customers, suggests that the importance of employee happiness in driving productivity growth is likely to rise in the coming years.

6.3 Further Limitations

While we show an effect of happiness on productivity, we are not able—given our data and setting—to adjudicate as to whether investing in schemes to enhance employee happiness makes good business sense for a firm. Any such adjudication is naturally dependent on both the costs of raising worker happiness as well as the potential benefits in terms of productivity (as well as other potential benefits through recruitment, retention, and so on). We have sought to provide evidence for the latter half of that equation.

A further limitation is that we are not able to observe what it is in our setting that affects workers' level of (un)happiness from week-to-week, beyond variation in weather. Thus we are not in a position to make any recommendations about what firms might do in order to boost happiness and (hope to) gain from any productivity effects. A natural question is the extent to which the local average treatment effect of weather-induced unhappiness on productivity is a useful (i.e. policy-relevant) parameter to estimate. Given that there has been nearly a century of empirical work on workplace well-being and performance, but with very little–if any–causal evidence in the field, we present our findings as an important step forwards in the literature. In doing so, we confirm laboratory findings on the same phenomenon (e.g. Oswald et al., 2015) in a real-world workplace setting. Our study is best thought of as a form of basic research in an applied setting. Although we examine data in a field setting, the analysis is focused primarily on isolating and estimating the causal effect of happiness on performance. The focus is not on management practices that may increase happiness and subsequent productivity, something that has been shown in various recent field experiments (Bloom et al., 2014; Breza et al., 2017; Gosnell et al., 2020).[34]

We show that week-to-week changes in employee happiness have significant causal effects on productivity. We are not able to say whether these short-run effects translate into any long-run effects, however. This is a significant further limitation since many of the management practices that firms would be likely to invest in are presumably mostly focused on improving the long-run happiness of workers.

7 Conclusion

The study of employee well-being and performance has a long history. In response to early scientific management theories— often loosely referred to as "Taylorism"—that were prevalent at the turn of the twentieth century, a large body of work began to try to account more fully for the "human element" in production.[35] Arguably, an analogous turn is currently occurring in personnel and labor economics nearly a century later, with researchers increasingly looking at behavioral aspects of work and workplaces, ranging from fairness concerns to social preferences and reference dependence, in an attempt to complement empirical and theoretical work on the ability of financial incentives to solve agency problems between firms and workers (e.g. Benabou and Tirole, 2003; Cassar and Meier, 2018; DellaVigna and Pope, 2017).

The large and long-running stream of research on job satisfaction, largely outside of economics, has produced mixed findings of a relationship with job performance (Judge et al., 2001). Nevertheless, a more recent turn in the literature away from satisfaction measures and onto workers' affective states, driven in large part by developments in psychology, has produced more robust causal findings – at least in laboratory settings where happiness can be manipulated in a controlled manner (e.g. Oswald et al., 2015). As yet, however, strong causal evidence in the field of an effect of happiness on performance is lacking. In this paper we leverage varia-

[34] Manipulating management practices in this way is a hugely useful endeavour but it is an approach that will struggle to provide clean evidence on the potential happiness mechanism, since almost all management interventions are likely to have at least some direct effect on performance even if they boost happiness at the same time.

[35] Given that focus, this body of work came to be known more commonly as the Human Relations movement, and is most frequently associated in the present day with experiments carried out at the Hawthorne plant of the Western Electric Company (see Mayo, 1993; Roethlisberger and Dickson, 1939).

tion in local weather conditions, together with a novel survey and detailed quantitative data on workplace behaviors and performance, in order to show a strong positive impact of employee happiness on productivity.

References

Aghion, P., Akcigit, U., Deaton, A., and Roulet, A. (2016). Creative destruction and subjective wellbeing. *American Economic Review*, 106(12):3869–97.

Ariely, D., Kamenica, E., and Prelec, D. (2008). Man's search for meaning: The case of legos. *Journal of Economic Behavior & Organization*, 67(3-4):671–677.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2015). Estimating the production function for human capital: Results from a randomized control trial in colombia. Technical report, National Bureau of Economic Research.

Baylis, P., Obradovich, N., Kryvasheyeu, Y., Chen, H., Coviello, L., Moro, E., Cebrian, M., and Fowler, J. H. (2018). Weather impacts expressed sentiment. *PloS one*, 13(4):e0195750.

Benabou, R. and Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(3):489–520.

Benjamin, D. J., Kimball, M. S., Heffetz, O., and Rees-Jones, A. (2012). What do you think would make you happier? what do you think you would choose? *American Economic Review*, 102(5):2083.

Black, S. E. and Lynch, L. M. (1996). Human-capital investments and productivity. *American Economic Review*, 86(2):263–267.

Black, S. E. and Lynch, L. M. (2001). How to compete: the impact of workplace practices and information technology on productivity. *Review of Economics and statistics*, 83(3):434–445.

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

Böckerman, P. and Ilmakunnas, P. (2012). The job satisfaction-productivity nexus: A study using matched survey and register data. *ILR Review*, 65(2):244–262.

Bond, T. N. and Lang, K. (2019). The sad truth about happiness scales. *Journal of Political Economy*.

Breza, E., Kaur, S., and Shamdasani, Y. (2017). The morale effects of pay inequality. *Quarterly Journal of Economics*, 133(2):611–663.

Brief, A. P. and Weiss, H. M. (2002). Organizational behavior: Affect in the workplace. *Annual Review of Psychology*, 53(1):279–307.

Bryson, A., Forth, J., and Stokes, L. (2017). Does employees' subjective well-being affect workplace performance? *Human Relations*, 70(8):1017–1037.

Capra, M. C. (2004). Mood-driven behavior in strategic interactions. *American Economic Review*, 94(2):367–372.

Cassar, L. and Meier, S. (2018). Nonmonetary incentives and the implications of work as a source of meaning. *Journal of Economic Perspectives*, 32(3):215–38.

Castanheira, F. and Chambel, M. J. (2010). Reducing burnout in call centers through hr practices. *Human Resource Management*, 49(6):1047–1065.

Connolly, M. (2008). Here comes the rain again: Weather and the intertemporal substitution of leisure. *Journal of Labor Economics*, 26(1):73–100.

Connolly, M. (2013). Some like it mild and not too wet: The influence of weather on subjective well-being. *Journal of Happiness Studies*, 14(2):457–473.

Coviello, D., Deserranno, E., Persico, N., and Sapienza, P. (2018). Effect of mood on workplace productivity. *Mimeo.*

DellaVigna, S. and Pope, D. (2017). What motivates effort? evidence and expert forecasts. *Review of Economic Studies*, 85(2):1029–1069.

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

Dolan, P., Layard, R., and Metcalfe, R. (2011). Measuring subjective well-being for public policy. Technical report, Office for National Statistics.

Durand, M. (2018). Countries' experiences with well-being and happiness metrics. In Sachs, J., editor, *Global Happiness Policy Report*, pages 200–245. Global Happiness Council.

Edmans, A. (2011). Does the stock market fully value intangibles? employee satisfaction and equity prices. *Journal of Financial economics*, 101(3):621–640.

Edmans, A., Garcia, D., and Norli, Ø. (2007). Sports sentiment and stock returns. *The Journal of Finance*, 62(4):1967–1998.

Elfering, A. and Grebner, S. (2010). A smile is just a smile: But only for men. sex differences in meaning of faces scales. *Journal of Happiness Studies*, 11(2):179–191.

Eren, O. and Mocan, N. (2018). Emotional judges and unlucky juveniles. *American Economic Journal: Applied Economics*, 10(3):171–205.

Erez, A. and Isen, A. M. (2002). The influence of positive affect on the components of expectancy motivation. *Journal of Applied Psychology*, 87(6):1055.

Estrada, C. A., Isen, A. M., and Young, M. J. (1997). Positive affect facilitates integration of information and decreases anchoring in reasoning among physicians. *Organizational Behavior and Human Decision Processes*, 72(1):117–135.

Feddersen, J., Metcalfe, R., and Wooden, M. (2016). Subjective wellbeing: why weather matters. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(1):203–228.

Fisher, C. D. (2000). Mood and emotions while working: missing pieces of job satisfaction? *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 21(2):185–202.

Frey, B. S. and Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, 15(5):589–611.

Frijda, N. H. (1986). *The emotions.* Cambridge University Press.

Gneezy, U. and Rustichini, A. (2000). Pay enough or don't pay at all. *Quarterly Journal of Economics*, 115(3):791–810.

Goldberg, L. S. and Grandey, A. A. (2007). Display rules versus display autonomy: emotion regulation, emotional exhaustion, and task performance in a call center simulation. *Journal of Occupational Health Psychology*, 12(3):301.

Gosnell, G. K., List, J. A., and Metcalfe, R. D. (2020). The impact of management practices on employee productivity: A field experiment with airline captains. *Journal of Political Economy.*

Haushofer, J. and Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: experimental evidence from kenya. *Quarterly Journal of Economics*, 131(4):1973–2042.

Heyes, A. and Saberian, S. (2019). Temperature and decisions: evidence from 207,000 court cases. *American Economic Journal: Applied Economics*, 11(2):238–65.

Hirshleifer, D. and Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance*, 58(3):1009–1032.

Hochschild, A. R. (1983). *The managed heart.* Berkeley CA: University of California Press.

Ichniowski, C. and Shaw, K. (2003). Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices. *Journal of Economic Perspectives*, 17(1):155–180.

Ichniowski, C., Shaw, K., and Prennushi, G. (1997). The effects of human resource management practices on productivity: A study of steel finishing lines. *American Economic Review*, pages 291–313.

Ifcher, J. and Zarghamee, H. (2011). Happiness and time preference: The effect of positive affect in a random-assignment experiment. *American Economic Review*, 101(7):3109–29.

Isen, A. M. (2001). An influence of positive affect on decision making in complex situations: Theoretical issues with practical implications. *Journal of Consumer Psychology*, 11(2):75–85.

Isen, A. M., Daubman, K. A., and Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, 52(6):1122.

Judge, T. A., Thoresen, C. J., Bono, J. E., and Patton, G. K. (2001). The job satisfaction–job performance relationship: A qualitative and quantitative review. *Psychological Bulletin,*, 127(3):376.

Kahneman, D. and Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1):3–24.

Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., and Stone, A. (2004). Toward national well-being accounts. *American Economic Review*, 94(2):429–434.

Kamstra, M. J., Kramer, L. A., and Levi, M. D. (2003). Winter blues: A SAD stock market cycle. *American Economic Review*, 93(1):324–343.

Killingsworth, M. A. and Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330(6006):932–932.

Krekel, C., Ward, G., and Neve, J.-E. D. (2019). Employee wellbeing, productivity and firm performance. *CEP Discussion Paper CEPDP1605*.

Krueger, A. B. and Schkade, D. A. (2008). The reliability of subjective well-being measures. *Journal of Public Economics*, 92:1833–1845.

Krueger, A. B. and Stone, A. A. (2014). Progress in measuring subjective well-being: moving toward national indicators and policy evaluations. *Science*, 346(6205):42–43.

Kunin, T. (1955). The construction of a new type of attitude measure 1. *Personnel Psychology*, 8(1):65–77.

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

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *American Economic Review*, 90(2):426–432.

Luttmer, E. (2005). Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics*, 120:963–1002.

Mayo, E. (1993). *The human problems of an industrial civilization. 2. Vol. 3.* MacMillan.

Moen, P., Kelly, E. L., Fan, W., Lee, S.-R., Almeida, D., Kossek, E. E., and Buxton, O. M. (2016). Does a flexibility/support organizational initiative improve high-tech employees' well-being? evidence from the work, family, and health network. *American Sociological Review*, 81(1):134–164.

Moen, P., Kelly, E. L., Lee, S.-R., Oakes, J. M., Fan, W., Bray, J., Almeida, D., Hammer, L., Hurtado, D., and Buxton, O. (2017). Can a flexibility/support initiative reduce turnover intentions and exits? results from the work, family, and health network. *Social Problems*, 64(1):53–85.

OECD (2013). Guidelines on measuring subjective wellbeing. *https://www.oecd.org/statistics/oecd-guidelines-on-measuring-subjective-well-being-9789264191655-en.htm*.

Oswald, A. J., Proto, E., and Sgroi, D. (2015). Happiness and productivity. *Journal of Labor Economics*, 33(4):789–822.

Papke, L. E. and Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145(1-2):121–133.

Roethlisberger, F. J. and Dickson, W. J. (1939). *Management and the Worker*. Harvard University Press.

Rothbard, N. P. and Wilk, S. L. (2011). Waking up on the right or wrong side of the bed: Start-of-workday mood, work events, employee affect, and performance. *Academy of Management Journal*, 54(5):959–980.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–65.

Tenney, E. R., Poole, J. M., and Diener, E. (2016). Does positivity enhance work performance?: Why, when, and what we don't know. *Research in Organizational Behavior*, 36:27–46.

Tice, D. M., Bratslavsky, E., and Baumeister, R. F. (2001). Emotional distress regulation takes precedence over impulse control: If you feel bad, do it! *Journal of Personality and Social Psychology*, 80(1):53.

Walsh, L. C., Boehm, J. K., and Lyubomirsky, S. (2018). Does happiness promote career success? revisiting the evidence. *Journal of Career Assessment*, 26(2):199–219.

Zapf, D., Seifert, C., Schmutte, B., Mertini, H., and Holz, M. (2001). Emotion work and job stressors and their effects on burnout. *Psychology & Health*, 16(5):527–545.

Supplementary Material

Does Employee Happiness Have an Impact on Productivity?

Clement S. Bellet
Erasmus University Rotterdam

Jan-Emmanuel De Neve
University of Oxford

George Ward
Massachusetts Institute of Technology

For Online Publication Only

Appendix A: Predictors of Participation in the Study

Table S1: Predictors of study participation: Extensive Margin

	Participated in the study = 1		
	(1)	(2)	(3)
Age	-0.001	-0.002	-0.003*
	(0.001)	(0.001)	(0.001)
Female	-0.020	-0.018	-0.002
	(0.021)	(0.022)	(0.022)
Left firm during study	-0.151***	-0.157***	-0.131***
	(0.037)	(0.036)	(0.035)
Tenure (months)	-0.032**	-0.032**	-0.021*
	(0.011)	(0.011)	(0.011)
Mean sales during study			-0.001
			(0.002)
Mean selling hours during study			0.010***
			(0.003)
Call center dummies	No	Yes	Yes
Observations	1793	1793	1793
R^2	0.037	0.042	0.058

Notes: Robust standard errors in parentheses, clustered on call centers. Participated =1 if worker responded to at least one survey. 1,438 workers (80.2%) participated in the study.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S2: Predictors of study participation: Intensive Margin

	# Waves responded to survey		
	(1)	(2)	(3)
Age (v. 15-20)			
Age	0.008***	0.008***	0.006***
	(0.002)	(0.002)	(0.002)
Female	-0.115***	-0.107***	-0.069**
	(0.037)	(0.036)	(0.032)
Left firm during study	-0.911***	-0.925***	-0.867***
	(0.047)	(0.046)	(0.035)
Tenure (months)	-0.075***	-0.086***	-0.055***
	(0.022)	(0.021)	(0.018)
Mean sales during study			-0.001
			(0.001)
Mean selling hours during study			0.024***
			(0.004)
Call center dummies	No	Yes	Yes
Observations	1438	1438	1438

Notes: Robust standard errors in parentheses, clustered on call centers. Poisson models reported. Sample is all workers who participated in the study.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix B: Attrition/Non-Response

Table S3: Predictors of attrition/non-response

	Responded to Survey = 1			
	(1)	(2)	(3)	(4)
Sales	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Selling time	0.012***	0.012***	0.012***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)
Bad weather index		0.001		0.001
		(0.002)		(0.002)
Team happiness (excl worker)			-0.001	-0.001
			(0.003)	(0.003)
Observations	33725	33725	33725	33725
R^2	0.436	0.436	0.436	0.436

*Notes: Robust standard errors in parentheses, clustered on individuals. Linear models reported. Unit of observation is worker*week. Individual and week fixed effects in all models.*
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix C: Fine-Grained Estimates using Daily Data

Table S4: Reduced Form Effects of Daily Weather on Sales

	Whole Sample		Non-Missing Happiness Data Only	
	(1)	(2)	(3)	(4)
Bad Weather Index (Daily)	-0.0100**		-0.0103*	
	(0.0044)		(0.0058)	
Bad Weather Index (Weekly)		-0.0071***		-0.0068**
		(0.0020)		(0.0028)
Observations	79,712	80,592	40,131	40,624
Individuals	1,188	1,189	1,184	1,185
Pseudo-R^2	0.473	0.473	0.430	0.431

Notes: Poisson-FE models reported. Robust standard errors in parentheses, clustered on individuals. All models include individual and date fixed effects, along with work schedule controls.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

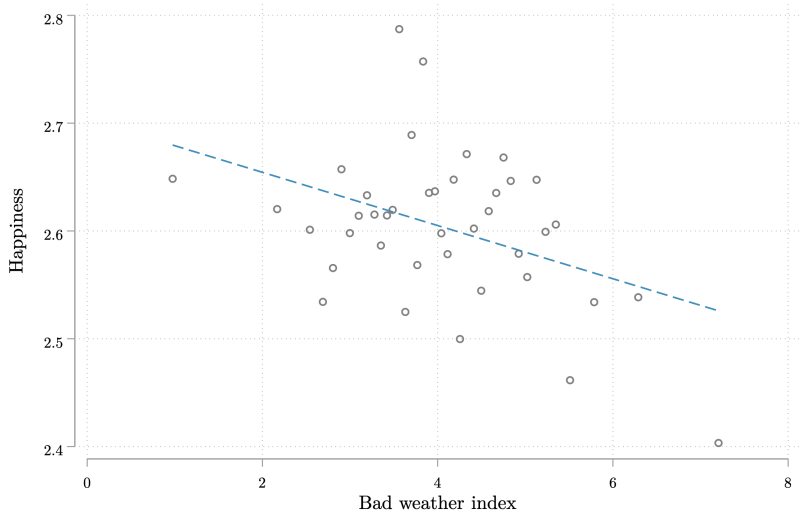
Table S5: Happiness and Sales Performance using Daily Data

	Sales (1) Poisson	(2) Poisson	(3) Poisson	(4) Poisson-IV	(5) Poisson-IV
Very Unhappy	-0.0794***				
	(0.0112)				
Unhappy	-0.0565***				
	(0.0107)				
Happy	0.0332***				
	(0.0107)				
Very Happy	0.0539***				
	(0.0155)				
Happiness (1-5)		0.0363***		0.2608**	
		(0.0036)		(0.1014)	
Happiness (ln)			0.0814***		0.5743**
			(0.0081)		(0.2334)
Observations	40,624	40,624	40,624	40,624	40,624
1st Stage F-Stat				12.59	13.00

Notes: Poisson-FE and Poisson-IV models reported. Robust standard errors in parentheses, clustered on individuals. All models include individual and date fixed effects, work schedule controls, and day of response to survey.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

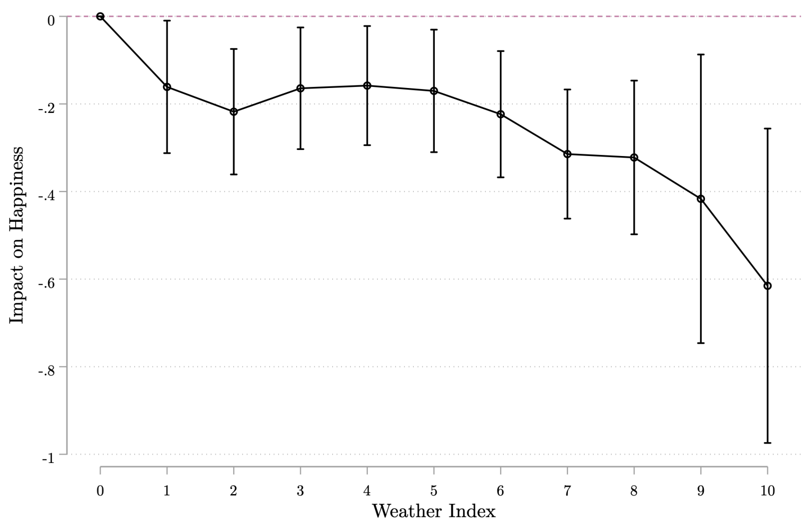
Appendix D: Alternative Instrumental Variable Definitions

Figure S1: Weather IV First Stage



Note: Figure shows the relationship between bad weather and happiness, adjusting for individual and week fixed effects as well as the full set of further controls. Both weather and emotion are first regressed on the full set of fixed effects and controls. The residuals from these regressions are binned across 40 quantiles and plotted as grey dots. The blue line shows the linear fit from an OLS regression using all of the data.

Figure S2: Weather IV First Stage using Bad Weather Index Dummies



Note: Figure reports coefficients and 90% confidence intervals from an OLS regression of sales on a series of dummies for each level of the bad weather index (omitting "zero" as the comparison category). The regression controls for individual and week fixed effects, as well as the full set of controls used in the main analysis.

Table S6: Alternative First-Stage Functional Forms

	IV: Weather^2			IV: Bad Weather =1		
	(1) Red.-Form	(2) 1st Stage	(3) 2nd Stage	(4) Red.-Form	(5) 1st Stage	(6) 2nd Stage
Happiness (1-5)			0.1811**			0.2058**
			(0.0911)			(0.0871)
Weather^2	-0.0004*	-0.0021***				
	(0.0002)	(0.0006)				
Bad Weather = 1				-0.0311**	-0.1527***	
				(0.0140)	(0.0415)	
Observations	12,282	12,282	12,282	12,282	12,282	12,282
1st Stage F-Stat		13.22			13.57	

Notes: Robust standard errors in parentheses, clustered on individuals in models 1, 2, 4 and 5. Bootstrapped standard errors reported for models 3 and 6, resampling across individuals to account for clustering. All models include individual and week fixed effects, work schedule controls, and dummies for day of week of response to survey. Bad weather dummy is equal to 1 if the index is 7 or above, 0 otherwise. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix E: Additional Evidence on IV validity

Table S7: Daily Call-Center-Level Regressions

	Local Demand	Local Attendance	Local Sickness
	(1)	(2)	(3)
Bad Weather Index	-0.1433	-0.1639	-0.1070
	(0.1256)	(0.0931)	(0.0920)
Observations	1,415	1,428	1,428
Call Centers	11	11	11
R^2	0.834	0.664	0.335

Notes: OLS-FE models reported. Robust standard errors in parentheses, clustered on call centers. All models include call center and date fixed effects. Local demand is the mean number of calls per employee-day in each call center. Local attendance and local sickness are the call-center mean.

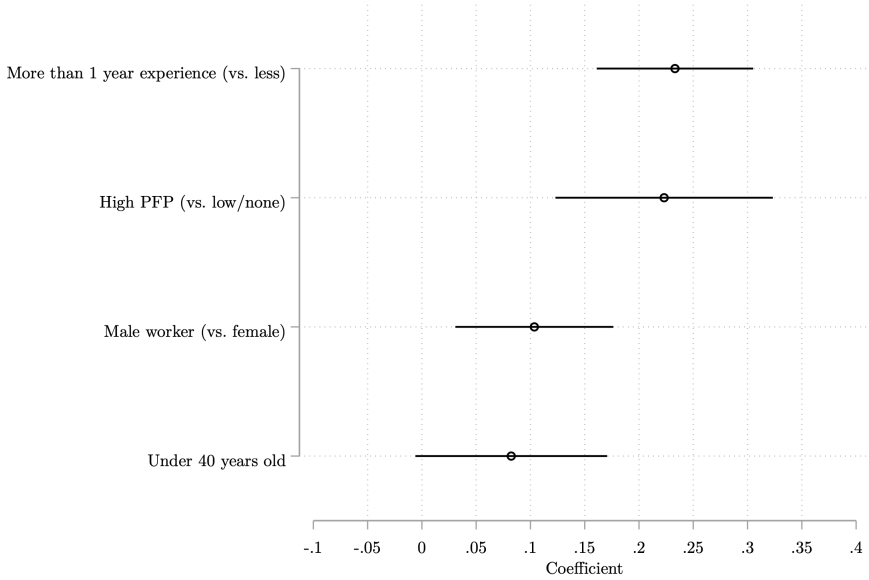
Table S8: Bad Weather Index and Happiness: Heterogeneous Treatment Effects

	Happiness					
	(1)	(2)	(3)	(4)	(5)	(6)
Bad Weather Index	-0.0247***	-0.0175*	-0.0129	-0.0263***	-0.0311***	-0.0284***
	(0.0076)	(0.0093)	(0.0236)	(0.0092)	(0.0097)	(0.0109)
Bad Weather Interaction:						
x Female Worker		-0.0180				
		(0.0143)				
x Worker's Age			-0.0003			
			(0.0006)			
x Worker's Tenure (Years)				0.0003		
				(0.0008)		
x Worker Avg Happiness > Median					0.0135	
					(0.0136)	
x # Sales						0.0002
						(0.0003)
Observations	12,282	12,282	12,282	12,282	12,282	12,282
Individuals	1,157	1,157	1,157	1,157	1,157	1,157
R^2	0.534	0.534	0.534	0.534	0.534	0.536

Notes: OLS-FE models reported. Robust standard errors in parentheses, clustered on individuals. All models include individual and week fixed effects, work schedule controls, and day of response to survey. Column (6) also controls for weekly sales as it varies within workers.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix F: Benchmarking of Magnitudes

Figure S3: Causal Effect of Happiness: Effect Size Comparisons



Note: Coefficients and 95% confidence intervals reported from a cross-sectional Poisson model. Dependent variable is the weekly number of sales. The model includes controls for the working hours, internal training time, and study week fixed effects.

Appendix G: Additional Evidence of Downward Bias in OLS

Table S9: Happiness and Number of Calls

	Happiness (1-5)	Happiness (ln)	Sales
	(1)	(2)	(3)
	OLS-FE	OLS-FE	Poisson-FE
Total number of weekly calls (ln)	-0.2087***	-0.0825***	0.0837**
	(0.0427)	(0.0171)	(0.0420)
Observations	12,282	12,282	12,282
Employees	1,157	1,157	1,157
R^2	0.535	0.538	
Pseudo-R^2			0.620
Week FEs	✓	✓	✓
Individual FEs	✓	✓	✓

Notes: Linear models reported in columns (1) and (2). Poisson model reported in column (3). Robust standard errors in parentheses, clustered on individuals. All models include controls for the working hours, internal training time, and day of week dummies for response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix H: Suggestive Evidence on Psychological Mechanisms

Table S10: Happiness and Performance-Related Pay

	# Sales			
	(1)	(2)	(3)	(4)
Happiness (1-5)	0.0352***	0.0347***		
	(0.0034)	(0.0070)		
Happiness × High Performance Pay		0.0006		
		(0.0079)		
Bad Weather Index			-0.0059**	-0.0167***
			(0.0026)	(0.0047)
Bad Weather Index × High Performance Pay				0.0146***
				(0.0050)
Observations	12282	12282	12282	12282
Individuals	1,157	1,157	1,157	1,157
Pseudo-R^2	0.621	0.621	0.620	0.620

Notes: Poisson-FE models reported. Robust standard errors in parentheses, clustered on individuals. All models include individual and week fixed effects, work schedule controls, and day of response to survey.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S11: Happiness and Multi-Tasking

	# Sales	
	(1)	(2)
Happiness (1-5)	0.0373***	
	(0.0040)	
Happiness (1-5) × Other Tasks (% Working Hours)	-0.0004	
	(0.0004)	
Bad Weather Index		-0.0090***
		(0.0030)
Bad Weather Index × Other Tasks (% Working Hours)		0.0006**
		(0.0003)
Other Tasks (% Working Hours)	-0.0003	-0.0033***
	(0.0013)	(0.0011)
Observations	12,255	12,255
Individuals	1,155	1,155
Pseudo-R^2	0.620	0.618

Notes: Poisson-FE models reported. Robust standard errors in parentheses, clustered on individuals. All models include individual and week fixed effects, work schedule controls, and day of response to survey.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S12: Happiness and Emotional Skills

	# Sales	
	(1)	(2)
Happiness (1-5)	0.0354***	
	(0.0034)	
Happiness (1-5) × Weekly Customer Satisfaction	-0.0043	
	(0.0027)	
Bad Weather Index		-0.0043*
		(0.0026)
Bad Weather Index × Weekly Customer Satisfaction		0.0105***
		(0.0019)
Observations	12,282	12,282
Individuals	1,157	1,157
Pseudo-R^2	0.621	0.620

Notes: Poisson-FE models reported. Robust standard errors in parentheses, clustered on individuals. All models include individual and week fixed effects, work schedule controls, and day of response to survey.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix I: Customer Satisfaction

Table S13: Happiness and Customer Satisfaction

	Customer Satisfaction (z-score)			
	(1) OLS-FE	(2) OLS-FE	(3) 2SLS	(4) 2SLS
Happiness (1-5)	0.0169*	0.0141	-0.3807	-0.3880
	(0.0094)	(0.0094)	(0.3177)	(0.3220)
# Sales		0.0030***		0.0067**
		(0.0011)		(0.0031)
Observations	10,059	10,059	10,059	10,059
Employees	1,081	1,081	1,081	1,081
R^2	0.173	0.174	-0.177	-0.178

Notes: Columns (1) and (2) are OLS-FE models. Column (3) and (4) report Second-stage 2SLS regressions. Robust standard errors in parentheses, clustered on individuals. Weekly customer satisfaction is z-scored. All models include individual and week fixed effects, work schedule controls, and day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix J: Production Function Approach

Table S14: Production Function using Weather IV

	Sales (Poisson-FE)		ln(Happiness) (OLS-FE)	Sales (Poisson-FE)
	(1) Non-IV	(2) Red.-Form	(3) 1st Stage IV	(4) 2nd Stage IV
Happiness (ln)	0.0783***			0.5204**
	(0.0077)			(0.2288)
Bad Weather Index		-0.0059**	-0.0116***	
		(0.0026)	(0.0034)	
Observations	12,282	12,282	12,282	12,282
1st Stage F-Stat			11.84	

Notes: Columns (1) and (3) are Poisson-FE models. Column (2) reports an OLS-FE regression. Column (3) reports the second stage of a Poisson-FE control function model. Robust standard errors in parentheses, clustered on individuals in models (1) and (2). Bootstrapped standard errors reported for model (3) - see text for detail. All models include individual and week fixed effects, work schedule controls, and day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S15: Production Function: Happiness and Labor Supply

	Attendance	Overtime	Paid Vacation	Paid Breaks	Unpaid Breaks
	(100% =1)	(Any = 1)	(Any = 1)	(Minutes)	(Minutes)
Happiness (ln)	-0.1397	-0.1816	-0.2942	-9.9953	-40.4520
	(0.3322)	(0.1551)	(0.2716)	(6.4319)	(44.0207)
Observations	12,279	12,282	12,282	12,282	12,282

Notes: Second-stage 2SLS models reported, using adverse weather index as an IV for (log) happiness. Robust standard errors in parentheses, clustered on individuals. All models include include individual and week fixed effects and dummies for day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S16: Production Function: Happiness and Labor Productivity

	Adherence (Met Target=1)		Minutes Per Call (Log)		Conversion Rate (Log)	
	(1)	(2)	(3)	(4)	(5)	(6)
Happiness (ln)	0.5579**	0.5598**	-0.3894***	-0.4069**	1.1283***	1.3529***
	(0.2726)	(0.2746)	(0.1500)	(0.1707)	(0.4137)	(0.4391)
Sales (ln)				0.0373**		
				(0.0185)		
Minutes per call (ln)						0.6776***
						(0.1006)
Observations	12,169	12,169	12,100	11,979	11,720	11,720

Notes: Second-stage 2SLS models reported, using adverse weather index as an IV for (log) happiness. Robust standard errors in parentheses, clustered on individuals. All models include individual and week fixed effects, and controls for log working hours, internal training time, and day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix K: Heterogeneity by Demographics

Table S17: Happiness and Sales Performance: Heterogeneity

	# Sales				
	(1)	(2)	(3)	(4)	(5)
Happiness (1-5)	0.0352***	0.0312***	0.0483***	0.0368***	0.0341***
	(0.0034)	(0.0042)	(0.0107)	(0.0042)	(0.0050)
Happiness (1-5) × Female Worker		0.0115*			
		(0.0067)			
Happiness (1-5) × Worker's Age			-0.0004		
			(0.0003)		
Happiness (1-5) × Tenure (Years)				-0.0004	
				(0.0005)	
Happiness (1-5) × Worker Avg Happiness > Median					0.0020
					(0.0066)
Observations	12,282	12,282	12,282	12,282	12,282
Individuals	1,157	1,157	1,157	1,157	1,157
Pseudo-R^2	0.621	0.621	0.621	0.621	0.621

Notes: Poisson-FE models reported. Robust standard errors in parentheses, clustered on individuals. All models include individual and week fixed effects, work schedule controls, and day of response to survey.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1654	Matej Bajgar Giuseppe Berlingieri Sara Calligaris Chiara Criscuolo Jonathan Timmis	Industry Concentration in Europe and North America
1653	Andrés Barrios Fernandez	Should I Stay of Should I Go? Neighbors' Effects on University Enrollment
1652	Emanuel Ornelas Laura Puccio	Reopening Pandora's Box in Search of a WTO-Compatible Industrial Policy? The Brazil – Taxation Dispute
1651	Giulia Giupponi	When Income Effects are Large: Labor Supply Responses and the Value of Welfare Transfers
1650	Jonas Jessen Sophia Schmitz Sevrin Waights	Understanding Day Care Enrolment Gaps
1649	S. Federico F. Hassan V. Rappoport	Trade Shocks and Credit Reallocation
1648	Max Nathan	Does Light Touch Cluster Policy Work? Evaluating the Tech City Programme
1647	Stuart Campbell Lindsey Macmillan Richard Murphy Gill Wyness	Inequalities in Student to Course Match: Evidence from Linked Administrative Data
1646	Cong Peng	Does E-Commerce Reduce Traffic Congestion? Evidence from Alibaba Single Day Shopping Event

1645	Dan Andrews Chiara Criscuolo Peter N. Gal	The Best versus the Rest: Divergence across Firms during the Global Productivity Slowdown
1644	Christopher Cornwell Ian M. Schmutte Daniela Scur	Building a Productive Workforce: The Role of Structured Management Practices
1643	Paul Dolan Georgios Kavetsos Christian Krekel Dimitris Mavridis Robert Metcalfe Claudia Senik Stefan Szymanski Nicolas R. Ziebarth	Quantifying the Intangible Impact of the Olympics Using Subjective Well-Being Data
1642	Xavier Jaravel Erick Sager	What are the Price Effects of Trade? Evidence from the US and Implications for Quantitative Trade Models
1641	Johannes Boehm Jan Sonntag	Vertical Integration and Foreclosure: Evidence from Production Network Data
1640	Teodora Borota Fabrice Defever Giammario Impullitti	Innovation Union: Costs and Benefits of Innovation Policy Coordination
1639	Monica Langella Alan Manning	Residential Mobility and Unemployment in the UK
1638	Christos Genakos Mario Pagliero	Competition and Pass-Through: Evidence from Isolated Markets
1637	Holger Breinlich Elsa Leromain Dennis Novy Thomas Sampson	Voting With Their Money: Brexit and Outward Investment by UK Firms

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE