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

July 20, 2022

Abstract

This paper provides evidence from a natural experiment on the relationship between positive affect and productivity. We link highly detailed administrative data on the behaviors and performance of all telesales workers at a large telecommunications company with survey reports of employee happiness that we collected on a weekly basis. We leverage variation in a worker's visual exposure to weather while at work in order to derive quasi-experimental estimates for the causal effect of happiness on productivity. We find a strong effect of happiness on sales performance, which is driven by changes in labor productivity – largely through workers converting more calls into sales, and to a lesser extent by making more calls per hour and adhering more closely to their schedule. We find no evidence in our setting of effects on measures of high-frequency labor supply such as attendance and break-taking.

*We thank Alex Bryson, Ed Diener, Alex Edmans, Paul Frijters, Sergei Guriev, John Helliwell, Micah Kaats, Caspar Kaiser, Erin Kelly, Tom Kochan, Armando Meier, Rob Metcalfe, Mike Norton, Paul Osterman, Alberto Prati, Alexandra Roulet, Mark Stabile, Anna Stansbury, Andrew Stephen, John Van Reenen, Ashley Whillans, and Nate Wilmers for helpful comments on manuscript drafts. We also thank seminar participants at Oxford, MIT, LSE, 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. De Neve is a research advisor to Butterfly AI. Corresponding author: George Ward, MIT Sloan, 100 Main Street, Cambridge MA. Email: wardg@mit.edu.

1 Introduction

A large number of employers are increasingly claiming to care about how their employees feel at work, and have begun to invest in management practices and services aimed at creating and maintaining a happier workforce. There may be various reasons for this – such as an increased ability to attract and retain high quality workers – but at least one motivation is a belief that happier workers will be more productive. When surveyed, for example, around 79% of U.S. managers reported an expectation that unhappiness in their workforce is likely to hurt productivity.[1] While this recent focus on employee happiness may seem a relatively new development, the relationship between happiness and productivity has in fact been the keen focus of both researchers and practitioners for many decades (see, e.g., Fisher and Hanna, 1931; Hersey, 1932, for early examples). One reason for this long-running interest is that the relationship has potentially key implications for how firms manage their workers and organize work (Edmans, 2012; Wright and Cropanzano, 2000). However, isolating the causal effect of happiness on productivity has remained an empirical challenge, particularly in field settings. Indeed, despite multiple generations of research on the topic, the literature has been bedeviled by inconsistent findings (Iaffaldano and Muchinsky, 1985; Judge et al., 2001; Tenney, Poole and Diener, 2016), such that claims surrounding the organizational benefits of worker happiness have frequently been met with skepticism (Wright and Cropanzano, 2000).

In this paper, we study the effects of employee positive affect on productivity. We use data on the universe of telesales workers at British Telecom (BT), allowing us to observe objective, granular information about the behaviors and performance of 1,793 workers at one of the United Kingdom's largest private employers. We link this administrative data to a survey instrument designed to measure the week-to-week affect of employees using a well-established survey measure of happiness. We use variation in exposure to visual weather conditions while at work, arising out of the interaction between weather and architecture, in order to provide a quasi-experimental test of the effect of happiness on sales performance.

We find that a one unit increase in happiness, on a standard 0-to-10 scale, leads to around 3 additional weekly sales, around a 12% increase over a base of 25. Using detailed data on worker behaviors, we rule out any effect going through employees working more (e.g. by taking fewer breaks or working overtime). Instead, we are able to study three potential productivity channels behind the effect of happiness on sales performance. First, workers could be better at organizing their time while happier at work and in doing so adhere more closely to their prescribed daily workflow schedule. Second, they could work faster when happier – that is, answer a higher number of calls per hour. Third, they could be more efficient at converting calls into sales when in a better mood. Although we find evidence for all three potential channels, the magnitude of the third channel is much stronger, such that the estimated effect on sales can be almost entirely explained by workers converting more of their calls into sales during weeks when they feel happier. We interpret this as suggesting that, at least in our context, much of the effect of happiness on productivity can be explained by better moods augmenting social and emotional

[1]The survey included 1,073 managers from a wide range of sectors and organization sizes, but was not a randomly drawn sample (see HBR Analytical Services, 2020, for a more extensive discussion of the survey and the beliefs and attitudes of managers in relation to employee wellbeing).

skills. In line with this, when breaking down the analysis by sales type, we find strong effects of happiness on sales when the worker is selling bundles of products and, in particular, when re-contracting, but negligible effects when doing more routine order taking. Moreover, we find suggestive evidence that the effect of employee happiness on sales is stronger when workers are more likely to be interacting on the phone with angry or dissatisfied customers.

Our identification strategy relies on four key features of the empirical setting. First, the 11 call centres are dispersed geographically across the whole of the United Kingdom, such that there is significant variation, within week, in weather in the local vicinity of each of the call centers. Second, workers take incoming calls that are allocated to them based on the call-type and handler availability, but not based in any way on the location of the caller – meaning that, in this two-sided market, workers based in a given location will take calls from customers all over the country. Third, the call centers vary significantly in terms of their architecture and, in particular, their window coverage – ranging from fully glass-covered tower buildings all the way to warehouse-style buildings with almost no windows at all. Finally, despite these major exterior differences, all are laid out internally in the same open-plan way, such that any given worker within a call center has roughly equal visual exposure to the external walls – be they fully windowed, solid, or anywhere in between. We hypothesize that only the most psychologically-relevant features of weather—namely, the extent to which it is bright or gloomy—will impact worker happiness. Such aspects are *visual* in nature, meaning that any effect should be contingent on a worker's observable exposure to them. We show that visibly gloomy weather has a strong negative impact on our measure of worker mood, on average, but that this effect is dependent on the window coverage of each call center – in other words, workers' exposure to it. We exploit this plausibly exogenous variation in happiness to provide a quasi-experimental test of the effect of happiness on sales performance in a real-world field setting.

We make three main contributions. First, we provide causal field evidence on the relationship between happiness and productivity. As we noted at the outset, in doing so we build on an extensive literature that spans many decades as well as several disciplines across the social and behavioral sciences. But despite a long line of research, the literature has been plagued with inconsistencies (Tenney, Poole and Diener, 2016). Although experimentally-induced positive affect has been shown to improve performance on stylized productivity tasks in the lab (see, e.g., Erez and Isen, 2002; Oswald, Proto and Sgroi, 2015), it remains unclear whether such effects translate to real-world, large-scale organizational settings – which are typically very different along a number of different dimensions.

Second, we make a methodological contribution in terms of how we identify mood effects in the field. While affective states are increasingly seen as a potentially important factor in driving economic behavior (Loewenstein, 2000), demonstrating this in natural settings has proven difficult. We join a growing literature that typically proceeds by estimating the effects of various proxies for mood (such as weather patterns) on outcomes like stock returns, consumption, real estate transactions, and voting (e.g. Agarwal et al., 2020; Edmans, Garcia and Norli, 2007; Hirshleifer and Shumway, 2003; Hu and Lee, 2020; Li et al., 2017; Meier, Schmid and Stutzer, 2019; Saunders, 1993). We label this existing method the reduced-form approach to mood effects in the field, and build on it in three ways. First, by studying participants in a two-sided market who are located in different places, we are able to identify weather-induced worker mood effects aside from overall shocks that may, for example, affect national demand or customer mood. Second, implicit to the reduced-form approach is an instrumental variables (IV) set up, whereby weather i) has an impact on mood and ii) affects behavior solely through that mood mechanism. However, whereas this is often assumed rather than fully estimated, we follow Guven (2012) in collecting a measure of happiness and estimating both stages of the IV specification (see also Coviello et al., 2020; Guven and Hoxha, 2015). This is particularly important since the prior literature on weather and happiness suggests it is not necessarily a given, casting doubt on the implied first stage (Feddersen, Metcalfe and Wooden, 2016; Frijters, Lalji and Pakrashi, 2020). Third, an ever-growing number of economic phenomena have been instrumented for using weather patterns, casting doubt on the validity of the exclusion restriction (Gallen, 2020; Mellon, 2020). Our novel use of variation in architecture coupled with a focus on only the visual and mood-relevant aspects of weather allows us to overcome many of the usual objections related to this assumption. We rely on detailed institutional knowledge of our workplace setting, gleaned from a mixture of image coding, employee surveys, semi-structured interviews with managers, and site visits. This mixed methods approach means we can design an identification strategy in which differential visual access to outside weather patterns allows us to "turn on and off" the treatment – and in doing so better isolate any causal effect.

Given that firms are increasingly at least claiming to focus on the happiness of their employees, the third main contribution of our paper is to provide more fine-grained and potentially useful evidence for managers on the source of the happiness-productivity relationship. A recent survey of U.S. executives suggests, for example, that while a large proportion of U.S. firms say they are considering investing practices designed to foster a happier workforce, only a small number currently have any sort of strategy in place to move in this direction (HBR Analytical Services, 2020).[2] Since work is one of the most unhappy activities people do in their day-to-day lives in countries like the USA and UK (Bryson and MacKerron, 2016; Krueger et al., 2009), there is ample room for improvement that could unlock potential productivity gains. In one sense, our empirical approach suggests the significance of a managerially important but often overlooked aspect of work life – space and the physical workplace environment. But, even more importantly, a long line of research has found a range of management and organizational practices influence the happiness of workers – suggesting a large annumber of potential levers for firms to pull.[3] By probing mechanisms through which happiness may impact productivity, we are able to discuss the types of tasks and jobs where happiness is most likely to be an important factor in explaining productivity differentials. In particular, with the number of jobs requiring workers to interact socially with customers increasing rapidly (Deming, 2017), our finding that

[2] 87% of executives agreed that workplace happiness can provide their firm with a competitive advantage, but only a third of the organizations in the survey of executives noted above say their organization sees employee wellbeing as a strategic priority. Not only this, fewer than 20% of these firms actually have any sort of strategy in place to measure or improve the wellbeing of their workforce.

[3] While evidence on the effectiveness of employee "wellness programs" is mixed (Gubler, Larkin and Pierce, 2018; Jones, Molitor and Reif, 2019), a growing literature demonstrates the more fundamental point that the ways in which work is managed and organized by firms—as well as the cultures they create—has a significant impact on employee wellbeing (Bloom et al., 2014; Clark, 2010; Gosnell, List and Metcalfe, 2020; Krekel, Ward and De Neve, 2019; Moen et al., 2016).

much of the happiness effect can be traced to improved social and emotional skills suggests that the importance of employee happiness in driving productivity growth is likely to rise in the coming years.

2 Background & Theory

For over a century, the relationship between happiness and productivity has been the keen focus of both researchers and practitioners (see, e.g., Fisher and Hanna, 1931; Hersey, 1932, for early examples). One reason for this long-running interest is that the relationship has potentially key implications for how firms manage their workers and organize work – as well as for the place, more broadly, of human resource management in firms' overall business strategy (Edmans, 2012; Wright and Cropanzano, 2000). Yet, despite a long line of research, the literature is bedeviled by inconsistent findings – many of which may be traceable to i) the multiplicity of ways that researchers approach—theoretically and empirically—the concepts of both happiness and performance, as well as ii) a number of inherent empirical difficulties in estimating any causal effect, particularly in natural field settings.

2.1 Happiness

Although the study of human happiness has a long history (see Diener et al., 1999), confusion can arise insofar as the term "happiness" is sometimes used loosely as a catch-all term referring to subjective wellbeing (SWB) – which has both affective and cognitive dimensions (Krueger et al., 2009). Cognitive measures of SWB are evaluative in nature, refer to global judgements people make about how things are going overall, and are typically assessed in the workplace context using survey questions on job satisfaction. Affective SWB, on the other hand, refers to people's emotional or hedonic experience. Kahneman, Wakker and Sarin (1997) refer to this as 'experienced utility,' in the Benthamite tradition, and note that it can be measured either in real time or via people's recollections – for example, by asking how happy they feel or have felt during a given day, week, or month.

One potential reason for the highly inconsistent nature of the literature on happiness and productivity is that researchers have approached the question often with differing notions of what "happiness" means (Wright and Cropanzano, 2000). While much of the early literature focused on measures of job satisfaction (e.g., Brayfield and Crockett, 1951; Fisher and Hanna, 1931; Lawler and Porter, 1967; Locke, 1969), a more recent body of work has generally turned toward the study of affect in the workplace (see, e.g., Barsade and Gibson, 2007; Brief and Weiss, 2002; Knight, Menges and Bruch, 2018). This distinction is particularly important as the theoretical links between happiness and performance are, as we will discuss below, arguably much stronger when thinking in terms of affect than satisfaction (Côté, 1999; Lucas and Diener, 2003).

In this paper, we focus on affective well-being – specifically, workers' feelings of happiness as they experience it week-to-week. We see this as a general measure of positive affect that can also reasonable be referred to as "mood." Within the broad category of affect or affective wellbeing, there is a key distinction between moods and emotions (see, e.g. Frijda, 1986). Emotions typically refer to a specific feeling that is a (relatively short-lived) reaction to a particular (and usually known) stimulus. Moods, on the other hand, are less specific and are typically less intense. They are not directed at a particular person, task, or situation, but are rather a more diffuse general feeling. While it is often easy for people to trace the root of a particular emotion, they are usually not aware of the source of a good or bad mood (Russell and Barrett, 1999). Given this, there is little reason to expect the effects of a weather-induced good or bad mood to be any different from the effects of a mood state induced by other factors, ranging from working arrangements to line manager behavior. This is important, since although we use weather-induced mood shocks for identification, we are nevertheless able say something more broadly about managerial and policy implications.

In addition to being influenced by a range of management and organizational factors (Gosnell, List and Metcalfe, 2020; Krekel, Ward and De Neve, 2019), subjective wellbeing is also influenced by the weather – one of the most pervading background variables in human life. However, the literature suggests that the link is not straightforward (Feddersen, Metcalfe and Wooden, 2016; Frijters, Lalji and Pakrashi, 2020), since the empirical relationship between weather and affect can be unstable (Denissen et al., 2008). A key explanation for the instability of these findings is that effects are contingent on the extent to which weather is visible. Keller et al. (2005) find weather affects mood, for example, but only when people are experimentally assigned to be outdoors. A related literature shows that the visual pleasantness of weather improves mood and prosocial behavior in studies with outdoor settings (Cunningham, 1979), though this does not replicate using time-series data on tipping in an indoor restaurant (Flynn and Greenberg, 2012).[4] Overall, one of the key findings of this literature is that the visibility of weather is key to any relationship with subsequent mood and behavior.[5] Given this contingent relationship, in order to convincingly use weather as a mood proxy or instrument, it is (a) useful to have variation in weather that is *visual* in nature and (b) even more useful to have variation in visual *exposure* to any given weather, in order to eliminate concerns related to the direct effect of weather on non-mood related drivers of productivity.

2.2 Performance

In addition to confusion surrounding conceptual definitions and measurement of happiness, a further potential explanation for the inconsistent state of the literature is that many different definitions of the performance *outcome* have been studied. Here we study labor productivity – that is, the residual variation in output that cannot be fully explained by observable inputs (Syverson, 2011). Call centers provide a particularly good setting in which to study this, since we not only observe detailed data on a large number of labor inputs, but workers in our setting also in the large part do the same telesales job, which is to take incoming calls from new and

[4] Rind (1996) also studies an indoor setting – a casino hotel in Atlantic City – in which hotel rooms all have dark, limo-tinted windows that make it look like it is cloudy outside regardless of the brightness of the weather outside. Having a server inform customers of the weather outside, experimentally varying how bright it is reported to be, the authors find that (a belief in) sunny skies increases tipping. That is, though this experiment does not vary visual exposure to weather, it does vary a related concept – the salience.

[5] This is in line with a large medical literature on seasonal affective disorder (SAD), which shows that experimental exposure to sunlight improves mood (e.g. Kripke, 1998), even among the non-depressed (e.g. Leppämäki, Partonen and Lönnqvist, 2002).

existing customers and sell them various products using the same phones and computer system.

A large number of existing studies rely on employee self-reports of productivity or subjective managerial evaluations (e.g. Staw and Barsade, 1993; Zelenski, Murphy and Jenkins, 2008). In the case of using self-reported performance, there are well-known empirical difficulties associated with regressing one subjective report on another, and in the case of managerial reports there is a strong possibility that performance will be subject to a 'halo' effect whereby the happier employee is rated more highly by the managerial rater precisely because they are happier and more agreeable, and not because of any real performance differences. Rather than having to rely on subjective outcome measures, we are able to use administrative data on a clear and objectively measurable output that is unambiguously positive for the firm: sales performance. Moreover, fine-grained objective data on worker behaviors also allows us to investigate channels through which any effect of happiness may translate into sales.

Given the various ways in which workers' mood may affect performance, having a clear understanding of which performance indicator is being used in any specific context is especially important. For instance, a task involving complex interactions with coworkers or customers may involve different skills than a more autonomous and repetitive task. The most recent field-experimental literature shows that management practices can have simultaneously positive impacts on i) productivity as well as ii) employee happiness and satisfaction for both low-skilled and high-skilled employees (Gosnell, List and Metcalfe, 2020). From pay inequality (Breza, Kaur and Shamdasani, 2017; Cullen and Perez-Truglia, 2019) to gift exchange (DellaVigna et al., 2020) and work autonomy (Bloom et al., 2014), this line of research suggests employee wellbeing as one possible channel through which workplace organization may feed through to productivity. However, it is unable to isolate the happiness channel in a causal manner. As a result, little discussion has been paid to the types of tasks or psychological channels through which improvements in worker well-being itself may mediate the positive effects of managerial practices on productivity.

2.3 Happiness and Performance

While much of the earlier work on happiness and workplace performance is based on cross-sectional comparisons, more recent studies have been able to leverage within-worker variation in longitudinal research designs that go some way to assuage concerns related to unobserved heterogeneity between workers, and in doing so has demonstrated that prior affective states predict subsequent performance (see, e.g. Koys, 2001; Miner and Glomb, 2010; Rothbard and Wilk, 2011; Staw and Barsade, 1993; Staw, Sutton and Pelled, 1994).[6] But although this temporal ordering is consistent with a (Granger) causal effect, it still may be the case that time varying third factors could be driving both. To infer causality, Oswald, Proto and Sgroi (2015) rely on a series of mood-inducement experiments in the lab,[7] using as their outcome an incentivized, math-based productivity task designed to be similar to something one might find

[6] For a much fuller treatment of this large literature, which we cannot review exhaustively here, see Walsh, Boehm and Lyubomirsky (2018) and the citations within.

[7] The authors randomly assign subjects to watch comedy or neutral video clips in their main experiment, and in a further test use a mood treatment involving fruit, chocolate, and bottled drinks. Finally, they also leverage recent real-world shocks like bereavement and family illnesses, which have an impact on happiness.

in a real-world work setting – and, ultimately, find a positive effect of happiness (see also Erez and Isen, 2002, for similar findings using a non-incentivized task).

Yet, while this evidence using a stylized, piece-rate productivity task moves the literature forward, the extent to which these affect-induced productivity effects translate into real-world employment settings remains an open question. Indeed, real-world jobs typically involve bundles of tasks, decisions on how to focus time and energy between them, as well as a number of other factors that make it difficult to generalize from a math task in the lab to a real job in the field. The external validity of existing lab-evidence can therefore be built upon, particularly for tasks involving social interactions with customers or coworkers – which are becoming ever more common in modern economies (Deming, 2017). In such cases, the positive effect of mood on performance found in a constrained lab-environment may either be reduced (e.g. if happier workers spend less time at work and more time socializing with co-workers) or magnified (e.g. if happier workers are better able to deal with customers). It is therefore important to be more specific about the exact mechanisms through which positive affect may lead to greater performance (for extensive discussion of the theoretical mechanisms between happiness and performance, see, Lucas and Diener, 2003; Tenney, Poole and Diener, 2016).

First, good mood may affect performance by augmenting cognitive skills, independently of the social interactions in which workers might be involved. In particular, positive affect can influence how we think and process information. The influential broaden-and-build framework suggests that positive affect signals to people experiencing it that the environment is non-threatening and that things are generally going well – as a result, positive mood states tend to broaden people's thought-action repertoires as well as allow them to build longer-lasting resources (Fredrickson, 2001). In line with this, laboratory evidence suggests that people induced into positive mood states tend to think in ways that are more flexible (Isen and Daubman, 1984), creative (Isen, Daubman and Nowicki, 1987), integrative (Isen, Rosenzweig and Young, 1991), open to information (Estrada, Isen and Young, 1997), and efficient (Isen and Means, 1983). Relatedly, it has been shown that the thoughts of happier people are less 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, Proto and Sgroi, 2015).

Second, a further potential channel for the happiness-performance link is that happier workers may be more motivated (Erez and Isen, 2002). People in a positive mood state may have greater prior expectations about the task, for example, and since people experiencing positive affect are more likely to expect to enjoy an upcoming task, they are likely to be more motivated to initiate or engage with it. Moreover, happier people might attribute their positive mood state to the task at hand, and in doing so make them feel that they are enjoying the task – and, ultimately, make them want to complete it (Forgas, 1995).

Third, positive affect may also influence behavior and outcomes by improving people's social and emotional skills. Outward indicators of happiness such as laughter signal that a person is friendly and open. People induced into positive mood states (in the lab) are more likely to engage in social contact with others (Isen, 1970) and people in happier moods tend to be more cooperative and less aggressive with others (Isen and Baron, 1991). The positive impact of affect may hence be even stronger when the task involves interacting directly with customers or coworkers. For instance, Carnevale and Isen (1986) show that positive affect improves bilateral negotiation skills in bargaining tasks, with participants employing less contentious tactics and also finding more integrative solutions. Related to this, the sociological literature on emotional labor (see, e.g., Hochschild, 1983) has long argued that in tasks involving interactions with customers, happier workers may be better at negotiating with angry customers, as it becomes easier to manage their emotions.

Given the differences in these psychological mechanisms, the extent to which affective happiness is likely to lead to greater productivity in real-world jobs will depend on a number of factors, including the type of tasks and the nature of the workplace in question. For example, greater creativity or being more integrative in thought (the more cognitive mechanisms) may be helpful to productivity in jobs where novel or more complex solutions to specific problems are performance-boosting, but less so where performance is best accomplished by following precise instructions or routines. However, to the extent that positive affect boosts motivation, task persistence, and energy, we may still expect it to lead to greater productivity in the latter cases. Equally, since happiness leads to greater social rewards it may be a particularly strong predictor of productivity in jobs where employees frequently interact with others, where social contact is beneficial, and where cooperation or negotiation with others is required in order to be successful. But in jobs where social contact is less integral to the successful carrying out of the job, or where there is little supervision or structure, happiness may be less important – or may even reduce productivity by providing a distraction. Improved sociability may even hurt performance if it leads workers to spend more time with co-workers doing unproductive tasks.

3 Institutional Setting & Data

Administrative Firm Data. We use detailed individual-level administrative data from BT, a large multinational telecommunications company based in the United Kingdom. We focus our attention on sales workers, whose job it is to take incoming calls and sell BT products, at 11 call centers across the United Kingdom (see Figure 1 for a map). 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 letters, online customer chats, and SMS messaging). Workers are paid a fixed hourly wage, with a potential bonus if they meet their target.[8] At the worker-day level we observe the number of sales, the distribution of which can be seen in panel (a) of Figure 2. As is typically the case with sales data, the distribution is both right-skewed and also contains some zero values. In addition, we also observe a host of information about worker behaviors such as break-taking, attendance, number of calls, average length of calls, and so on.

[8]This is neither an explicit piece-rate pay schedule nor is it a commission-based pay system. In each of these instances, one might expect each individual sale to bring with it a psychological reaction. Rather, the pay system is a much more slow-moving bonus scheme, in which the majority of pay is paid through a base salary.

Figure 1: Spatial Distribution of Call Centers

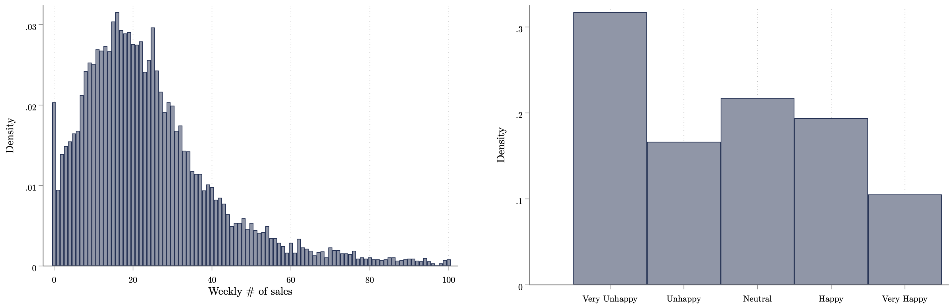
Note: Map shows the location of the 11 call centers in the study, as well as the number of sales workers in each.

Affective Well-Being Survey. We link this performance data at the worker-week level with a survey we administered to capture positive affect. We use a succinct happiness question that was designed following the OECD's guidelines on the measurement of SWB (see OECD, 2013). Employees were asked "*Overall, how happy did you feel this week?*" over a six-month period. This is an affective question, and, following Kunin (1955) and decades of subsequent work in psychology, we offer five response categories as a Faces Scale that ranges from very sad to very happy androgynous faces. 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). The survey was sent by email every week on Thursday afternoon, and is shown in Figure 3. The single-item survey question could be quickly answered within the email. Workers were assured that their individual happiness responses were being collected externally for the sole purpose of academic research, and would not be shared with management. Workers were also offered the opportunity to opt-out of the study at any time, via a simple email click-through. The study ran for 6 months – the first and final emails were sent on the 20^{th} of July 2017 and the 18^{th} of January 2018, respectively.

Benjamin et al. (2021) note the importance of being explicit about what happiness researchers are measuring. It is not only important in terms of wellbeing notion – job satisfaction versus affect in the workplace, for example – but also the time horizon.[9] We are purposefully specific in the question about the time frame, measuring affect during a week-long period, allowing us to match up with contemporaneous weekly data on productivity. One concern is that respondents may answer solely based on their current mood at the end of the week (or the peak during it), given that recall of emotions can be biased (Thomas and Diener, 1990). While affec-

[9] For example, Finkelstein, Luttmer and Notowidigdo (2013) ask people about their happiness "during the past week", Allcott et al. (2020) variously measure people's happiness "right now", "over the last ten minutes" and "over the past four weeks," and Kahneman et al. (2004) utilize the Day Reconstruction Method to measure people's emotional experience during a particular episode of the previous day.

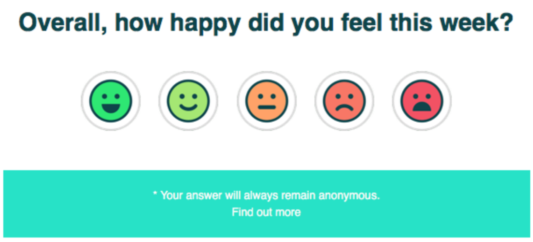
Figure 2: Distribution of Happiness and Sales



(a) Sales (b) Happiness

Note: Panels (a) and (b) show the overall distribution of sales and happiness. Each observation is a worker-week.

Figure 3: Happiness Survey Email



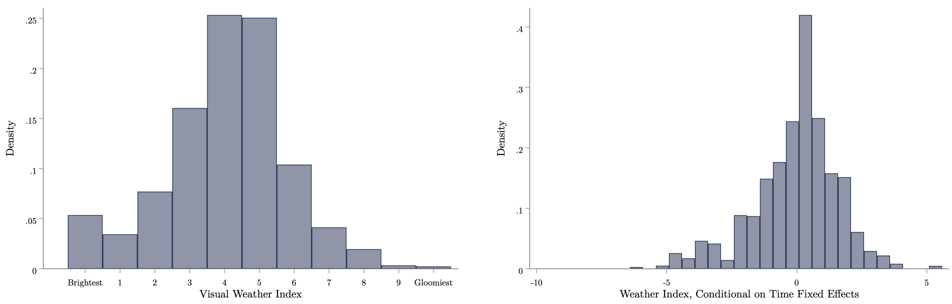
Notes: Screenshot of the mood survey, which was sent weekly over a six month period to all workers. Respondents had to click a face within the email for their response to be registered. See text for more details.

tive recall is not always fully accurate, Kemp, Burt and Furneaux (2008) nevertheless show that subjects do respond to the time prompts in such questions – by surveying daily mood over a week-long period and then subsequently asking about total weekly mood, they are able to show that people are able to provide a good reconstruction of the affective states (more generally, see Krueger and Stone, 2014; Krueger and Schkade, 2008, for further discussion of the validity and reliability of SWB measures topic).[10]

The distribution of responses is shown in panel (b) of Figure 2. Happiness in our setting is low, with the modal answer being the most unhappy. We use our happiness responses both ordinally and cardinally, depending on specification. When using happiness as a continuous measure—as is typically done in the SWB literature—we assign the five categorical happiness states equally-spaced numerical values between 0 and 10 (0, 2.5, 5, 7.5, 10) from least happy happy to most happy in order to be aligned with a number of scales used in similar survey measures (for example, in household panel surveys widely used in the literature). When doing so, the mean response is around 4, with a within-person standard deviation of 2.4 (see Table A1).

[10] In Figure S7 we are able to provide an empirical test of the temporal nature of our happiness measure. Using daily weather data (described in more detail below), we show that the mood measure taken on a Thursday afternoon or Friday is meaningfully related to weather exposure not only on that same day, but a few days before within that same week. But that the relationship strongly declines in the days *after* the mood report – lending intuitive support to the weekly nature of the measure.

Figure 4: Distribution of Visual Weather Index



(a) Raw Data (b) Within-Week Variation

Note: Visual weather index is a weekly measure that counts the total number of daily instances of fog, rain, and snow in the vicinity of each call centre. Panel (a) shows the raw distribution, while panel (b) shows the distribution of residuals from a regression of the index on week fixed effects (bin width=0.5).

Sample construction and characteristics. We aggregate all of the administrative data to the Monday-to-Friday working week. All 1,793 workers were invited to take part in the study and were sent weekly mood surveys. Of these employees, 1,438 (around 80%) participated by answering at least one survey over the subsequent 6 months. Conditional on participating in the study, workers responded to a mean of 10.3 waves ($SD = 7.1$). The weekly response rate of workers who participated was on average around 37%, which rises to 50% if we focus only on workers who work on Thursday or Friday.[11] We drop any participants who responded to only one survey wave, since we rely on within-worker variation over time. This leaves us with a final sample of 1,157 employees. Summary statistics for this final sample are shown in Table A1.

4 Empirical Results

We are interested in whether positive affect has any causal impact on weekly performance at work. We estimate equations of the following general form:

$$S_{ijt} = \beta PA_{ijt} + \gamma X_{ijt} + \nu_i + \tau_t + \varepsilon_{ijt}, \quad (1)$$

where S_{ijt} corresponds to the sales performance of worker i in call center j during week t, and PA_{ijt} is their positive affect during that same period. 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 mood and sales. Finally, we include a vector of controls X_{ijt} for two time-varying labor inputs – the (logged) total number of selling hours and the fraction of time spent at work in the week on mandatory non-productive activities. We adjust the error term to account for two-way clustering on individuals and location-week.[12]

[11]See Appendix A for discussion of non-response and attrition. Reassuringly, neither weekly sales performance nor team average happiness (minus the focal worker) is significantly related to non-response within-individuals over time.

[12]For an empirical example of a similar approach using panel data, see Acemoglu and Pischke (2003), who cluster standard errors at the individual and at the region-time level. A week is defined in our case by the date on which a worker answered our survey and location by the call center.

4.1 Reduced Form Effect of Exposure to Weather on Performance

We begin by following the existing literature on mood-in-the-field effects by estimating the impact of weather on sales. That is, we first assume weather to be a suitable proxy for PA_{ijt} in the above equation. Given that sales is a right-skewed count variable including some zero values, we use a Poisson quasi-maximum likelihood estimator. Using local weather station data that we matched to the locations of the 11 call centers, we coded a Visual Weather Index that corresponds to the total number of daily incidences of fog, rain, and snow during the working week (see Appendix J for more details and Figure 4 for the distribution).[13]

In column (1) of Table 1 Panel A, we show that the Visual Weather Index has a negative, though imprecisely estimated, impact on sales. Going beyond this, we further isolate the mood effect by using variation not only in weather but also in people's visual exposure to it, making use of two additional factors critical to our identification strategy. In our setting, while all workstations are open-plan across buildings, the type of building varies significantly – all the way from warehouse-style workplaces with almost no exposure to outdoor conditions to glass tower buildings or large window buildings with full exposure to the outdoor weather. Coding the proportion of external walls that is covered by glass windows using image processing software, we confirm there is large variation across call centers – with the window share ranging from .03 to .59 (see Table A1).[14] Figure 5 illustrates this variation in architecture with photos taken from the outside and inside of Doncaster and Swansea call centers, two ends of the external window coverage spectrum.

In column (2) of of Panel A in Table 1, we interact the weather index with visual exposure to it. We find that the negative effect on sales is much stronger in situations where weather is *visible*. The effect of window share itself is not estimated since it is captured by the worker fixed effects, but recall that it lies between 0 and 1. The coefficient on the main effect of the weather index thus gives the effect of weather for buildings that have no windows. Here we find no significant impact. The interaction with window share, on the other hand, suggests that the effect is much larger in magnitude (i.e. more negative) for buildings that are fully windowed.

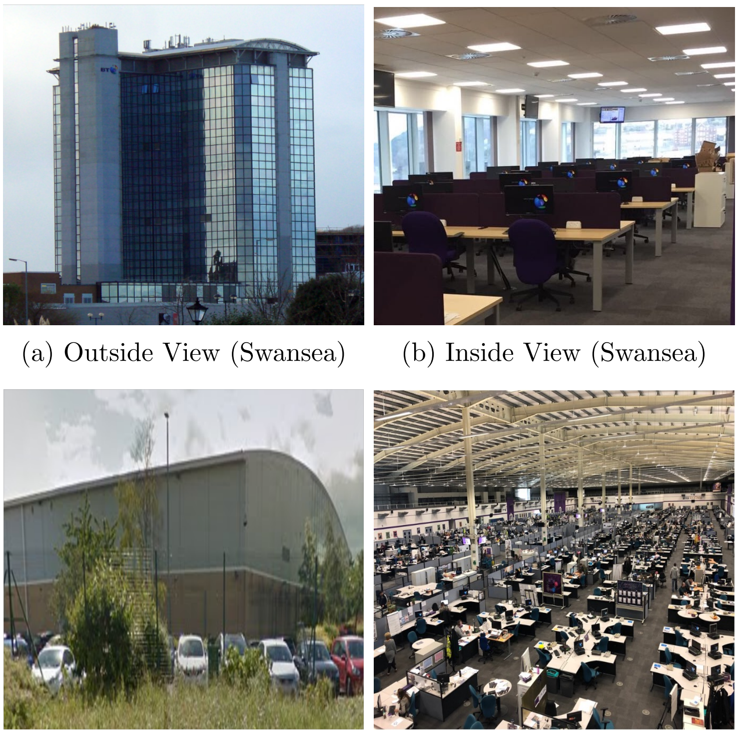
To further illustrate what these results mean in terms of magnitude, we regress our visual weather index within equally-sized groups of call center workers by window shares, namely centers with below median window share (whose average window share is only 9%), versus centers with above median window share (whose average window share reaches 32%). As can be seen in Figure 6a, the point estimate on sales is close to zero in call centers with very few windows but significantly negative in centers with many windows.[15]

[13]This in theory has a range of 0 to 15, where 0 would mean a likely bright day with no rain, snow, or fog on any day of the week at all and 15 would mean that all three happened on every single day of the week. Importantly for our identification, the United Kingdom has sufficiently volatile weather that there is significant variation across call centers, even within a given week (Panel (b) of Figure 4).

[14]For an objective measure of external window-coverage, we first collect all wall photos from each of the call centers in the dataset using Google Street View. For each building, we then code the percentage of wall surface that is covered by glass windows using the *ImageJ* software (see Appendix H for more details). We supplement this data below with a worker-level survey asking about the number of windows and subjective experience of natural light.

[15]Here in the above median share group, every 1-point increase on the weather scale (which is coded such that higher numbers mean gloomier weather and has a standard deviation of 1.36) lowers weekly sales by 1.5%. We also look for potential asymmetries, testing whether sales are more or less sensitive to visual exposure to bright or gloomy weather. Though both effects go in the expected opposite direction, we find no systematic evidence of

Figure 5: Photos of External Walls and Internal Offices



(a) Outside View (Swansea) (b) Inside View (Swansea)

(c) Outside View (Doncaster) (d) Inside View (Doncaster)

4.2 First Stage Impact of Exposure to Weather on Happiness

The finding of a significant interaction in the reduced form-equation between weather and visual exposure to it lends noteworthy credence to the idea that we are picking up a mood effect. However, without observing mood itself, alternative explanations may still be possible. In column (1) of Panel B in Table 1, we show there is, on average, a negative effect of gloominess But column (2) again suggests that this effect is much stronger in buildings with high glass coverage (and non-existent in buildings without windows). Another way to see the contingent relationship of weather on happiness is shown in Figure 6b, which shows that the negative effect of gloomy weather on happiness is concentrated within the buildings with high window coverage. This means at least two things. First, by using exposure to visual weather, we are effectively relying on variation in positive affect rather than any physical effect of weather. Second, we provide evidence that the source of the mood shock is occurring while at work.[16]

Our identification comes here not from either windows or weather, but rather from the interaction between the two. Note that the main effect of windows is not estimated, since it is subsumed into the worker fixed effects. Indeed, we are not here principally interested in the effect of windows, particularly since we do not have random variation in the architecture of these buildings – which may, for example, be correlated with local economic conditions. Nevertheless, we show that the presence of windows is—on average—positively correlated to workers' mood (Figure S1). More important for our identification is the within-worker variation in happiness

an asymmetrical effect in terms of magnitude (Figure S5).

[16]In this case, just like in the case of sales, we find no strong evidence of any asymmetrical effect on mood between the effect of bright and gloomy weather (Figure S5).

Table 1: Impact of Exposure to Visual Weather on Sales and Happiness

	(1)	(2)	(3)
Panel A: Reduced Form (DV: Sales)			
Visual Weather Index	-0.0059	0.0069	
	(0.0036)	(0.0074)	
Weather × Proportion Windows		-0.0666**	
		(0.0333)	
Visual Exposure to Weather (SDs)			-0.0258**
			(0.0104)
Observations	12,282	12,282	12,282
Log Likelihood	-41,657	-41,643	-41,646
Panel B: First Stage (DV: Happiness)			
Visual Weather Index	-0.0617***	0.0139	
	(0.0207)	(0.0377)	
Weather × Proportion Windows		-0.3572***	
		(0.1325)	
Visual Exposure to Weather (SDs)			-0.1950***
			(0.0442)
Observations	12,282	12,282	12,282
R^2	0.534	0.534	0.534
Panel C: CF Poisson-IV (DV: Sales)			
Happiness			0.1331**
			(0.0532)
Observations	12,282	12,282	12,282

Notes: Panel A: Poisson-FE models reported with weekly sales as dependent variable. Panel B: OLS-FE models reported with weekly happiness as dependent variable. Panel C: two-stage control function Poisson-IV model. Standard errors in parentheses, adjusted for two-way clustering on individuals and location-week in Panels A and B and calculated using the delta method in Panel C (see text for more details and alternative SEs using a bootstrap). All models include individual and week fixed effects, work schedule controls, and indicator variables for day of week of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

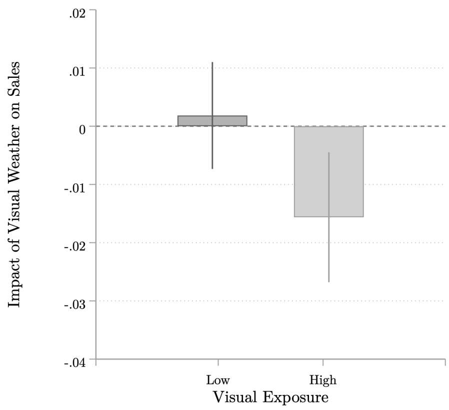
across the buildings. Consistent with the intuition behind our empirical strategy, the mood of workers (and their performance) in more heavily glass buildings is more variable than those who work in call centres that block visual access to outside weather (see Figure S1).[17]

4.3 Putting it Together: IV Estimates

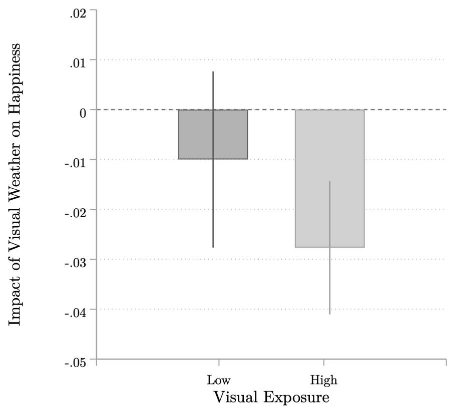
Control Function Poisson-IV Estimates. Column (3) of Table 1 shows the elements of our main IV approach. For our main IV specifications, we prefer to use one single instrument – since there are well-known issues with strength when using multiple instruments. We will thus rely on the interaction term by itself in those models (such that this measure can be thought of as a window-weighted or window-adjusted weather index). To aid interpretation when using it by itself, we z-score it – such that, in Panel A, the coefficient suggests that a one standard deviation change in visual exposure to weather leads to a 2.5% shift in sales. To be valid as an instrument, it must (at least) have a sufficiently strong impact on happiness in the first stage.

[17] The variance in sales across call centers is also generally higher in locations with windows (see Figure S2). The overall relationship between windows and sales is negative, though less systematic than the positive relationship between windows and happiness (see Figure S2).

Figure 6: Impact of Weather on Sales and Happiness by Visual Exposure



(a) Sales (Poisson-FE)



(b) Happiness (OLS-FE)

Note: Coefficients are reported together with 95% confidence intervals. Panel (a): Poisson-FE models reported with weekly sales as dependent variable. Panel (b): OLS-FE models reported with weekly happiness as dependent variable. Each dot corresponds to a separate regression for low visual exposure (below median window share) vs. high visual exposure (above median window share). Average window share within each equally-sized group of workers is 9% and 32%, respectively. Full set of fixed effects and controls are included, as in main specification.

In the third column of Panel B, we find that our preferred instrument significantly depresses mood (with an F-statistic of 19.44). A one standard deviation increase in visual exposure to weather leads to around a 0.2 point decrease in happiness on the 0-10 scale.

The IV estimate logically arises from the ratio between the reduced form and first stage estimates presented in column (3) of Table 1. Following Lin and Wooldridge (2019), in Panel C we provide an initial IV estimate using a control function approach. Here we estimate the first stage via OLS and then add the first-stage residuals as a control in the second-stage Poisson regression.[18] A one point increase in happiness (on a 0-10 scale)—which amounts to around a 42% of a standard deviation increase in within-worker happiness—leads to a .13 log point increase in weekly sales. We estimate an average marginal effect of 3.36 additional sales ($SE = 1.34$; 95% CIs: $0.73, 5.99$). In line with there being a strong downward bias in the non-instrumented model (see Appendix B for further discussion and evidence on the strength and direction of this bias), the non-IV estimate is smaller than the quasi-experimental one. Indeed, in the equivalent Poisson-FE regression of sales on happiness, with the same set of fixed effects and controls, the coefficient on happiness is 0.0141 ($SE = 0.0014; p < .001$). We return in more detail to this point in our extended discussion of magnitudes below.

2SLS Estimates. While the Poisson approach is attractive with a count outcome variable like sales, there are nevertheless a number of benefits to using a linear two-stage least squares (2SLS) estimator, which is more tractable and whose properties are well-understood in the literature.

[18]We use the delta method to calculate the standard errors in this case, as suggested by Lin and Wooldridge (2019). As an alternative, we can also bootstrap the whole two-stage procedure using a panel bootstrap (i.e. that re-samples individuals). When doing so, we find a smaller standard error ($SE = .0483$). Estimating 2SLS regressions, as below, we are able to more straightforwardly allow for two-way clustering on individuals and location-week.

Table 2: 2SLS Estimates of the Effect of Happiness on Sales Performance

	Weekly Data										Daily Data	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Happiness	0.142***	0.157***	0.193**	0.142***	0.122***	0.134***	0.124***	0.106***	0.167***	0.122***	0.176**	0.168*
	(0.047)	(0.053)	(0.087)	(0.047)	(0.039)	(0.049)	(0.047)	(0.039)	(0.054)	(0.042)	(0.075)	(0.099)
Observations	12,282	12,282	12,282	12,282	12,282	12,282	12,282	12,282	12,282	12,282	43,883	11,545
Individual + Time FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Work Schedule	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls:												
Temperature	No	Yes	No	No	No	No	No	No	No	No	No	No
Visual Weather Index	No	No	Yes	No	No	No	No	No	No	No	No	No
Detailed Schedule	No	No	No	Yes	No	No	No	No	No	No	No	No
England × Week FEs	No	No	No	No	Yes	No	No	No	No	No	No	No
Alternative IVs:												
Subjective Windows	—	—	—	—	—	Yes	—	—	—	—	—	—
Subjective Light	—	—	—	—	—	—	Yes	—	—	—	—	—
Weather^2	—	—	—	—	—	—	—	Yes	—	—	—	—
asinh(Weather)	—	—	—	—	—	—	—	—	Yes	—	—	—
Dis-Aggregated Weather	—	—	—	—	—	—	—	—	—	Yes	—	—
Daily Sample:												
Response Day Only	—	—	—	—	—	—	—	—	—	—	No	Yes

Notes: 2SLS models estimated using asinh(sales-per-hour) as the dependent variable. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include individual and time fixed effects, work schedule controls, and indicator variables for day of week of response to survey. Columns (1) to (10) use weekly data on sales and exposure to weekly weather. Columns (11)-(12) use daily data on sales and exposure to weather. Time fixed effects are study week in columns (1) to (10) and calendar date in columns (11) and (12). Column (11) assumes that responses to the happiness question applies equally to each of the weekdays. Column (12) only uses observations from the day of response to the happiness survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Given that sales is right-skewed and contains zeros, we use as our main outcome in this case the inverse-hyperbolic sine (asinh) of sales-per-hour – a transformation that allows us to estimate equations that are roughly log-linear, while preserving the zeros and general interpretability of our estimates. We thus estimate a 2SLS equation of the following form:

$$asinh(S_{ijt}) = \beta PA_{ijt} + \gamma X_{ijt} + \nu_i + \tau_t + \varepsilon_{ijt}. \tag{2}$$

This equation includes the same set of fixed effects and time-varying controls as above and adjusts the error term for two-way clustering on individual and location-week. Here we use a direct survey measure of positive affect rather than a proxy, but in order to deal with endogeneity, PA_{ijt} is instrumented for using Z_{ijt}, where:

$$Z_{jt} = Weather_{jt} \times PropWindows_j. \tag{3}$$

The first-stage equation is the same as above in column (2) of Panel B in Table 1, where we showed that visual exposure to weather has a strong effect on worker happiness. Using this variation, we show in column (1) of Table 2 a 2SLS estimate of 0.14 (95% CIs: $.05, .23$). This is similar to the Poisson-IV estimate shown above using a control function approach.

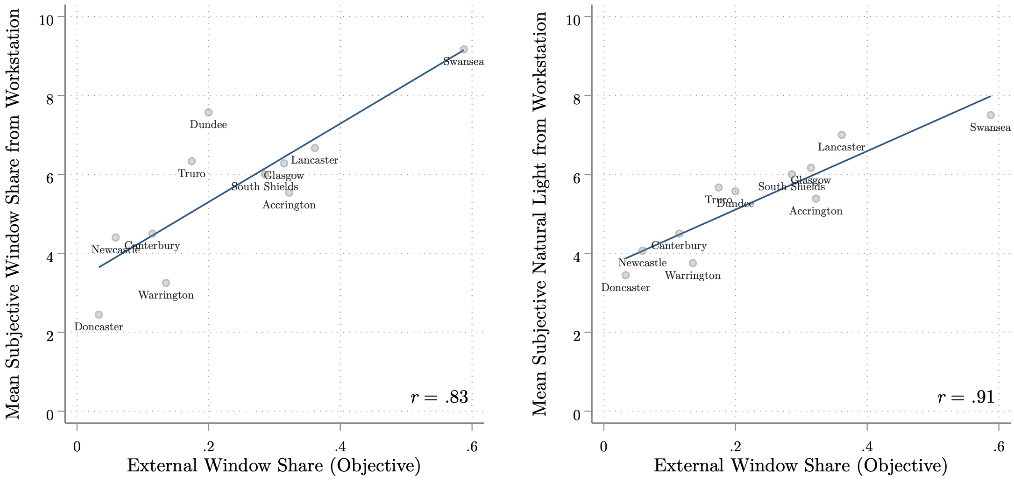
4.4 Robustness

Window-Based Exposure. Given that our identification relies on workers' differential window-induced exposure to weather, one concern with the approach of relying on objectively-measured external window coverage is that this may not translate to a worker's visual exposure to outdoor conditions. For example, even if a building is fully externally glass, some workers inside may nevertheless have work-spaces with no windows. In our context this is unlikely to be a problem, given the internal layout is similarly open-plan across buildings. This means the inside is always one large space in which any given worker will have roughly equal visual access, on average, to the walls – be they glass or solid. Moreover, as a fully open space, natural light (if it is let in through windows) will spread through the space.

We combine our objective data on window coverage with two additional sources of data: semi-structured interviews with managers and a short supplementary survey of workers (see Appendix G for more details on this additional data collection). The survey enabled us to confirm, from the the perspective of workers themselves, what we (a) were told by management as well as (b) observed on site visits: that despite significant exterior variation, all of the buildings are open-plan (see Figure 5 for examples of this in both high- and low-window buildings). We find from the survey that 95.6% (N=318) report working in an open office space as opposed to a closed office space.

Moreover, the supplemental worker survey allowed us to ask workers directly about their subjective experience of windows and light. Here we first ask workers to imagine sitting at their typical workstation, and then ask "*Do you see few or many windows?*" – with a 0 to 10 answer scale where 0 is labelled as "*no windows at all*" and 10 "*fully glass building.*" As a further test to this, we also ask how much natural light they have access to from their workstation, on a scale where 0 is "*no natural light at all*" and 10 is "*like sitting outside.*" As can be seen in Figure 7, we

Figure 7: Subjective and Objective Measures of Visual Exposure to External Conditions



Notes: Objective window share is calculated from photos and using image processing software. Call-center-averages displayed for responses to worker-level survey questions, each on a 0 to 10 scales, with the following wordings. Subjective windows: "Imagine sitting at your typical workstation. Do you see few or many windows? Use the slider below. Imagine a 10 being a completely glass office and 0 being a room with no windows at all." Subjective light: "While at work, how much natural light do you have access to from your workstation? Use the slider below. Imagine a 10 being like seating outside while working and 0 being an office with no access to natural light at all."

find that workers' subjective impression of windows and natural light is very highly correlated with our objective measure of external window coverage ($r = 0.83$ and $r = 0.91$, respectively). Using these subjective measures instead of the objective measure in order to construct our main instrument, we also show that this does not affect our main 2SLS estimates (see columns (6)-(7) of Table 2).

Product Demand Effects. A further concern in relation to the exclusion restriction is that weather may have an effect on customer demand (either directly or indirectly by affecting customers' mood). This would be a direct threat to the exclusion restriction – namely, that visual exposure to weather will only be a valid IV in this 2SLS equation under the condition that weather has no direct impact on productivity other than through its effect on mood. However, in this two-sided market, we have call centre workers in fixed locations, and customers calling in from all over the country. As we noted in the Introduction, one of the four key factors in our identification strategy is that our call centers are sufficiently well dispersed across a country that has variable weather across space, even within a given week. For a map of the spatial distribution of these call centers, see Figure 1, which shows this dispersion across three nations of the United Kingdom. We are also able to show in panel (b) of Figure 4 that, conditional on the time fixed effects, there is variation in the visual nature of the weather across the locales.[19]

The inclusion of time fixed effects ensures that our key piece of identifying variation is exposure to weather shocks across call centers but within week. 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. Given that call centers do not field calls based on originating location,[20] *local* weather in the vicinity of the focal call center should be independent, on average, of customer demand.[21]

Although we cannot observe the origin of calls directly—our data is at the worker-day level rather than the call level—we provide direct evidence that local demand pressure at a particular call center measured by the average daily number of incoming calls per worker and the average duration of a call is not affected by daily weather that is local to that particular call center (Table S5).[22] Moreover, if it were the case that calls were only routed within each of the four nations of the UK, such that all calls coming from English customers were only routed towards workstations located within England, then the addition of England-week fixed effects should significantly affect our point estimate. We find in column (5) of Table 2, however, that this is not the case.

Additional Controls. Adverse weather could physically be affecting sales performance through changes in temperature or pollution. Temperature has been shown, for example, to affect student learning and academic test scores (e.g. Park et al., 2020) as well as investment decisions (Huang, Xu and Yu, 2020).[23] We show in column (2) of Table 2 that, when controlling for local temperatures, all of our main findings are robust. The temperature variable itself does not enter into the equation in a statistically significant way. Relatedly, air pollution can have a direct impact upon worker productivity (e.g. Chang et al., 2016, 2019; Graff Zivin and Neidell, 2012). Although we are not able to measure air pollution directly, pollution correlates with temperature, and there is little reason to suspect that glass-clad buildings would be more susceptible to air pollution effects than warehouse-style ones.

In addition to controlling for temperature, we are able in column (3) of Table 2 control for the main effect of the visual weather index, while using the interaction of weather and windows

[19]Although weather also varies within a day across locations, the (normalized) distribution of the weather index within a day is narrower than within a week (Figure S8). Measuring employee mood on a weekly basis is hence also consistent with the spatial variability of weather being higher within weeks than within days.

[20]The company have since our study introduced a new system (in Spring 2019) that does allow for geographical allocation; however, no such technological capability was in place at the time of our study (July 2017 - January 2018).

[21]Using data on customer satisfaction aggregated at the location-week level, we confirm that a small (negative) residual correlation between our weather index and customer satisfaction across workstations nationally (r=-0.105) is entirely nullified after adding week fixed effect (r=0.001).

[22]Local supply shocks could also affect productivity if, for instance, gloomy weather makes it harder for workers to answer calls. However, this should affect all call centers regardless of the window coverage of the building. These effects are more likely to be driven by snowfalls than by fog or rain, but we find the effects are not driven solely by snowfall (Table S6).

[23]In an empirical strategy somewhat analogous to ours, Park et al. (2020) show null effects in schools that have air conditioning (i.e. instead of turning visual weather on and off with windows, they turn heat on and off with AC). Nevertheless, Heyes and Saberian (2019) present evidence to suggest a mood effect using high temperature and judges' decisions (though without measuring mood), who are inside in climate controlled buildings – arguing that they may 'import' the effect of outdoor temperature when they move indoors. In terms of studying mood effects, we prefer to focus clearly on visual aspects of weather coupled with observable variation in exposure to be more sure of the mood mechanism taking place.

as our instrument. This analysis ensures our identification comes solely from the interaction, and when doing so we find consistent results though a slightly higher point estimate. As was suggested by the reduced form and first stage evidence shown above in column (2) of Table 1, the main effect of weather is not statistically different from zero in this case.

While we favor a parsimonious specification, our main 2SLS estimate also remains largely unchanged if we include a much more exhaustive set of controls capturing the detailed daily work schedule of workers and additional labor supply controls (see column (4) of Table 2).[24]

Sickness. Adverse weather conditions may cause sickness among workers, and impair their ability to work effectively if they attend work while ill. Looking directly in the data, the local weather turns out to be unrelated to the local share of workers under sick leave in any given day or week (Table S7). We also consider the possibility that such a relationship may occur with some lag. Weather conditions a day (or a week) before remains unrelated to the frequency of sick leaves a day (or week) after. Most importantly, however, any sickness argument would apply whether or not the call center had many windows. Given the relationships shown in Figure 6, it seems clear that the variation our instrument picks up on is related to psychological rather than physical health.

Sorting Effects. It is assumed throughout the paper that no other factors correlated with the share of windows but unrelated to visual exposure itself may explain heterogeneous sensitivity to weather. The issue could arise if certain types of workers happen to be more negatively affected than others by adverse weather conditions (e.g. older or sicker people), and if those same workers tend to be systematically working in call centers with more windows. Of course, one should only be worried about such sorting effects if there exists important sources of heterogeneity in worker sensitivity to weather in the first place, other than through visual salience. We investigate this possibility directly, looking at whether weather (non-adjusted for exposure) affects workers' mood differently across a number of important measurable characteristics. We look at basic demographics (gender, age and workers' tenure), the total number of weekly sales, and how frequently the worker takes sick leave. We find no evidence of heterogeneity across any of these dimensions (Table S8).[25]

Functional Form of Happiness. One concern that we noted above in relation to our happiness survey is that answers are given on an ordinal scale. We make the assumption that this can be cardinalized into a continuous measure of happiness. We provide a test for the reasonableness of this assumption by replacing continuous happiness with indicator variables for different

[24]This includes a full set of 35 indicator variables capturing the detailed daily work schedule of workers. For each day of the week, we control for whether workers started (ended) their shift in the morning (7:00-12:00), afternoon (12:00-17:00), or evening (17:00-24:00), or whether they did not work at all that day. We also add to this list more detailed working supply controls, namely i) the total amount of hours spent on breaks, ii) the total amount of hours spent working overtime, and iii) whether the worker was reported sick during the week.

[25]It may still be that workers in call-centers with more windows would have a tendency to report a better (or worse) mood when they can visualize good (or bad) weather. This effect would be consistent with the effect of visual weather on mood being stronger in call centers with more windows. However, it is not a threat to our identification strategy as long as (i) this effect is not biased towards reporting only good or bad mood and (ii) reporting behavior is correlated to an actual shift in mood, which is further confirmed by the fact visual exposure to weather does have clear behavioral effects on performance.

levels of happiness in a Poisson regression of sales on happiness (including the full set of fixed effects and controls, as in our main IV specification). Figure S3 reports the coefficients from this exercise. 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.

First Stage Functional Form. In Figure S4, we show a graphical representation of our first stage regression of happiness on visual exposure to weather. The relationship looks roughly linear. But in order to explore and account for possible non-linearities more fully, we test alternative functional forms for our instrumental variable. In Table 2 we use the squared value of the exposure to poor weather index as well as the inverse hyperbolic since of weather (roughly equivalent to the log, but allowing for zeros). We find in columns (8) and (9) of Table 2 that the resultant second-stage coefficients are consistent with our main results.[26] Related to this, concerns related to weak instruments should be mitigated by the reduced form evidence provided in this paper. Indeed, reduced form estimates remains unbiased estimates, even if the instruments are weak (Angrist and Krueger, 2001).

Heterogeneous Responses. Instead of identifying an average treatment effect (ATE), a valid instrument identifies a local average treatment effect (LATE) in the second stage (Angrist, Imbens and Rubin, 1996) – that is, the effect driven by those whose mood can be most easily manipulated – or in our case, sensitive to visual exposure to weather.[27] Assuming away the possibility of heterogeneous treatment effects can be problematic when the causal effect of the endogenous variable is directly related to the individual's own choice (Angrist and Imbens, 1995). While this issue has been widely discussed, for instance when estimating the returns to education (Angrist and Krueger, 1991), in our case, mood movements are largely "external" to an individual: one does not have a direct control over them. Heterogeneous treatment effects arising from employees' selection on the productivity gains of good mood are unlikely to occur in our context. Table S9 shows the first stage of our IV strategy this time interacting *visual* exposure to weather with each of the six main characteristics described earlier. We find no evidence of heterogeneity across any of these dimensions.[28] We also conduct sub-group analysis splitting between generally unhappy and happy workers as well as generally low and high productivity workers (below/above median in each case). We find no statistical difference

[26]We also show a 2SLS regression in column (10) of Table 2 where we include three separate instruments, one each for the incidence of the three weather phenomena that make up the index. Here we run into potential problems of weak instruments, since the combined F-statistic is much lower in this case – as is typically the case when using multiple instruments (see Table S6 for the first stage). However, results remain consistent overall.

[27]As pointed by Angrist, Imbens and Rubin (1996), the LATE (or causal effect of weather-induced happiness on sales) corresponds to the ratio of the impact of visual exposure to weather on sales (the reduced-form, which may be referred to as an intent-to-treat effect) and the impact of visual exposure to weather on happiness (i.e. the fraction of workers whose mood is sensitive to weather) – or alternatively it can be calculated using an IV estimator. The LATE is of specific policy-relevance if the goal is to target those individuals whose affective state is most likely to be affected by any policy change.

[28]One reason for the higher salience of weather in buildings with more windows may result from workers paying more attention to weather over time in those locations as they can more easily blame the bad weather for their low productivity. Such learning effects should build up over time, so that workers with longer tenure should be more sensitive to visual exposure to weather, which is not what we find (Table S9). We also found no evidence for seasonality in the effect of visual exposure to weather for both happiness and sales (see Table S10).

in the sensitivity of productivity to weather-induced mood effects across those groups (Table S11).

Timing of Happiness Response. Respondents are explicitly asked to report their average mood state over the course of the week. To test whether the response to the happiness question indeed captures respondents' mood during the week in which the response is reported (and not the week after or before), we first check whether weekly reported happiness significantly relates to visual exposure to weather a week before or a week after the happiness response week. None of those significantly relate to reported happiness, and the coefficient estimated on visual exposure to weather during the week in which the response is reported remains unchanged (Table S12). We find the same result looking at reported happiness and weekly sales. Moreover, looking at those relationships within the same week (with reported happiness applied equally to each day up to 3-4 days before vs. after the response day) using daily weather and sales data around the response day, reported happiness about "this week" is significantly and strongly related to visual exposure to weather and sales on the response day and a few days before, but much less so (or not significantly) a few days after (Figure S7).

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 table S13 we regress daily sales on daily weather exposure at work, 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 daily visual exposure to weather has a negative and significant effect on daily sales performance.

We provide two further analyses, this time using the weekly happiness data combined with the daily sales and performance data. First, 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 (Figure S7 demonstrates this assumption is supported by the data itself). We then estimate the 2SLS regression using a daily dataset, and include date (instead of week) fixed effects. When doing so, in column (11) of Table 2 we find that our results are unchanged. In column (12) we make the more restrictive assumption that the weekly happiness question applies only to the day of response. Restricting only to the day of response, we again find similar results. The effects are less precisely estimated, consistent with the lower variance in weather within a given day across locations, relative to within a given week, as documented in Figure S8.

5 Evidence on Channels

Section 2.3 discussed three broad ways in which happiness may affect work performance: the impacts of positive affect on cognitive processing (in particular being more creative, efficient or integrative in thought), work motivation, and social or emotional skills. We provide suggestive evidence on the relative importance of those various mechanisms in our context.

5.1 Cognitive Mechanisms

Though we cannot measure cognition directly, an impact of positive affect on the ways in which workers think could be reflected in three major labor productivity measures: adherence to daily workflow schedule, speed, and call-to-sale conversion. First, adherence to workflow – where, for example, positive affect may lead to greater flexibility in thought and a better ability to multi-task as well as effectively plan and switch between tasks. In our setting, workers attend and have their day's workflow scheduled for them and displayed on their terminal screen (for example, they may 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. We code our adherence outcome variable here as 1 if the firm's target is met, zero otherwise.[29] How does workers' mood affect this outcome? In Table 3 Panel A, we present reduced form evidence using our measure of visual exposure to weather. In Panel B, we show the results from the second stage of a 2SLS regression, in which happiness is instrumented for using visual exposure to weather. We find that happier workers adhere more closely to the workflow that has been set out for them (see column (1) in Table 3). In our setting, however, conditional on the total number of hours spent at work (selling or doing other internally scheduled non-productive activities), adherence happens to have little influence on sales (see Table S14). In other settings with different types of work, however, where adherence may be more critical to performance, this mechanism could well be more consequential.

Second, improved efficiency in information processing, brought about by higher positive affect, should translate into an ability to work faster. We observe, on a daily basis, the total number of minutes spent on incoming calls as well as the number of calls taken, and show in column (2) of Table 3 that in happier weeks workers work faster. This "speed" measure is what would typically be used as a labor productivity metric in the manufacturing industry. However, in our setting, and in the service industry more generally, it is not clear that taking more, shorter calls will be beneficial when the goal is selling.[30] Table S14 confirms the total number of calls per hour is not a good predictor of productivity – and, if anything, it is associated with a reduction in the number of weekly sales per worker. Though, again, in settings where speed may be more crucial, this channel could have more of an effect.

Relative to adherence or speed, a third productivity metric is more clearly linked to problem-solving and creativity: call-to-sale conversion. Column (3) shows that in happier weeks, workers convert more of their calls to sales. Out of the three cognitive mechanisms, workers' improved ability to solve customers' problems seems most likely to explain our main productivity effect. Indeed, the point estimate for adherence and calls per hour are relatively small: Though a 1-point increase in happiness leads to a rise in the average number of calls per hour from 5 to 5.3 calls, speed is unable to explain much of an increase in sales. The dominance of the

[29]The measure is continuous, out of 100. However, occasional deviance from this workflow may be beneficial if the worker has to stay on a call to complete a sale, for example. As such, the firm sets a loose target of 91% adherence each week, which is the cut-off we use.

[30]This speed-quality trade-off is particularly salient in call center settings (Singh, 2000). Indeed, faster calls may displease customers and make them less likely to buy if the operator is too blunt or quick with them. Furthermore, sales calls are likely to be mechanically longer, due to the time it takes to complete an order, take payment details, and so on.

Table 3: Happiness and Labor Productivity

	Adherence (Met Target=1)	Calls Per Hour (Log)	Conversion Rate (Log)	
	(1)	(2)	(3)	(4)
Panel A: Reduced Form				
Visual Exposure to Weather (SDs)	-0.0155**	-0.0119***	-0.0564***	-0.0650***
	(0.0068)	(0.0033)	(0.0143)	(0.0151)
Adherence (Met Target=1)				0.0084
				(0.0105)
Calls per hour (ln)				-0.8782***
				(0.0703)
Observations	12,169	12,100	11,720	11,672
Panel B: 2SLS Model				
Happiness (0-10)	0.0822**	0.0632***	0.2832***	0.3254***
	(0.0382)	(0.0227)	(0.0965)	(0.1047)
Adherence (Met Target=1)				-0.0628*
				(0.0347)
Calls per hour (ln)				-0.5623***
				(0.1414)
Observations	12,169	12,100	11,720	11,672

Notes: 2SLS models reported, using visual exposure to weather as an IV for happiness. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include individual and week fixed effects, and controls working hours, internal shrinkage, and day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

conversion channel is more apparent in Column (4), where we control for adherence and the number of calls per hour. It confirms that the average number of calls per hour is negatively, not positively, correlated with the conversion rate.

We saw that an important higher-order mechanism explaining how positive affect may lead to higher performance is through its impact on workers' ability to find creative or integrative solutions to problems. In routine jobs or tasks where there is little space for creativity, this channel may be much smaller than in solving more complex tasks (or even go in the other direction). Although we do not observe the amount of time spent (or number of calls) selling different products, we are able to examine the effect of happiness on different *types* of sales. When doing so, we find that, although all of our estimates are less precise, the magnitude is close to zero for regular order-taking (see Figure 8). This is consistent with the main mechanism being call-to-sales conversion rather than working faster (or more efficiently), since line sales are largely mechanical order-taking. More strongly positive effects are found for TV and cell-phone contracts, which are also more technical and involve selling bundles with multiple different options, as well as for re-contracting sales.

5.2 Motivational Mechanisms

Besides purely cognitive mechanisms, positive affect may also increase performance through higher work motivation – in particular, as noted in Section 2.3, by making work more enjoyable, happiness may incite workers to put more effort on the least enjoyable tasks and spend more time at work generally (or take fewer breaks). We have two reasons to believe higher motivation

Table 4: Happiness and High-Frequency Labor Supply

	Sell Time (ln)	Attendance (100% =1)	Overtime (Any = 1)	Paid Vacation (Any = 1)	Break Time (ln)
	(1)	(2)	(3)	(4)	(5)
Panel A: Reduced Form					
Visual Exposure to Weather (SDs)	0.0185	0.0052	0.0066	0.0011	0.0104
	(0.0127)	(0.0103)	(0.0041)	(0.0083)	(0.0067)
Observations	12,282	12,279	12,282	12,282	12,282
Panel B: 2SLS Model					
Happiness (0-10)	-0.0958	-0.0270	-0.0342	-0.0055	-0.0539
	(0.0663)	(0.0542)	(0.0216)	(0.0432)	(0.0378)
Observations	12,282	12,279	12,282	12,282	12,282

Notes: 2SLS models reported, using visual exposure to weather as an IV for happiness. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include include individual and week fixed effects and indicator variables for day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

is unlikely to be the main driver of the happiness-productivity effect in our setting, however. First, we saw that happiness had no impact on simple order-taking, which is the most routine and likely least fulfilling selling task. Since baseline motivation on those tasks is low, if higher motivation were a major mechanism then those particular sales should be most reactive to happiness, which is not what we find. Second, in addition to labor productivity, we can also investigate the impact of positive mood on short-run labor supply decisions. If higher motivation is what drives our main effect on sales, we may also expect to find that workers would spend more time working.

We first look at total number of selling hours, which we used as our main control of labor inputs throughout the paper. In both the reduced form and the second-stage of a 2SLS equation, shown in column (1) of Table 4, we find no robust evidence of any happiness effects on the amount of time spent selling. Given workers' overall labor supply, we also observe additional high-frequency data on the allocation of time between work and non-work (or leisure). First, 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. Here too, we find no robust evidence of any significant mood effects on attendance. This rules out the possibility that happier workers would find work more enjoyable, hence be motivated to attend work more often. We also find negative (but non-significant) coefficients for over-time working. Finally, we also observe whether workers took any paid vacation during the week, and the number and length of breaks taken by workers. Both coefficients are negative but non-significant. Taken 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 want to over-interpret this evidence as strongly suggesting that happiness does not affect labor supply decisions (and the motivation to work) 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 effectively provides us with an ideal setting in which to test for the pure productivity effects.

5.3 Social and Emotional Mechanisms

Finally, positive mood could also augment social and emotional skills, especially in a real-life workplace where workers interact with customers and co-workers. Figure 8 shows that the strongest positive effect of happiness is on re-contracting sales. In these situations, workers are negotiating and, consistent with the experimental literature discussed in Section 2.3, seem to be finding more integrative solutions in happier weeks or being better able to persuade – likely because of relying less on contentious tactics.

The effect of positive affect on re-contracting likely involves both a cognitive (better problem-solving and creative thought) and a social (friendliness and negotiation) channel. But which dominates in our setting? We find evidence more in line with the latter. In a customer-facing setting, this is expected – indeed, 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.[31] Consistent with those hypotheses, we find that our effect is stronger during weeks where customers are, on average, most unsatisfied. Table S16 replicates our main analysis by terciles of weekly customer satisfaction.[32] In weeks where national customer satisfaction is low (bottom tercile), being in a good mood has a much stronger positive effect on sales than during weeks where customer satisfaction is high (middle and top terciles). Going beyond this, we also find that this effect only manifests in the most complex task of upgrade a re-contracting sales: in particular, in weeks where the average customer is satisfied, the effect of happiness on those sales is null, which suggests the type of skills driving the effects of positive affect in our setting are not simply cognitive but also social (Figure 8). We cannot fully adjudicate in this paper as to whether this positive social effect is due to being nicer, better at negotiation (e.g. via persuasion), or a mix of both. The dominance of persuasion in negotiation skills would suggest an overall negative effect on customer satisfaction as customers would accept solutions that may not be in their best interest. Using the weekly average customer satisfaction response for each worker, we show in Table S15 that, within-workers over time, happiness has a negative (although non-significant effect) on customer satisfaction. We are hesitant to over-interpret this result, however, given that the customer satisfaction data is very noisy when used at the worker level, since each employee receives very few feedback ratings, on average, per week.[33]

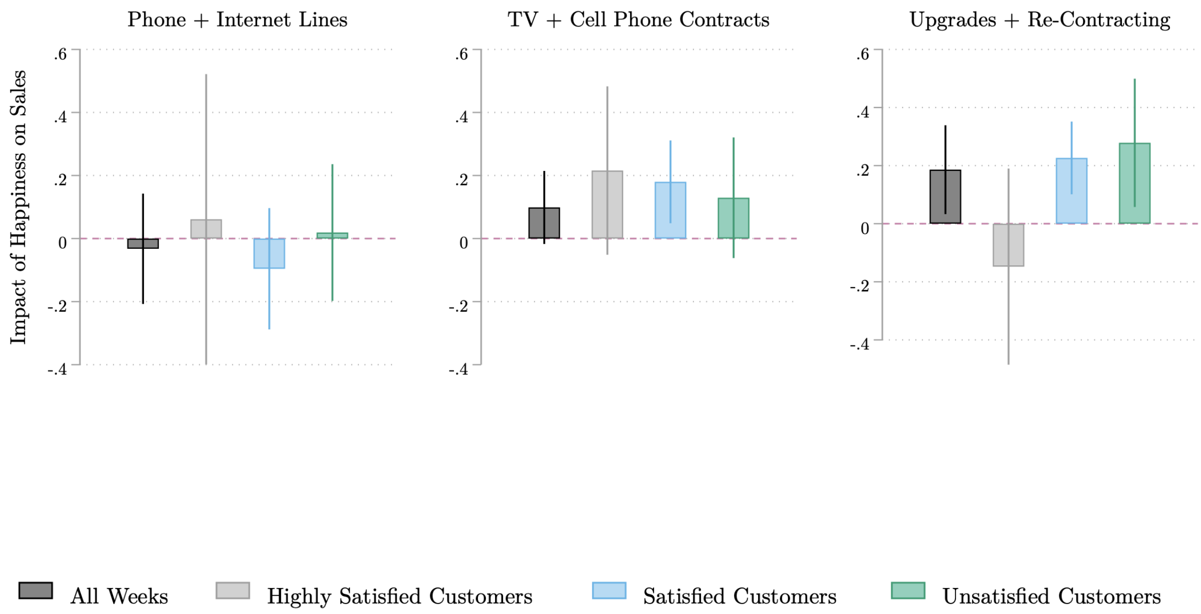
Finally, as we noted in Section 2.3, happiness may here have an ambiguous theoretical relationship with performance given that the more social mechanisms may in some contexts lead to distraction or loafing (cf. Coviello et al., 2020). Indeed, higher sociability may incite workers to spend less time working and more time socializing with their co-workers while at work, which would predict negative happiness effects on productivity. In our particular setting, however, where team work is absent and workers are monitored, we find no evidence suggesting

[31] Related to this work, the psychological broaden-and-build framework similarly suggests that positive affect not only broadens people's thought-action repertoires but also enables them to build resources, meaning that happier people over time amass a greater stock of resources to deal with shocks (Fredrickson, 2001).

[32] 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, which we aggregate here to the national level.

[33] The mean number of satisfaction responses a worker receives per week in our sample is 2.66. 18% of worker-week observations have 0 customer responses and the modal number of responses is 1.

Figure 8: Effects by Average Weekly Customer Satisfaction and Type of Sales



Note: Coefficients and 95% confidence intervals shown from Poisson-IV models, in which happiness is instrumented using visual exposure to weather. The outcome in all cases is the number of sales, but separate regressions are estimated according to the type of sale. Customer satisfaction is aggregated by week to the national level. The sample is then split by terciles of this national customer satisfaction measure.

sociability may be detrimental to productivity. On the contrary, happier workers take calls slightly faster and the effect of happiness on break taking is negative, effectively ruling out any large-scale effect whereby happier workers take longer breaks to chat with their co-workers.

6 Discussion

6.1 Managerial Implications

We employ variation in visual exposure to weather level as an econometric device in order to isolate the causal effect of happiness on productivity – and to explore potential mechanisms. Our study is thus best thought of as a form of *basic research*, albeit in an applied setting. As such, we are not able to say definitely whether windows, in and of themselves, are good or bad for firms. We show that workers are, on average, happier in buildings with more windows; however, without plausibly exogenous variation in window coverage or a setting where workers routinely switch between buildings, it is difficult to make any strong claims. Nevertheless, our findings do demonstrate the importance of an often-overlooked aspect of work, namely the physical environment – and the ways that it is an important factor in mediating the effects of environmental factors on workers.

A natural question is whether or not the effect of weather-induced happiness on productivity is a useful (i.e. policy- or managerially-relevant) parameter to estimate. One initial thing to note here is that although there has been over a century of empirical work on the issue of employee well-being and performance, there remains little causal evidence in the field. Thus our confirmation of earlier laboratory findings (e.g. Erez and Isen, 2002; Oswald, Proto and Sgroi, 2015) is an important step forwards in the literature. In particular, managers may worry that findings inside the lab may not replicate or scale easily in real-life settings, and may not be able to relate the experimental evidence to the types of tasks, skills and workplace environments that their workers face. Our paper not only provides confirmation that the type of experimental evidence generated in the lab replicates in the field, but shows those effects may be even more relevant when looking at more complex tasks involving real social interactions with customers. Conversely, our results should lower managerial concerns that the benefits of happiness on productivity may necessarily be smaller in real-life settings as more sociable workers would simply tend to work less.

Although the root of an emotion is typically easily traceable for people, the source of moods is not. There is little reason to expect the effects of a weather-induced mood to be any different from the effects of a mood state induced by other factors, many of which are managerially relevant. Indeed, a great deal of research has shown workplace mood and happiness can be influenced by a range of management practices and other organizationally-relevant factors (Clark, 2010; Jencks, Perman and Rainwater, 1988; Krekel, Ward and De Neve, 2019). Interestingly, this is a point that is already well understood by firms themselves: in a recent survey of a large sample of U.S. executives, 95% believed that they have “some” or “a high degree” of control when it comes to influencing the happiness of their employees (see HBR Analytical Services, 2020). A growing body of evidence suggests that employee happiness is at least partially determined by structural factors related to how firms organize and manage work as well as the workplace cultures they create. Recent field-experimental evidence shows, for example, significant effects on wellbeing of various management practices – including monitoring, performance information feedback, personal targets, and pro-social incentives (Gosnell, List and Metcalfe, 2020). Further work on aspects of work life—such as manager support and flexibility (Moen et al., 2016), pay inequality (Breza, Kaur and Shamdasani, 2017; Cullen and Perez-Truglia, 2019), gift exchange (DellaVigna et al., 2020), and worker autonomy (Bloom et al., 2014)—has also shown that management practices can have impacts on employee wellbeing.

6.2 Magnitudes & Comparisons with Existing Literature

We estimate an average marginal effect of 3.36 additional sales for each one unit increase in happiness (from a base of around 25 sales per week), suggesting around a 12% effect. This is for a one unit increase in the 0-10 happiness scale, which has a within-worker standard deviation of 2.37 during our study period. It is worth bearing in mind when thinking of magnitudes that a 1 point change would be a relatively large increase. Indeed, the first stage and reduced form estimates in our IV equation show that a one standard deviation increase in visual exposure to weather has around a 0.2 point effect on happiness and around a 2.6% increase in sales per week.

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). In Figure S6, 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. Male workers and workers under 40 in general make around 10% more sales per week on average. Finally, we can compare our instrumented estimate with the non-instrumented one. The non-IV within-person partial correlation estimate is smaller, which is in line with the fact that—as we discuss in more detail in Appendix B—the non-IV equation is likely to be severely downward biased.

We can also compare out estimates with previous studies, in both lab and field. In the experimental set-up of Oswald, Proto and Sgroi (2015), 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 (See Appendix I for further detail). In our setting, a one standard deviation increase in happiness (equivalent to a 2.4 point increase on the 0-10 scale), leads to around a 30 percent increase in sales. In addition to lab experiments, we can compare our estimates with the results of field experiments in which management practices simultaneously impacted employee happiness and productivity. In these contexts we cannot 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 on the impact of working-from-home. 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.

Finally, we can compare our estimates with studies in the field using observational data such as our own. Coviello et al. (2020) also study call center workers in the USA, but with a number of differences to our study. First, the authors instrument for worker engagement—on a 1-to-5 scale corresponding to feeling 'unstoppable', 'good', 'so so', 'exhausted', and 'frustrated'—using weather patterns. Second, as opposed to sales, the outcome is short-run labor supply decisions (i.e. percentage of unproductive time at work) and productivity in terms of work speed (i.e. number of calls per hour). Third, the study focuses largely on customer service representatives rather than sales workers, who perform a somewhat different task. In apparent contradiction to our findings, the authors find that more engaged workers answer fewer calls per hour and a higher percentage of time at work not working. We see the paper as complementary to ours, since despite a number of differences between both settings, they also find positive (though non-significant) effects within their sub-sample of sales workers. One possibility is that positive productivity effects of happiness at work are likely to dominate any negative effect on labor supply in contexts where tasks are well-defined (sales advisors have clear targets, for example). However, for the various reasons outlined in Section 2, productivity effects of happiness are likely to be stronger or weaker in different industries and occupations.[34]

[34]A further possibility in this comparison is that the performance gains from positive mood may play a stronger

6.3 Further Limitations

There are a number of further areas where future research is required in order to better understand the relationship between positive affect and performance. While we study weekly data over a six-month period, future research may benefit from studying higher-frequency measures such as daily surveys – though there are usually trade-offs between frequency of mood data collection and the possible length of the study period, with higher-frequency data typically being more difficult to collect in real-world settings and requiring a much shorter study period. One potential avenue of future research in this direction may be to use things like natural language processing of call transcripts to measure mood, or tonal analysis of recordings, in order to obtain higher-frequency measures of mood without having to interrupt workers with surveys. Indeed, this type of approach may also allow for the measurement of mood on the other side of the market – in our case, customers — opening up a range of further potential research possibilities.

We rely for identification on exposure to weather using a measure of window coverage at the call center level. Unfortunately, the firm does not have data on floor plans and does not routinely collect data on who is sitting where in each building on any given day – such data may in the future provide more power by providing individual-level (possibly even time-varying) variation in proximity to windows. In our setting, however, we found that 96% of employees reported working in an open office environment, such that even though being closer to the window may provide some useful within-call-center variation in visual exposure to weather, this variation is likely to be small relative to the between-call-center variation in window coverage we rely on. Indeed, the very high correlation we obtain between our objective measure of external window share and the subjective measure of natural light (r=0.91), which is even higher than the correlation with subjective internal window share (r=0.83) provides further support for those claims (see Figure 7).

In addition, while we demonstrate a causal effect using visual exposure to weather, further research would benefit from using a range of different mood shocks in order to improve generalizability, particularly using directly managerial shocks – though these typically violate the exclusion restriction by having potential direct effects on happiness, hence the usefulness of using an exogenous shock like weather to pin down the causal happiness channel. Moreover, future research may usefully study the extent to which happiness determines performance in different types of settings – for example in jobs with different kinds of tasks, different levels of monitoring, or in different locations and cultures. As we discussed in Section 2, there are theoretical reasons to believe that happiness may have a stronger effect on some types of tasks than others (and even have a negative effect in certain contexts). Relatedly, there may also be diminishing marginal returns to happiness – that is, future research may explore the question of whether there is some sort of optimal happiness level for productivity or whether more is always better.

Our use of windows for identification, which essentially allow us to "turn on and off" the

role at lower levels of happiness. Indeed, a key difference is that while a significant fraction of workers in our sample report feeling unhappy, 70% of workers in Coviello et al. (2020) report feeling "good" or "unstoppable." However, this is unlikely to explain the difference: as discussed earlier, we do not find significant heterogeneity in the effect of happiness between generally unhappy and generally happy workers (Table S11).

effect of visual aspects of weather, allows us to rule out a number of issues related to the use of weather in mood research. However, there may always be alternative explanations that we are not able to fully rule out using our data and setting and that will, ultimately, require further research. For example, rain may affect productivity through making noise, the extent to which may vary between glass windows and solid walls (though we are able to replicate our effects when dis-aggregating the weather index into parts other than rainfall, such as fog).[35]

7 Conclusion

A long-running literature has sought to explain heterogeneity in productivity across individuals as well as firms. In this paper, we contribute to an extensive literature in management and elsewhere on the relationship between happiness and performance. We show a strong positive impact of employee positive affect on productivity, in a real-world field setting. We follow the existing mood-effects-in-the-field literature by first presenting reduced-form evidence of the effects of weather. We extend this literature to the domain of productivity, and build on it by providing three key pieces of further evidence. First, we clearly pin down mood shocks arising from weather by exploiting variation in exposure to it. Second, we demonstrate using survey data that visual exposure to weather while at work has a significant impact on workers' week-to-week happiness. Third, we are able to provide quasi-experimental estimates of the effect of happiness on sales. Of course, the latter requires us to make a number of further assumptions, which we have sought to discuss as clearly as possible; however, we note that these estimates are subject to usual critiques of IV strategies to estimate causal effects using observational data. Nevertheless, taking the three pieces of evidence as whole, we interpret the findings as suggestive of a strong causal effect of positive affect on productivity.

References

Acemoglu, Daron, and Jörn-Steffen Pischke. 2003. "Minimum wages and on-the-job training." *Research in Labor Economics*, 22(2003): 159–202.

Agarwal, Sumit, Souphala Chomsisengphet, Stephan Meier, and Xin Zou. 2020. "In the mood to consume: Effect of sunshine on credit card spending." *Journal of Banking & Finance*, 121: 105960.

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow. 2020. "The welfare effects of social media." *American Economic Review*, 110(3): 629–76.

Angrist, Joshua D, and Alan B Krueger. 1991. "Does compulsory school attendance affect schooling and earnings?" *The Quarterly Journal of Economics*, 106(4): 979–1014.

Angrist, Joshua D, and Alan B Krueger. 2001. "Instrumental variables and the search for identification: From supply and demand to natural experiments." *Journal of Economic perspectives*, 15(4): 69–85.

[35]In addition to noise, there may be sleep-related effects if window coverage affects people's circadian rhythms – though we rule out physical channels in general, in a more coarse way, by studying sickness and attendance. Moreover, it is not clear why this physical effect should not also play a role when performing more routine tasks, or when dealing with satisfied customers, which is where we find our effects to be smaller.

Angrist, Joshua D, and Guido W Imbens. 1995. "Two-stage least squares estimation of average causal effects in models with variable treatment intensity." *Journal of the American statistical Association*, 90(430): 431–442.

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of causal effects using instrumental variables." *Journal of the American statistical Association*, 91(434): 444–455.

Barsade, Sigal G, and Donald E Gibson. 2007. "Why does affect matter in organizations?" *Academy of management perspectives*, 21(1): 36–59.

Benjamin, Daniel J, Jakina Debnam Guzman, Marc Fleurbaey, Ori Heffetz, and Miles S Kimball. 2021. "What do happiness data mean? Theory and Survey Evidence." *NBER Working Paper 28438*.

Bloom, Nicholas, James Liang, John Roberts, and Zhichun Jenny Ying. 2014. "Does working from home work? Evidence from a Chinese experiment." *Quarterly Journal of Economics*, 130(1): 165–218.

Brayfield, Arthur H, and Walter H Crockett. 1951. "An index of job satisfaction." *Journal of Applied Psychology*, 35: 307–311.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani. 2017. "The morale effects of pay inequality." *Quarterly Journal of Economics*, 133(2): 611–663.

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

Bryson, Alex, and George MacKerron. 2016. "Are you happy while you work?" *The Economic Journal*, 127(599): 106–125.

Carnevale, Peter JD, and Alice M Isen. 1986. "The influence of positive affect and visual access on the discovery of integrative solutions in bilateral negotiation." *Organizational behavior and human decision Processes*, 37(1): 1–13.

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2016. "Particulate pollution and the productivity of pear packers." *American Economic Journal: Economic Policy*, 8(3): 141–69.

Chang, Tom Y, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2019. "The effect of pollution on worker productivity: evidence from call center workers in China." *American Economic Journal: Applied Economics*, 11(1): 151–72.

Clark, Andrew E. 2010. "Work, jobs, and well-being across the millennium." *International differences in well-being*, 436–468.

Côté, Stéphane. 1999. "Affect and performance in organizational settings." *Current Directions in Psychological Science*, 8(2): 65–68.

Coviello, Decio, Erika Deserranno, Nicola Persico, and Paola Sapienza. 2020. "Effect of Mood and Worker Incentives on Workplace Productivity." *Mimeo*.

Cullen, Zoë, and Ricardo Perez-Truglia. 2019. "How much does your boss make? the effects of salary comparisons."

Cunningham, Michael R. 1979. "Weather, mood, and helping behavior: Quasi experiments with the sunshine samaritan." *Journal of personality and social psychology*, 37(11): 1947.

DellaVigna, Stefano, John A List, Ulrike Malmendier, and Gautam Rao. 2020. "Estimating social preferences and gift exchange with a piece-rate design." *CEPR Discussion Paper No. DP14931.*

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

Denissen, Jaap JA, Ligaya Butalid, Lars Penke, and Marcel AG Van Aken. 2008. "The effects of weather on daily mood: a multilevel approach." *Emotion*, 8(5): 662.

Diener, Ed, Eunkook M Suh, Richard E Lucas, and Heidi L Smith. 1999. "Subjective well-being: Three decades of progress." *Psychological Bulletin,*, 125(2): 276.

Edmans, Alex. 2012. "The link between job satisfaction and firm value, with implications for corporate social responsibility." *The Academy of Management Perspectives*, 26(4): 1–19.

Edmans, Alex, Diego Garcia, and Øyvind Norli. 2007. "Sports sentiment and stock returns." *The Journal of Finance*, 62(4): 1967–1998.

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

Estrada, Carlos A, Alice M Isen, and Mark J Young. 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, John, Robert Metcalfe, and Mark Wooden. 2016. "Subjective wellbeing: why weather matters." *Journal of the Royal Statistical Society Series A*, 179(1): 203–228.

Finkelstein, Amy, Erzo FP Luttmer, and Matthew J Notowidigdo. 2013. "What good is wealth without health? The effect of health on the marginal utility of consumption." *Journal of the European Economic Association*, 11(suppl_1): 221–258.

Fisher, Cynthia D. 2000. "Mood and emotions while working: missing pieces of job satisfaction?" *Journal of Organizational Behavior*, 21(2): 185–202.

Fisher, Vivian Ezra, and Joseph Vorhees Hanna. 1931. *The dissatisfied worker.* New York:Macmillan.

Flynn, Sean Masaki, and Adam Eric Greenberg. 2012. "Does Weather Actually Affect Tipping? An Empirical Analysis of Time-Series Data 1." *Journal of Applied Social Psychology*, 42(3): 702–716.

Forgas, Joseph P. 1995. "Mood and judgment: the affect infusion model (AIM)." *Psychological Bulletin,*, 117(1): 39.

Fredrickson, Barbara L. 2001. "The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions." *American psychologist*, 56(3): 218.

Frijda, Nico H. 1986. *The Emotions.* Cambridge University Press.

Frijters, Paul, Chitwan Lalji, and Debayan Pakrashi. 2020. "Daily weather only has small effects on wellbeing in the US." *Journal of Economic Behavior & Organization*, 176: 747–762.

Gallen, Trevor. 2020. "Broken Instruments." *SSRN Working Paper 3671850.*

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

Graff Zivin, Joshua, and Matthew Neidell. 2012. "The impact of pollution on worker productivity." *American Economic Review*, 102(7): 3652–73.

Gubler, Timothy, Ian Larkin, and Lamar Pierce. 2018. "Doing well by making well: The impact of corporate wellness programs on employee productivity." *Management Science*, 64(11): 4967–4987.

Guven, Cahit. 2012. "Reversing the question: Does happiness affect consumption and savings behavior?" *Journal of Economic Psychology*, 33(4): 701–717.

Guven, Cahit, and Indrit Hoxha. 2015. "Rain or shine: Happiness and risk-taking." *The Quarterly Review of Economics and Finance*, 57: 1–10.

HBR Analytical Services. 2020. "Cultivating Workforce Well-Being to Drive Business Value." *https://bg.hbr.org/resources/pdfs/comm/workplacewellbeing.pdf.*

Hersey, Rexford B. 1932. "Rate of Production and Emotional State." *Personnel Journal*, 335–364.

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

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

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

Huang, Jiekun, Nianhang Xu, and Honghai Yu. 2020. "Pollution and performance: Do investors make worse trades on hazy days?" *Management Science*, 66(10): 4455–4476.

Hu, Maggie Rong, and Adrian D Lee. 2020. "Outshine to Outbid: Weather-Induced Sentiment and the Housing Market." *Management Science*, 66(3): 1440–1472.

Iaffaldano, Michelle T, and Paul M Muchinsky. 1985. "Job satisfaction and job performance: A meta-analysis." *Psychological Bulletin,*, 97(2): 251.

Isen, Alice M. 1970. "Success, failure, attention, and reaction to others: The warm glow of success." *Journal of personality and social psychology*, 15(4): 294.

Isen, Alice M, and Barbara Means. 1983. "The influence of positive affect on decision-making strategy." *Social cognition*, 2(1): 18–31.

Isen, Alice M, and Kimberly A Daubman. 1984. "The influence of affect on categorization." *Journal of personality and social psychology*, 47(6): 1206.

Isen, Alice M, Andrew S Rosenzweig, and Mark J Young. 1991. "The influence of positive affect on clinical problem solving." *Medical Decision Making*, 11(3): 221–227.

Isen, Alice M, and Robert A Baron. 1991. "Positive affect as a factor in organizational-behavior." *Research in organizational behavior*, 13: 1–53.

Isen, Alice M, Kimberly A Daubman, and Gary P Nowicki. 1987. "Positive affect facilitates creative problem solving." *Journal of personality and social psychology*, 52(6): 1122.

Jencks, Christopher, Lauri Perman, and Lee Rainwater. 1988. "What is a good job? A new measure of labor-market success." *American Journal of Sociology*, 93(6): 1322–1357.

Jones, Damon, David Molitor, and Julian Reif. 2019. "What do workplace wellness programs do? Evidence from the Illinois workplace wellness study." *The Quarterly Journal of Economics*, 134(4): 1747–1791.

Judge, Timothy A, Carl J Thoresen, Joyce E Bono, and Gregory K Patton. 2001. "The job satisfaction–job performance relationship: A qualitative and quantitative review." *Psychological Bulletin,*, 127(3): 376.

Kahneman, Daniel, Alan B Krueger, David A Schkade, Norbert Schwarz, and Arthur A Stone. 2004. "A survey method for characterizing daily life experience: The day reconstruction method." *Science*, 306(5702): 1776–1780.

Kahneman, Daniel, Peter Wakker, and Rakesh Sarin. 1997. "Back to Bentham? Explorations of Experienced Utility." *Quarterly Journal of Economics*, 112(2): 375–406.

Keller, Matthew C, Barbara L Fredrickson, Oscar Ybarra, Stéphane Côté, Kareem Johnson, Joe Mikels, Anne Conway, and Tor Wager. 2005. "A warm heart and a clear head: The contingent effects of weather on mood and cognition." *Psychological science*, 16(9): 724–731.

Kemp, Simon, Christopher DB Burt, and Laura Furneaux. 2008. "A test of the peak-end rule with extended autobiographical events." *Memory & Cognition*, 36(1): 132–138.

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

Knight, Andrew P, Jochen I Menges, and Heike Bruch. 2018. "Organizational affective tone: A meso perspective on the origins and effects of consistent affect in organizations." *Academy of Management Journal*, 61(1): 191–219.

Koys, Daniel J. 2001. "The effects of employee satisfaction, organizational citizenship behavior, and turnover on organizational effectiveness: A unit-level, longitudinal study." *Personnel Psychology*, 54(1): 101–114.

Krekel, Christian, George Ward, and Jan-Emmanuel De Neve. 2019. "What makes for a good job? Evidence using subjective wellbeing data." In *The Economics of Happiness.* , ed. Mariano Rojas, 241–268. Springer.

Kripke, Daniel F. 1998. "Light treatment for nonseasonal depression: speed, efficacy, and combined treatment." *Journal of affective disorders*, 49(2): 109–117.

Krueger, Alan B., and Arthur A. Stone. 2014. "Progress in Measuring Subjective Well-Being: moving toward national indicators and policy evaluations." *Science*, 346(6205): 42–43.

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

Krueger, Alan B, Daniel Kahneman, David Schkade, Norbert Schwarz, and Arthur A Stone. 2009. "National time accounting: The currency of life." In *Measuring the subjective well-being of nations: National accounts of time use and well-being.* 9–86. University of Chicago Press.

Kunin, Theodore. 1955. "The Construction of a New Type of Attitude Measure 1." *Personnel Psychology*, 8(1): 65–77.

Lawler, Edward E, and Lyman W Porter. 1967. "The effect of performance on job satisfaction." *Industrial relations: A journal of Economy and Society*, 7(1): 20–28.

Leppämäki, S, T Partonen, and J Lönnqvist. 2002. "Bright-light exposure combined with physical exercise elevates mood." *Journal of Affective Disorders*, 72(2): 139–144.

Li, Chenxi, Xueming Luo, Cheng Zhang, and Xiaoyi Wang. 2017. "Sunny, rainy, and cloudy with a chance of mobile promotion effectiveness." *Marketing Science*, 36(5): 762–779.

Lin, Wei, and Jeffrey M Wooldridge. 2019. "Testing and correcting for endogeneity in nonlinear unobserved effects models." In *Panel Data Econometrics.* , ed. Mike Tsionas, 21–43. Elsevier.

Locke, Edwin A. 1969. "What is job satisfaction?" *Organizational Behavior and Human Performance*, 4: 309–336.

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

Lucas, Richard E, and Ed Diener. 2003. "The happy worker: Hypotheses about the role of positive affect in worker productivity." *Personality and work*, 30–59.

Meier, Armando N, Lukas Schmid, and Alois Stutzer. 2019. "Rain, emotions and voting for the status quo." *European Economic Review*, 119: 434–451.

Mellon, Jonathan. 2020. "Rain, Rain, Go away: 137 potential exclusion-restriction violations for studies using weather as an instrumental variable." *SSRN Working Paper 3715610.*

Miner, Andrew G, and Theresa M Glomb. 2010. "State mood, task performance, and behavior at work: A within-persons approach." *Organizational Behavior and Human Decision Processes*, 112(1): 43–57.

Moen, Phyllis, Erin L Kelly, Wen Fan, Shi-Rong Lee, David Almeida, Ellen Ernst Kossek, and Orfeu M Buxton. 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.

OECD. 2013. "Guidelines on Measuring Subjective Wellbeing." *https://www.oecd.org/statistics/oecd-guidelines-on-measuring-subjective-well-being-9789264191655-en.htm.*

Oswald, Andrew J, Eugenio Proto, and Daniel Sgroi. 2015. "Happiness and productivity." *Journal of Labor Economics*, 33(4): 789–822.

Park, R Jisung, Joshua Goodman, Michael Hurwitz, and Jonathan Smith. 2020. "Heat and learning." *American Economic Journal: Economic Policy*, 12(2): 306–39.

Rind, Bruce. 1996. "Effect of beliefs about weather conditions on tipping." *Journal of Applied Social Psychology*, 26(2): 137–147.

Rothbard, Nancy P, and Steffanie L Wilk. 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.

Russell, James A, and Lisa Feldman Barrett. 1999. "Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant." *Journal of personality and social psychology*, 76(5): 805.

Saunders, Edward M. 1993. "Stock prices and Wall Street weather." *American Economic Review*, 83(5): 1337–1345.

Schneider, Daniel, and Kristen Harknett. 2019. "Consequences of routine work-schedule instability for worker health and well-being." *American Sociological Review*, 84(1): 82–114.

Singh, Jagdip. 2000. "Performance productivity and quality of frontline employees in service organizations." *Journal of marketing*, 64(2): 15–34.

Staw, Barry M, and Sigal G Barsade. 1993. "Affect and managerial performance: A test of the sadder-but-wiser vs. happier-and-smarter hypotheses." *Administrative Science Quarterly*, 304–331.

Staw, Barry M, Robert I Sutton, and Lisa H Pelled. 1994. "Employee positive emotion and favorable outcomes at the workplace." *Organization Science*, 5(1): 51–71.

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

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

Thomas, David L, and Ed Diener. 1990. "Memory accuracy in the recall of emotions." *Journal of personality and social psychology*, 59(2): 291.

Walsh, Lisa C, Julia K Boehm, and Sonja Lyubomirsky. 2018. "Does happiness promote career success? Revisiting the evidence." *Journal of Career Assessment*, 26(2): 199–219.

Wright, Thomas A, and Russell Cropanzano. 2000. "Psychological well-being and job satisfaction as predictors of job performance." *Journal of occupational health psychology*, 5(1): 84.

Zelenski, John M, Steven A Murphy, and David A Jenkins. 2008. "The happy-productive worker thesis revisited." *Journal of Happiness Studies*, 9(4): 521–537.

Appendix

Table A1: Summary Statistics

	N	Mean	Standard Deviation		
			Overall	Between	Within
Happiness	12,282	4.01	3.44	2.55	2.37
Sales	12,282	25.25	19.16	14.89	12.41
Selling Time	12,282	19.99	8.18	5.69	6.38
Internal Shrinkage	12,282	10.52	13.30	10.11	10.74
Customer Satisfaction	10,120	7.90	2.26	1.20	2.06
Adherence	12,174	91.99	5.48	3.58	4.34
Calls per Hour	12,282	4.96	1.62	1.16	1.14
Conversion Rate	11,850	26.63	17.98	15.77	10.59
Sick Leave	12,279	1.58	7.73	3.37	7.26
Attendance	12,279	92.58	14.14	6.55	13.15
Breaks (hrs)	12,282	3.77	1.49	1.15	1.09
Overtime (hrs)	12,282	0.15	0.98	0.48	0.89
Paid Time Off (hrs)	12,282	1.07	2.99	1.36	2.78
Visual Weather Index	12,282	4.06	1.78	1.18	1.36
Visual Exposure to Weather	12,282	0.77	0.62	0.53	0.32
Windows (% of wall surface)	1,157	0.20		0.14	
Age	1,157	33.81		10.41	
Female	1,157	0.41		0.50	
Tenure	1,157	4.99		7.17	
Left Firm During Study	1,157	0.04		0.26	

Online Supplementary Materials

Appendix A Enrollment/Attrition/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. One concern is 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 S1 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, local gloomy weather (multiplied or not by window share) 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. Non-response is, however, positively related to the number of hours worked during the week and whether or not they work on Thursdays or Fridays, 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.[36]

Table S1: Predictors of attrition/non-response

	Responded to Survey = 1					
	(1)	(2)	(3)	(4)	(5)	(6)
Sales	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Selling time	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Not working on thursday/friday	-0.153***	-0.153***	-0.153***	-0.153***	-0.153***	-0.153***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Weather index		0.001			0.001	
		(0.002)			(0.002)	
Weather * windows			0.002			0.002
			(0.008)			(0.008)
Team happiness (excl. worker)				-0.001	-0.001	-0.001
				(0.001)	(0.001)	(0.001)
Observations	33725	33725	33725	33725	33725	33725
R^2	0.448	0.448	0.448	0.448	0.448	0.448

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

[36] One approach to dealing with non-response to the survey would be to impute 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. 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.

Table S2: Predictors of study participation: Extensive Margin

	Participated in the study = 1				
	(1)	(2)	(3)	(4)	(5)
Age	-0.002*	-0.003**	-0.003**	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female	-0.003	-0.005	-0.005	-0.003	-0.002
	(0.019)	(0.022)	(0.022)	(0.023)	(0.025)
Left firm during study	-0.127***	-0.131***	-0.133***	-0.135***	-0.131***
	(0.034)	(0.036)	(0.037)	(0.035)	(0.034)
Tenure (months)	-0.019*	-0.018	-0.018	-0.020	-0.020
	(0.010)	(0.012)	(0.012)	(0.012)	(0.012)
Mean selling hours during study	0.009***	0.010***	0.010***	0.010***	0.007**
	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Mean sales during study		-0.001	-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.002)	(0.002)
Mean team happiness (excl. worker)		-0.007	-0.007	-0.004	-0.005
		(0.006)	(0.006)	(0.007)	(0.007)
Mean bright/gloomy weather (call center)			-0.006		
			(0.005)		
Building window coverage			-0.073		
			(0.060)		
Not working on thursday/friday					-0.206***
					(0.051)
Call center dummies	No	No	No	Yes	Yes
Observations	1793	1762	1762	1762	1762
R^2	0.054	0.057	0.058	0.061	0.069

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

Table S3: Predictors of study participation: Intensive Margin

	# Waves responded to survey				
	(1)	(2)	(3)	(4)	(5)
Age	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Female	-0.080**	-0.083**	-0.084**	-0.071**	-0.071**
	(0.037)	(0.036)	(0.037)	(0.034)	(0.031)
Left firm during study	-0.862***	-0.850***	-0.848***	-0.857***	-0.865***
	(0.039)	(0.044)	(0.045)	(0.041)	(0.041)
Tenure (months)	-0.041**	-0.045**	-0.048**	-0.057***	-0.063***
	(0.020)	(0.021)	(0.022)	(0.021)	(0.021)
Mean selling hours during study	0.022***	0.022***	0.022***	0.025***	0.018***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)
Mean sales during study		-0.000	-0.000	-0.001	-0.001
		(0.002)	(0.001)	(0.001)	(0.001)
Mean team happiness (excl. worker)		-0.027	-0.027	-0.021	-0.022
		(0.017)	(0.017)	(0.019)	(0.019)
Mean bright/gloomy weather (call center)			0.019		
			(0.012)		
Building window coverage			0.028		
			(0.094)		
Not working on thursday/friday					-0.546***
					(0.126)
Call center dummies	No	No	No	Yes	Yes
Observations	1438	1413	1413	1413	1413

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 Evidence of Downward Bias in Baseline Equation

There are a number of reasons to be concerned that β in equation (1) may be biased. One initial reason is measurement error in the survey, which will bias the coefficients downward. A second reason is the existence of omitted variables that may affect both reported happiness and productivity week-to-week. A significant further empirical concern is that, even within-workers over time, a change in SWB is likely to be endogenous to performance. 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 monetary or non-monetary rewards, for example from their colleagues or managers or simply enjoy successfully completing tasks. This alone could explain their higher happiness, in which case the β coefficient will be biased upward. Second, what makes workers happier could lower their productivity, depressing the true β coefficient. One major candidate is the quantity of work. Over-work can lead to stress and anxiety, which are both strongly negatively correlated with happiness. Equally, doing more work may simply be less enjoyable. If this is the case, the coefficient will be biased even further downward.

Table S4: Evidence of Downward Bias in OLS: Happiness and Number of Calls

	Happiness	Sales
	(1)	(2)
	OLS-FE	Poisson-FE
Total number of weekly calls (ln)	-0.5217***	0.0837*
	(0.1073)	(0.0462)
Observations	12,282	12,282
Employees	1,157	1,157
R^2	0.535	
Pseudo-R^2		0.620
Week FEs	✓	✓
Individual FEs	✓	✓

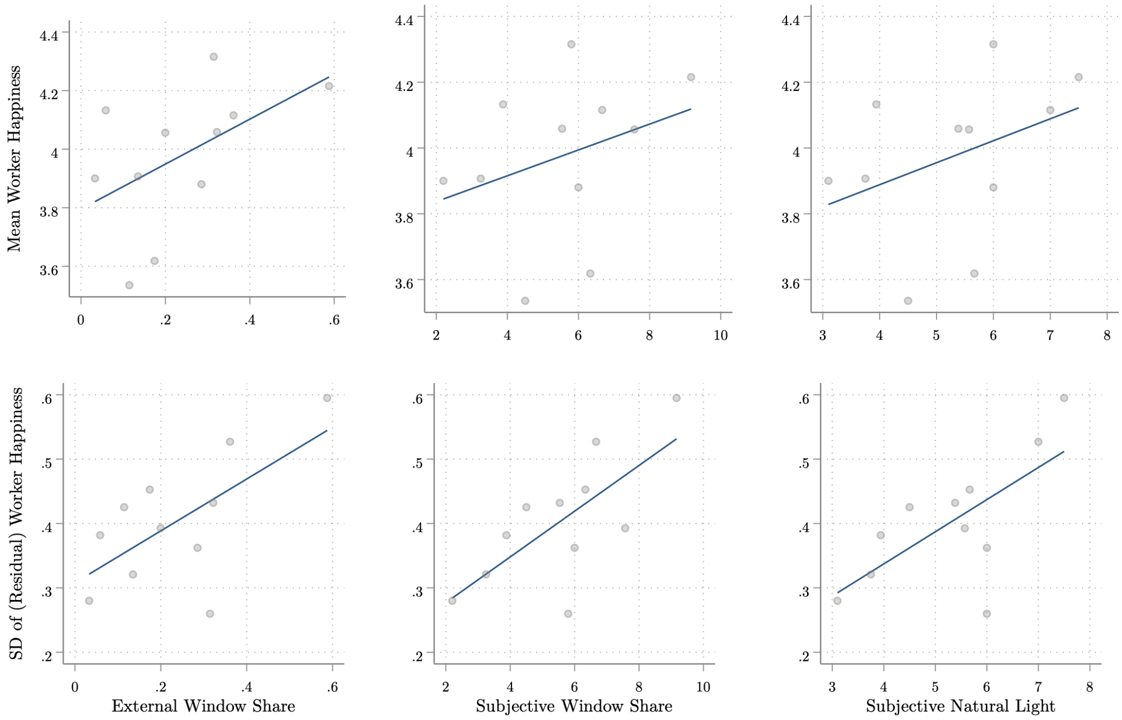
Notes: Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include scheduling controls and day of week dummies for response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

In a customer-facing retail setting, we expect a strongly downward bias. In call centers, there is a pressure to deal with a constant flow of incoming calls, often initiated by unhappy customers (e.g. re-contracting), which led management scholars (Singh, 2000) to discuss the conflict between work pressure (e.g., calls per hour) and the quality of work (e.g., call conversion). Although pay is partly based on performance in our setting, this is a slow moving bonus scheme that makes up only a relatively small amount of total earnings, and so is unlikely to play a major role. We thus expect higher work volumes to lead to lower happiness among workers. We use our data and setting to provide direct evidence of this. In Table S4, 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 included in equation (1). We document a strong negative impact on workers' happiness of total number of weekly calls answered. Answering twice as many calls leads to a fall in happiness of nearly 0.52 points, which corresponds to a happiness drop of about 20% of a standard deviation within workers. This number even goes to a 30% drop if we control for sales to capture the fraction of "successful" calls. Because happier workers are also answering fewer calls, our initial estimates from equation (1) are thus likely to be strongly biased downward.[37]

[37] Note that despite the positive relationship between total number of calls and sales, we still find that happiness

Appendix C Additional Tables and Figures for IV Set-Up

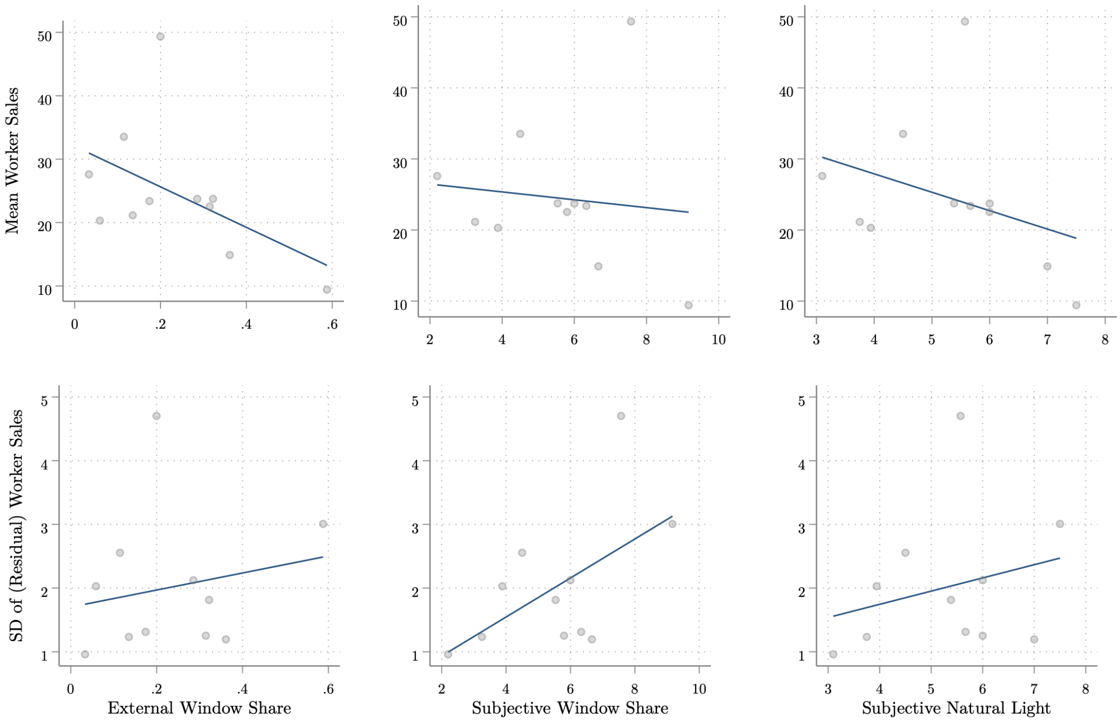
Figure S1: Correlation Between Window Share, Mean Happiness and Standard Deviation in Week-to-Week (Residual) Happiness at Location



Note: Upper-panels: correlation between mean worker happiness and windows at location. Lower-panels: correlation between standard-deviation week-to-week (residual) worker happiness and windows at location. The residual happiness is estimated from a regression with worker fixed effects and time-varying controls other than local weekly weather shocks. Each dot represents a location.

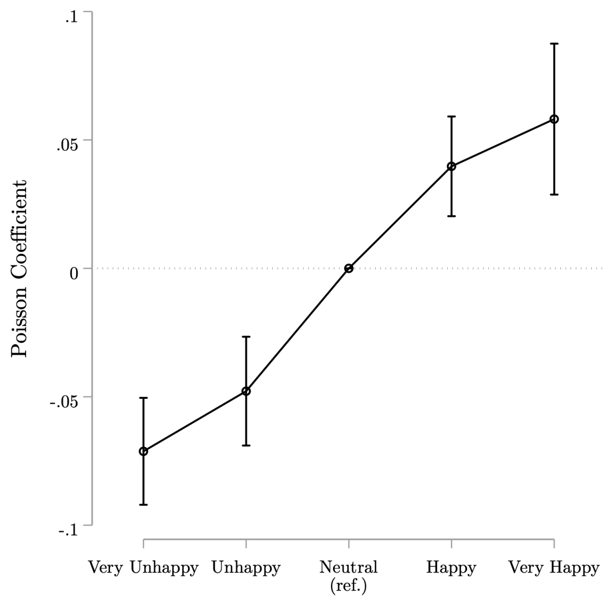
improves sales using equation (1). This is consistent with sales calls generating less emotional distress than non-sales calls, or with happiness affecting performance mostly through the conversion of calls into sales, a possibility we later explore in the paper.

Figure S2: Correlation Between Window Share, Mean Sales and Standard Deviation in Week-to-Week (Residual) Sales at Location



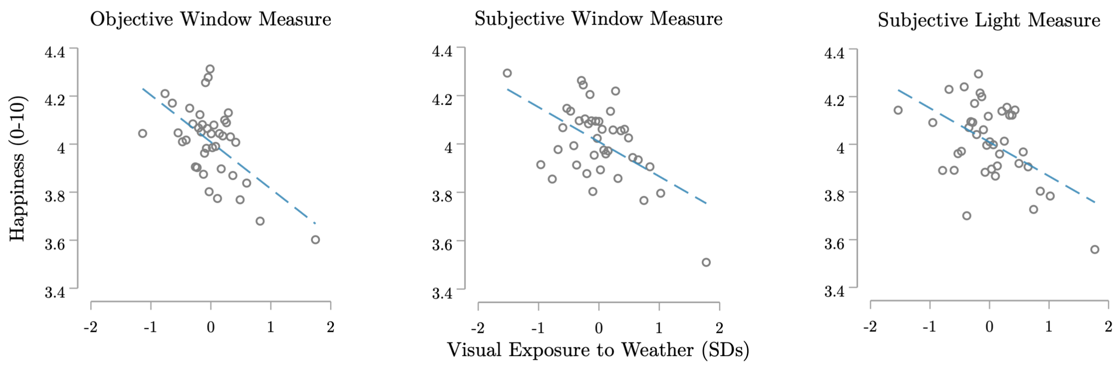
Note: Upper-panels: correlation between mean worker sales and windows at location. Lower-panels: correlation between standard-deviation week-to-week (residual) worker happiness and sales at location. The residual happiness is estimated from a regression with worker fixed effects and time-varying controls other than local weekly weather shocks. Each dot represents a location.

Figure S3: 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.

Figure S4: IV First Stage: Graphical Representation



Note: Figure shows binned scatter plots of the relationship between visual exposure to weather and happiness, adjusting for individual and week fixed effects as well as the full set of further controls. Visual exposure is the weather index multiplied by window coverage, which is measured in three different ways (1 objective and 2 subjective) – see text for more details.

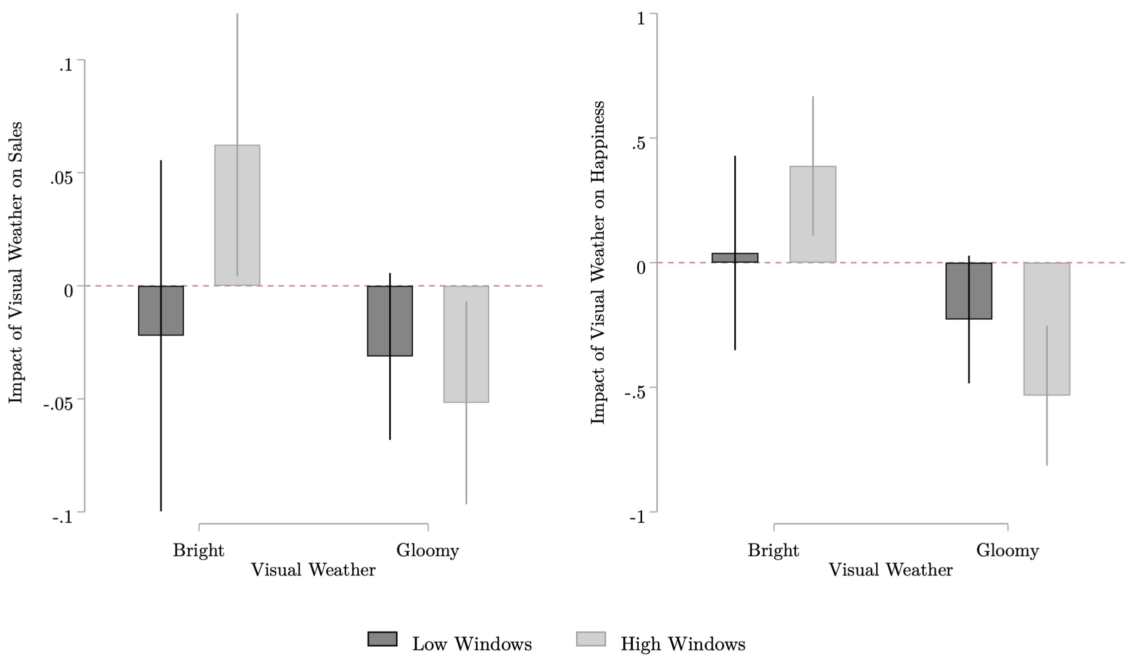
Appendix D Additional Robustness

Table S5: Daily Call-Center-Level Regressions

	Local Demand (Poisson-FE)	Local Speed (OLS-FE)
Visual Weather Index	-0.0054	0.0055
	(0.0048)	(0.0049)
Observations	1,415	1,415
R^2		0.839
Pseudo-R^2	0.118	

Notes: Poisson-FE and 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 speed is the mean length of call per employee-day in each call center.

Figure S5: Impact of Visually Bright or Gloomy Weather on Weekly Happiness and Sales



Note: Left: Poisson-FE models reported with weekly sales as dependent variable. Right: OLS-FE models reported with weekly happiness as dependent variable. Mostly bright weather defined as less than 2 on the weather index and mostly gloomy weather defined as more than 6. The models are estimated separately on each sub-sample of high vs. low window coverage locations, defined as above vs. below median window coverage. All models include individual and week fixed effects, work schedule controls, and dummies for day of week of response to survey. 90% confidence intervals are constructed using robust standard errors, adjusted for two-way clustering on individuals and location-week.

Table S6: Results When Using Weather Index Dis-Aggregated

	1st Stage IV	2nd Stage IV
	(1)	(2)
Happiness (0-10)		0.1219***
		(0.0422)
Exposure to Fog	-0.4198***	
	(0.1404)	
Exposure to Rain	-0.2578**	
	(0.1069)	
Exposure to Snow	-0.3084	
	(0.2421)	
Observations	12,282	12,282
1st Stage F-Stat	6.95	

Notes: 2SLS model reported. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include individual and week fixed effects, work schedule controls, and dummies for day of week of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S7: Visual Weather Index vs. Local Share of Sick Employees

	Local Sickness	
	(1)	(2)
Panel A: Daily Call-Center-Level Regressions		
Visual Weather Index (Daily)	-0.1070	
	(0.1052)	
Visual Weather Index (Daily, 1-Day Lag)		-0.1191
		(0.1180)
Observations	1,428	1,139
R^2	0.335	0.336
Panel B: Weekly Call-Center-Level Regressions		
Visual Weather Index (Weekly)	-0.1116	
	(0.0713)	
Visual Weather Index (1-Week Lag)		-0.1012
		(0.0873)
Observations	289	276
R^2	0.199	0.213

Notes: OLS-FE models reported. Robust standard errors in parentheses, clustered on call centers and time. All models include call center and time fixed effects.

Table S8: Sensitivity to Weather: Heterogeneous Effects

	Happiness					
	(1)	(2)	(3)	(4)	(5)	(6)
Gloomy Weather	-0.0617***	-0.0438*	-0.0324	-0.0658***	-0.0850***	-0.0711**
	(0.0207)	(0.0253)	(0.0596)	(0.0247)	(0.0287)	(0.0295)
Weather Interaction:						
× Female Worker		-0.0449				
		(0.0358)				
× Worker's Age			-0.0008			
			(0.0016)			
× Worker's Tenure (Years)				0.0007		
				(0.0021)		
× Worker Avg Sickness > Median					0.0477	
					(0.0353)	
× # Sales						0.0004
						(0.0009)
Observations	12,282	12,282	12,282	12,282	12,280	12,282
R^2	0.534	0.534	0.534	0.534	0.534	0.536

Notes: OLS-FE models reported. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. 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 S9: Visual Exposure to Weather: 1st Stage Treatment Effects (LATE)

	Happiness					
	(1)	(2)	(3)	(4)	(5)	(6)
Visual Exposure to Weather (SDs)	-0.1950***	-0.1763***	-0.0841	-0.1751***	-0.2331***	-0.2079***
	(0.0442)	(0.0568)	(0.1423)	(0.0501)	(0.0630)	(0.0562)
Weather Interaction:						
× Female Worker		-0.0490				
		(0.0827)				
× Worker's Age			-0.0032			
			(0.0038)			
× Worker's Tenure (Years)				-0.0045		
				(0.0053)		
× Worker Avg Sickness > Median					0.0776	
					(0.0795)	
× # Sales						0.0008
						(0.0020)
Observations	12,282	12,282	12,282	12,282	12,280	12,282
R^2	0.534	0.534	0.534	0.534	0.534	0.537

Notes: OLS-FE models reported. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. 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 S10: Analysis of Potential Seasonal Effects

	Happiness		Sales	
	(1)	(2)	(3)	(4)
Visual Exposure to Gloomy Weather	-0.3122***	-0.2827**	-0.0413**	-0.0501*
	(0.0654)	(0.1128)	(0.0187)	(0.0290)
Interaction with Month (Ref. Month: January 2018)				
December 2017		-0.0357		0.0058
		(0.1287)		(0.0277)
November 2017		-0.0096		0.0246
		(0.1293)		(0.0301)
October 2017		-0.1376		-0.0234
		(0.1343)		(0.0325)
September 2017		0.0517		0.0380
		(0.1547)		(0.0348)
August 2017		0.1016		0.0156
		(0.1518)		(0.0307)
July 2017		0.1679		0.0575
		(0.2062)		(0.0375)

Notes: Columns (1)-(2) estimated using OLS-FE. Columns (3)-(4) estimated using Poisson-FE. All models include individual and week fixed effects, work schedule controls, and day of week dummies for response to survey. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S11: Impact of Happiness on Sales Performance: Sub-Group Analysis

	All Workers	Unhappy Workers	Happy Workers	Unproductive Workers	Productive Workers
	(1)	(2)	(3)	(4)	(5)
Happiness (0-10)	0.1376***	0.1290**	0.1486**	0.1467*	0.1307***
	(0.0469)	(0.0525)	(0.0652)	(0.0757)	(0.0490)
Observations	12,239	6,142	6,097	6,031	6,208
1st Stage F-Stat	18.85	12.06	9.30	6.85	14.42

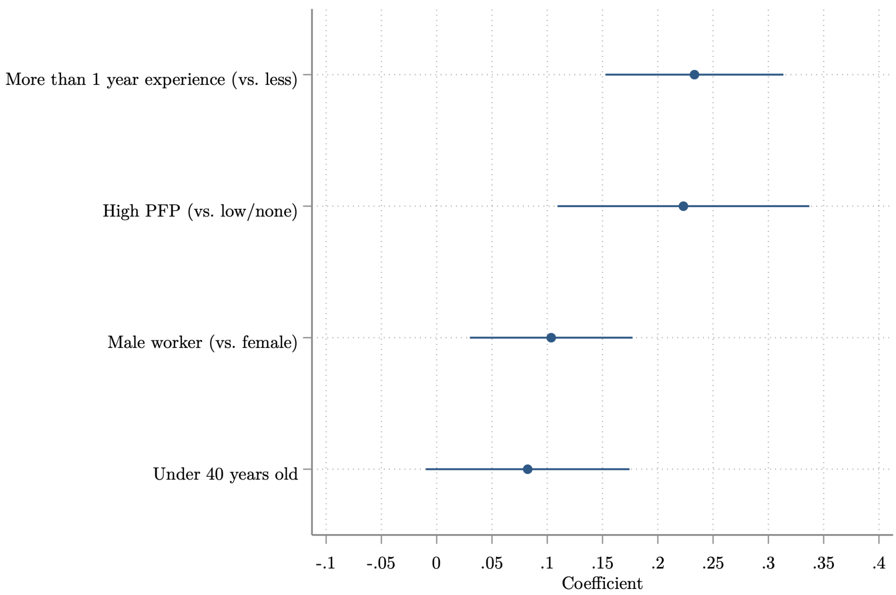
Notes: Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. Poisson-FE models reported. All models include individual and week fixed effects, work schedule controls, and dummies for day of week of response to survey. Sample is split above/below median within each call center in terms of average sales per hour and happiness over the whole study period. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S12: Placebo Regressions: Week Before (Lag) vs. After (Lead)

	Happiness (weekly)		Sales (weekly)	
	(1)	(2)	(3)	(4)
Visual Exposure to Weather (SDs)	-0.1950***	-0.1925***		
	(0.0442)	(0.0443)		
Visual Exposure to Weather (SDs, 1-week lag)		0.0437		
		(0.0457)		
Visual Exposure to Weather (SDs, 1-week lead)		-0.0282		
		(0.0548)		
Happiness (0-10)			0.0141***	0.0168***
			(0.0014)	(0.0022)
Happiness (1-week lag)				-0.0014
				(0.0023)
Happiness (1-week lead)				0.0019
				(0.0020)
Observations	12,282	12,137	12,282	4,522

Notes: Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include individual and week fixed effects, work schedule controls, and day of week dummies for response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure S6: Effect Size Bench-marking



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

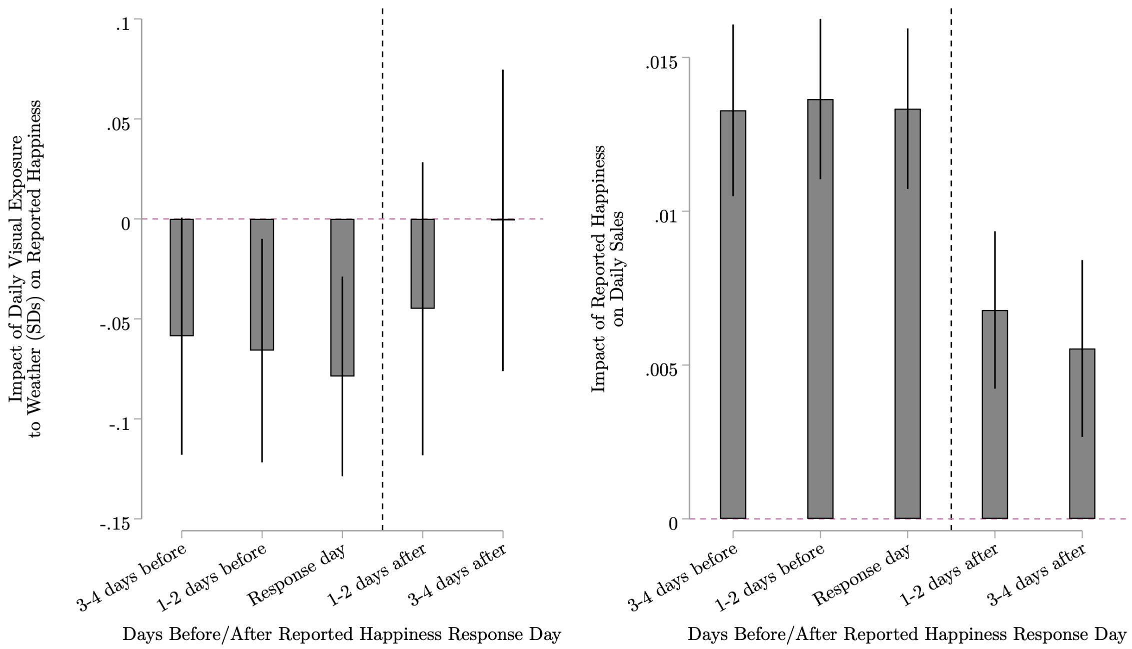
Appendix E Analysis using Daily Data

Table S13: Reduced Form Effects of Daily Weather on Daily Sales

	Daily Sales (Poisson-FE)			
	(1)	(2)	(3)	(4)
Visual Exposure to Weather (Daily)	-0.0094*	-0.0086*		
	(0.0048)	(0.0050)		
Visual Exposure to Weather (Weekly)			-0.0238**	-0.0234**
			(0.0109)	(0.0109)
Observations	42,111	42,111	42,111	42,111
Time FEs	Week	Day	Week	Day

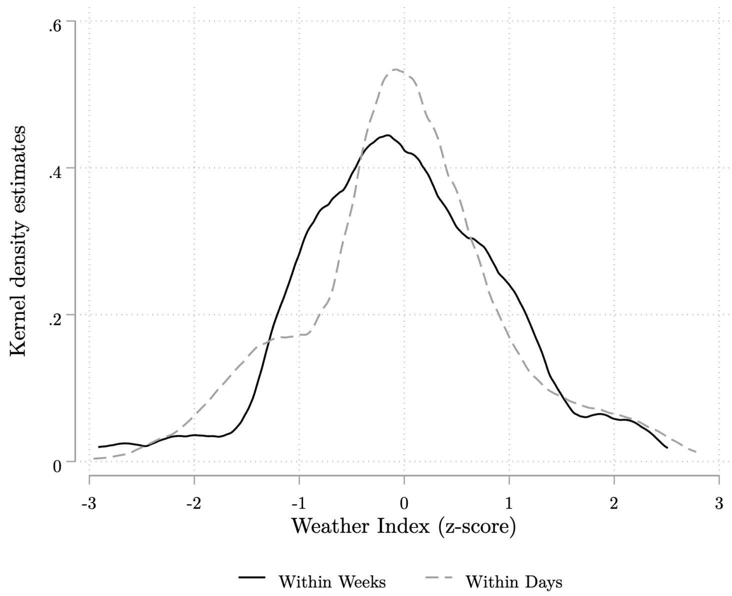
Notes: Poisson-FE models reported. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include individual and date fixed effects, along with daily work schedule controls. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure S7: Impact of Visual Exposure to Weather on Happiness Before/After Response Day to Happiness Survey



Note: Left: Impact of daily visual exposure to weather (SDs) on reported happiness estimated using OLS-FE. Right: Impact of reported happiness on daily sales estimated using Poisson-FE. 90% confidence intervals are constructed using robust standard errors, adjusted for two-way clustering on individuals and location-week.

Figure S8: Distribution of (Residual) Weather Index Across Locations (Within Weeks vs. Days)



Note: Kernel density distribution of residual weather index across locations within weeks vs. within days from a regression of the (normalized) weather index on week and location fixed effects (bin width = 0.2).

Appendix F Suggestive Evidence on Mechanisms

Table S14: Impact of Adherence and Speed on Sales Conditional on Labor Supply

	Sales		
	(1)	(2)	(3)
Labor Productivity:			
Adherence (Met Target=1)	0.0034		0.0034
	(0.0079)		(0.0079)
Total number of calls per hour (ln)		-0.0579	-0.0616*
		(0.0365)	(0.0365)
Labor Supply:			
Total number of selling hours (ln)	0.9831***	0.9926***	0.9967***
	(0.0184)	(0.0135)	(0.0135)
Internal shrinkage	0.0007	0.0000	0.0007
	(0.0005)	(0.0005)	(0.0005)
Observations	12,169	12,100	12,033

Notes: Poisson-FE models reported. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include individual and week fixed effects, scheduling controls, and day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S15: Happiness and Customer Satisfaction

	Customer Satisfaction (z-score)			
	(1) OLS-FE	(2) OLS-FE	(3) 2SLS	(4) 2SLS
Visual Exposure to Weather (SDs)	0.0168	0.0176		
	(0.0172)	(0.0171)		
Happiness (0-10)			-0.0824	-0.0886
			(0.0864)	(0.0878)
# Sales		0.0032***		0.0052**
		(0.0011)		(0.0023)
Observations	10,059	10,059	10,059	10,059
R^2	0.173	0.174	-0.055	-0.060
1st Stage F-Stat			20.16	18.59

Notes: Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include scheduling controls and day of week dummies for response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table S16: Effects by Average Weekly Customer Satisfaction

	Sales (Poisson-FE)	
	Red.-Form	CF Poisson-IV
Panel A: Unsatisfied Customers Weeks		
Happiness (0-10)		0.2753***
		(0.0792)
Visual Exposure to Weather (SDs)	-0.0623***	
	(0.0180)	
Observations	4,161	4,161
Panel B: Satisfied Customers Weeks		
Happiness (0-10)		0.1321***
		(0.0511)
Visual Exposure to Weather (SDs)	-0.0382***	
	(0.0147)	
Observations	4,081	4,081
Panel C: Highly Satisfied Customers Weeks		
Happiness (0-10)		-0.0873
		(0.1417)
Visual Exposure to Weather (SDs)	0.0152	
	(0.0243)	
Observations	3,462	3,462

Notes: The first column reports reduced-form Poisson-FE regressions of sales on visual exposure to weather; standard errors are reported in parentheses and are adjusted for two-way clustering at the individual and location-week level. The second column reports control function Poisson-IV models in which happiness is instrumented for using visual exposure to weather; standard errors are computed using the delta method. Sample is divided by terciles of national weekly customer satisfaction. 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 G Supplementary Qualitative Data

In addition to the email survey coordinated with the firm during the study period, we collected supplementary data from both managers and workers.

Worker Survey. We use the worker-level data largely in order to further validate our empirical approach leveraging differential visual exposure to weather conditions. In this instance, we designed a very short survey using *Qualtrics* and then recruited BT call center workers (N=318), following innovate work done by Schneider and Harknett (2019), using targeted advertising on social media. We use two channels. First, for *LinkedIn*, we use the targeting features that allow us to advertise the survey based on users' employer and occupation. Second, for *Facebook*, we contacted the administrators of informal BT employee "groups," some of whom posted a link to our survey so that their members could see it. In our advertisement, we say that "The University of Oxford's Wellbeing Research Centre is conducting a short survey of contact centre workers, past and present, around the UK. We would like to invite you to submit a small amount of information regarding your time working at a contact centre. All information is entirely anonymous. All entrants have the chance to win £100 as a thank your for your participation."[38]

Having clicked through to the survey, respondents are again told the survey will be strictly anonymous and are given general information about the survey, their rights, how their data will be handled, IRB approval (which was sought separately for this small survey), whom to contact in the case of doubts or questions, and so on. Having consented to take part, they are asked if they currently or previously work at BT. In order to be eligible for the survey, using initial screen questions, the respondent has to be over 18 and i) either currently working for BT in a contact center or ii) has done so in the past 5 years.

We then ask the location of the call center that the employee works at or previously worked at, using a drop-down list of all of BT's call centres. In addition, we ask a short series of questions on their working environment. The wordings are as follows:

- "*Are you working in an open office space (i.e. where you can easily see the rest of your co-workers) or in a closed office space (i.e. where you are separated from your co-workers by internal walls)?*" (potential answers: open office, closed office, other)
- "*Imagine sitting at your typical workstation. Do you see few or many windows? Use the slider below. Imagine a 10 being a completely glass office and 0 being a room with no windows at all.*" (0 to 10 scale with a slider is presented, with 0 labeled "No windows" and 10 labeled "Full glass office.")
- "*While at work, how much natural light do you have access to from your workstation? Use the slider below. Imagine a 10 being like seating outside while working and 0 being an office with no access to natural light at all.*" (0 to 10 scale with a slider is presented, with 0 labeled "No natural light" and 10 labeled "Like sitting outside.")

In the cases where the respondent reports being a former employee, we alter the wording accordingly such that the questions are in the past tense.

Managerial Interviews. We conducted semi-structured interviews with managers at BT. This included interviews with a total of 16 different managers, which took place, broadly speaking, in three main stages. First, an initial round was held in 2017 prior to the launch of the survey. The aim of these interviews was to better understand the functions of workers at BT,

[38] The survey ran from 2022-04-28 to 2022-06-20. £100 was paid to one randomly drawn participant at the end of the survey period.

understand what the firm's main performance outcomes were, and so on. Second, a round of interviews was held in 2018 – largely with more technical managers in charge of data who helped to understand the nuances of the rich administrative data on worker behaviors and outcomes. Finally, an additional set of interviews was conducted in 2021 and 2022. These interviews focused more on eliciting additional information about the physical work environment. The interviews included a mixture of high-level and lower-level managers from various different departments across the business. This included three broad management areas: people in charge of health and welfare across the firm (for example, the Chief Medical Officer, Health & Well-being Lead, and Culture and Engagement Lead), personnel in charge of data and analytics in contact centers (including a Performance and Insight Manager, Principal Research Scientist, and Reporting and Analysis Manager) as well as managers running call centers (such as the Managing Director Customer Care, HR Director of Customer Care, and Head of Customer Insight). We also held more informal conversations (i.e. interviews without any pre-planned questions) during site visits to the contact centers. Finally, in designing elements of the survey, we held meetings with union representatives, though we did use these to interview them or elicit specific information related to the study.

Appendix H Objective Calculation of Proportion of Wall Surface with Glass Windows

This appendix describes the construction of our window share variable. We follow three main steps described in details below.

Step 1. For each of the 11 call center in our dataset, we first recover the building map from Google Maps. This allows us to identify the number of wall pictures to be collected, along with the relative size of each wall. If a long wall is twice as long as a short wall, we give it a weight that is twice as large in the final wall surface measure. The number of wall pictures per call center varies depending on the building type and whether it has walls in common with other neighboring buildings.

Step 2. We then collect wall photos for each call center using Google Street View. In the few cases where all walls are not fully visible with Google Street View, we can still identify them by symmetry, as the same architectural rules apply to various sides of the same building. There exists a large heterogeneity in building types between call centers, and hence exposure to natural light. While the Swansea call center is located within a tall glass tower building, with a lot of light, the Newcastle call center is located within a warehouse set-up, with almost no windows at all. To compute the share of wall surface with windows, we use the open source image processing software ImageJ (see `https://imagej.nih.gov/ij/`). ImageJ can calculate area and pixel value statistics of user-defined selections and intensity-thresholded objects, like the color of wall surface or windows. We first compute the pixel value of the windows within each wall, then the pixel value of the entire wall. We take the ratio between these two measures to obtain the share of wall surface with windows. Window pixel surface is captured in red. The corresponding values obtained are, respectively, 10% (wall 2, Newcastle), 0% (wall 3, Newcastle), 89% (wall 3, Swansea), and 26% (wall 3, Accrington).

Step 3. As a final step, we compute a weighted average at the call center level based on (i) the number of walls and their respective size and (ii) the share of windows for each wall belonging to the call center. The resulting call-center level measure of "share of wall surface with windows" ranges from 3% (Doncaster) to 59% (Swansea), with a mean of 23% and a standard deviation of 16%.

Appendix I Extra Detail on Effect Size Comparisons

Oswald, Proto and Sgroi (2015): The figure we quote of a 1 SD increase of happiness leading to 29 to 35 percent more productivity (in terms of incentivized additions) is implied by the results reported from Experiment 2. 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.

Bloom et al. (2014): 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.

Appendix J Extra Detail on Visual Weather Index

We determine the latitude and longitude of each call center, and match each center to the closest weather station available in the NOAA Global Surface Summary of the Day database, which is on average 14km away. We construct for each call center location a Visual Weather Index, corresponding to the total number of daily incidences of fog, rain and snow during the working week for which individuals reported their happiness. This in theory has a range of 0 to 15, where 0 would mean a likely bright day with no rain, snow, or fog on any day of the week at all and 15 would mean that all three happened on every single day of the week. This is a measure that is visual in nature, ranging from very bright to very gloomy. For robustness, we also collect weekly mean temperature data from the same source and control for it in some empirical specifications. Figure 4 shows the distribution of this visual weather index.