Happiness at Work: Essays on subjective wellbeing in the workplace and labor market

by

George Ward

B.A., University College London (2011)
M.Sc., University College London (2013)
S.M., Massachusetts Institute of Technology (2021)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Management

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2022

Signature of Author: . . . .
Department of Management
April 29, 2022

Certified by: . . . .
Erin L. Kelly
Sloan Distinguished Professor of Work and Organization Studies
Thesis Supervisor

Accepted by: . . . .
Catherine Tucker
Sloan Distinguished Professor of Management
Professor, Marketing
Faculty Chair, MIT Sloan PhD Program

Happiness at Work: Essays on subjective wellbeing in the workplace and labor market

by

George Ward

Submitted to the Department of Management
on April 29, 2022, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Management

Abstract

This dissertation consists of three essays studying the extent to which subjective wellbeing shapes behaviors and outcomes in the workplace and labor market. The first essay studies an information-provision field experiment conducted on a large online jobs platform. The study collects data on the self-reported affective happiness of millions of workers across the USA, aggregates this to the level of companies, and then shows this information to some (randomly-allocated) job seekers on the platform but not others. I find that job seekers respond behaviorally to the provision of information about the happiness of incumbent workers at different organizations, and in doing so they re-allocate their applications away from low-happiness and towards higher-happiness companies. In the second essay, I build on these findings by conducting a survey experiment that provides people with hypothetical choices between jobs at companies with varying levels of i) wage and ii) employee happiness. The results of this analysis suggest that people value workplace happiness and are, on average, willing to trade off wages in order to work at happier companies. In the final chapter, I investigate the relationship between positive affect and productivity. Studying the universe of call center sales workers at one of the UK's largest employers, this research measures the happiness of workers on a week-to-week basis and links it to detailed administrative data on behavior and performance. Exploiting exogenous variation in employee happiness arising from differential visual exposure to bright or gloomy weather while at work, the results show a causal effect of worker happiness on sales in a field setting. Taken together, the findings of three chapters suggest that employee happiness has the potential to promote organizational performance by raising productivity, reducing turnover, and aiding recruitment.

Thesis Supervisor: Erin L. Kelly
Title: Sloan Distinguished Professor of Work and Organization Studies

Acknowledgments

I am indebted to Erin Kelly for her generous, patient, and sage supervision throughout the PhD. I am also very grateful to committee members Tom Kochan and Paul Osterman, who have always provided sympathetic support and astute suggestions during my time at MIT. More broadly, I thank the rest of the students and faculty at the Institute for Work and Employment Research, including Emilio Castilla, Barbara Dyer, Mahreen Khan, Will Kimball, Sandy Kowalski, Claire McKenna, Bob McKersie, Arrow Minster, Michael Piore, Hye Jin Rho, Christine Riordan, Soohyun Roh, Mary Rowe, MacKenzie Scott, Anna Stansbury, Di Tong, and Nate Wilmers. During my time at MIT, I have also benefited greatly from the kindness, comments, and assistance of a great many others – including Clément Bellet, Alan Benson, Brittany Bond, Andrew Clark, Avi Collis, Jared Curhan, Ed Diener, Dean Eckles, Johannes Eichstaedt, Roberto Fernandez, Simon Friis, Caroline Fry, Valère Fourel, Carolyn Fu, David Holtz, John Horton, Simon Jäger, Micah Kaats, Summer Jackson, Richard Layard, Jackson Lu, Sandra Matz, Jim Mellody, Rob Metcalfe, Jan De Neve, Mike Norton, Jenna Myers, Andrew Oswald, Ethan Poskanzer, Drazen Prelec, Ray Reagans, James Riley, Frank Schilbach, Martin Seligman, Susan Silbey, Aaron Sojourner, Basima Tewfik, Lyle Ungar, Hagay Volvosky, Ashley Whillans, Nate Wilmers, Duanyi Yang, Heather Yang, Ting Zhang, and Ezra Zuckerman. I am grateful to *Indeed* for the exciting opportunity to run the field experiment reported in Chapter 1, and am indebted to a large number of people there for their extremely generous help and support, including Rae Albee, Anna Choi, Andy Chiang, Will Johnson, Taylor Sansom, Katie Stevenson, and, especially, Janeane Tolomeo. I am thankful to *British Telecom* for the chance to design and run the study described in Chapter 3 as well as to *Butterfly AI* for their technical support in implementing the employee survey. Finally, I am extremely grateful to my friends, family, and partner, Edlane, for all of their encouragement and support. Thank you.

Contents

List of Figures

List of Tables

Introduction

This dissertation is composed of three chapters that investigate the role of employee wellbeing in shaping behaviors and outcomes in the workplace and labor market. The essays form part of a broader research agenda that focuses on the *lived experience of workers*, and in doing so attempts to understand i) what leads to the large-scale variation across organizations, even within tightly-defined industries and locations, that exists in aspects of the subjective work experience including employee happiness, and ii) what the downstream consequences of this variation are – both for individuals as well as organizations more broadly.

For over a century, both researchers and practitioners have been interested in the topic of employee wellbeing, broadly understood. In a long-running literature stretching back many decades, one of the most extensively studied aspects of workplace wellbeing has been its impact on subsequent behavior and outcomes – employee and organizational performance, in particular. Indeed, early examples of this type of work can be found at least as far back as the 1920s and 1930s (see, e.g., Fisher and Hanna, 1931; Hersey, 1932a,c; Hoppock, 1937; Houser, 1927; Kornhauser and Sharp, 1932; McMurry, 1932; Super, 1939). One of the reasons for this long-running interest is that the relationship has potentially significant implications for the ways in which firms treat their workers and, more generally, for the role of human capital management in a firm's overall business strategy. Perhaps because of this, the relationship between worker wellbeing and performance has even be referred to by some as the "holy grail" of organizational research (Landy, 1985; Wright and Cropanzano, 2004).

While the idea that employee wellbeing can promote organizational success may be far from new, the past few years have seen a resurgence of interest in the topic.[1] A growing number of companies are now at least claiming to care about the wellbeing of their employees, and may be looking to invest in management and organizational practices designed to foster a happier workforce. For example, in a recent large-scale survey of 1,073 business executives in the USA, around 87% agreed that workplace happiness can provide their firm with a competitive advantage (HBR Analytical Services, 2020).[2] Even higher percentages of these managers say they believed that unhappiness has a harmful effect on productivity and that creating happier places to work would likely make it easier for firms to attract and retain talented workers. Moreover, among the factors that these executives say they see as being the most important to the success of their organization, happiness of the workforce comes second only to the happiness of customers – and above a range of other potential factors such as continuous innovation, investing in the best technologies, and having the right market strategy.

But despite this apparent revival of interest in—and enthusiasm for—employee wellbeing, two things are nevertheless worth noting. The first is that although firms increasingly profess to valuing the wellbeing of their workforce, very few actually do anything to promote it. 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, maintain, or improve the wellbeing of their workforce (HBR Analytical Services, 2020). The second thing to note is not therefore very surprising. Work is still very far from a happy experience for most people (De Neve and Ward, 2017). In fact, research using highly detailed time-use data in both the UK and USA finds that work is, on average, one of the most unhappiness-producing activities that

[1]Major news outlets have begun to focus on the issue, for example, including articles written in Forbes (Malcolm, 2021), the Wall Street Journal (Chen and Smith, 2021), the Financial Times (Hill, 2019), CNN (Achor, 2012), and the Harvard Business Review (Seppälä and Cameron, 2015). In addition, the topic has begun to become the focus of attention for influential management consultancies such as McKinsey (Segel, 2021) and Deloitte (Hampson and Jacob, 2020), while several high-profile companies have even gone so far as to appoint executives focused on employee wellbeing, or so-called "Chief Happiness Officers" (Hill, 2022).

[2]For more details of the survey of managers, which included responses from a wide range of sectors and organization sizes, but was not a randomly drawn sample, see HBR Analytical Services (2020).

people do. Paid work is ranked lower in terms of happiness than almost any other activity that people engage in during their day-to-day lives—the only thing worse is being sick in bed—and, somewhat tellingly, the most unhappy moments of people's days are when they are with their boss or supervisor (see Bryson and MacKerron, 2016; Krueger et al., 2009). Put simply, the way that work is currently organized and managed does not appear, on average, to be at all conducive to employee wellbeing.

So, there is a striking disconnect between what firms *say* and *do* in relation to the wellbeing of their employees. In one sense, the survey answers of US executives appear to suggest that the overwhelming majority of managers subscribe to what one might call the human relations viewpoint (cf. MacGregor, 1960; Maslow, 1943; Vroom, 1964) – the idea that employees are key organizational assets and that employee wellbeing can create value by improving productivity, reducing turnover, and boosting recruitment. Some may even argue that these things are so self-evident as not to be worth wasting time and energy writing a dissertation on. But the actions of the majority of firms suggest, though, that the prevailing view is a more traditional—or zero-sum—one in which employee happiness is seen largely as being the result of workers being either under-worked or over-paid. Which viewpoint is closer to the truth has significant implications for the ways in which firms organize and manage work.

One potential reason for the apparent reticence of firms to look after worker wellbeing is that business leaders may yet to be convinced of its potential organizational benefits. That is to say, the evidence on the relationship between wellbeing and performance may be lacking – or at least incomplete. Indeed, despite early enthusiasm for the topic, it is worth noting that the scientific findings were ambivalent at best. Reviews of the early literature on individual-level job satisfaction and productivity, for example, found decidedly mixed results (see, e.g. Brayfield and Crockett, 1955; Schwab and Cummings, 1970). A meta-analysis by Iaffaldano and Muchinsky (1985) concluded that job satisfaction and performance were, at best, "slightly related to each other" and, ultimately, described the relationship as largely being "an illusory correlation." Even though reviews of later literature on the topic found somewhat stronger evidence for the downstream organizational benefits of job satisfaction (Judge et al., 2001) and happiness (Tenney et al., 2016; Walsh et al., 2018), it is perhaps not surprising, given the history of scholarship on the topic, that employee wellbeing is frequently approached by some, both in industry and academia, with more than a hint of skepticism. Ultimately, for many, the idea that employee wellbeing may promote organizational performance is little more than a management fad or a prime example of management mythology.

In this dissertation, I return to these issues in three essays that provide new evidence on the downstream effects of employee subjective wellbeing (SWB). In doing so, I focus my attention specifically on one particular aspect of employee subjective wellbeing: Happiness. SWB is typically thought of as having two main components – an evaluative one and an affective one (for more detailed treatments, see, e.g., Diener et al., 1999; Kahneman and Krueger, 2006; Kahneman et al., 1997). Evaluative SWB in the workplace refers to how people *think* about their jobs. It is an overall cognitive judgement, which is frequently measured using the highly influential concept of job satisfaction (e.g., Freeman, 1978).[3] Affective SWB, on the other hand, refers to how people *experience* their jobs day-to-day. That is, how they feel at work, the positive and negative emotions that they experience.[4] Happiness is typically thought of as a measure of the latter – that is, a positive measure of affective (or "hedonic") wellbeing. While much of the early research on employee wellbeing focused on evaluative measures of job satisfaction (e.g., Brayfield and Crockett, 1951; Fisher and Hanna, 1931; Lawler and Porter, 1967; Locke, 1969), in this dissertation I build on—and follow in the footsteps of—more recent organizational scholarship that has turned more clearly toward the study of affect in the workplace (see, e.g., Barsade and Gibson, 2007; Brief

[3]Job satisfaction is often also thought of, particularly in the organizational behavior literature, as being a job attitude (Wright, 2006).

[4]A third dimension, more closely aligned with evaluative SWB, is increasingly also recognized and studied. This is eudaimonic well-being refers to sense of purpose (see, e.g., Cassar and Meier, 2018; Chadi et al., 2017; Gartenberg et al., 2019).

and Weiss, 2002; Knight et al., 2018; Walsh et al., 2018).

In Chapter 1, I study the extent to which workplace affective SWB has an impact on a company's attractiveness to potential workers. That is, whereas much of the prior work on the role played by employee happiness in shaping overall organizational performance has focused principally on the potential effects of happiness on i) productivity and ii) retention, in this chapter I study the effects of workplace happiness on a third channel: A firm's ability to compete in the labor market in the first place. I provide evidence from a pre-registered field experiment on a large online jobs platform in the USA, *Indeed*, in which treated job seekers are shown crowd-sourced information about the happiness of incumbent workers at the companies to which they are considering applying. I show that displaying information about happiness has an impact on subsequent job search behavior, with treated job seekers increasing their application selectivity and redirecting applications away from low-happiness companies. These labor supply effects are driven largely by job seekers "screening out" unhappy firms from their job search, a finding that is replicated in subsequent field experiments on the platform in Canada and the United Kingdom. The results suggest that emerging large-scale sources of crowdsourced data, which are increasingly displayed on online platforms like *Indeed*, *Glassdoor*, and *LinkedIn* have the potential to be a source of *power* for workers. The provision of information by new labor market institutions affects behavior. Ultimately, in the case of the information shown in this study, it suggests that employers face incentives to invest in organizational and management practices that are conducive to worker happiness, to the extent that they value a larger applicant pool.

In Chapter 2, I build on this field work by conducting a pre-registered survey experiment on a broadly nationally representative sample of the labor force in the USA. Here I attempt to elicit people's willingness to trade off salary in order to work at organizations with happier workforces (or avoid working at miserable ones). I use a discrete choice task involving two jobs, where I vary i) the happiness level of incumbent workers at the company as well as ii) the wage offered for the job. Survey respondents are first introduced to the idea of a company-level happiness scale (measure out of 100, with an average of 65). In all cases, people are offered two jobs, one of which is always at a company with a happiness level of 65 and offers the same wage as the respondent's existing job. The second job offered varies both the wage (in terms of a percentage higher or lower than the respondent's current salary) and the happiness level (45, 55, 75, or 85) of the company's existing employees. I find that, in order to work at company with a happiness level of 75 (as opposed to 65), people are willing, on average, to forgo around 13% of their current salary (and in order to work at an even happier company, with a score of 85, this increases to around 17%). The findings suggest that workers value happiness in the workplace, with at least two main implications. First, these results suggest that firms maintaining happier workforces are likely to experience lower turnover, on average. Second, in line with the results of Chapter 1, firms with happier workforces are likely to find it easier to compete in the labor market, particularly as increasing digitization of labor market institutions continues to facilitate the flow of information about subjective aspects of organizational life—like worker happiness—from incumbent to prospective workers.

Finally, Chapter 3 returns to the long-studied relationship between SWB and productivity. Affective measures of happiness have been shown to have a causal effect on productivity in stylized laboratory settings (e.g. Erez and Isen, 2002; Isen, 2001; Oswald et al., 2015), and this Chapter sets out to build on this work by using a quasi-experimental research design to estimate causal effects in a real-world field setting. This co-authored work begins by measuring the happiness of the universe of call center sales workers at one of the United Kingdom's largest private employers—British Telecom—over a 6-month period, and linking these reports with highly detailed administrative data on workplace behaviors and objective measures of productivity. Exploiting exogenous variation in employee happiness arising from differential visual exposure to bright or gloomy weather while at work (arising from the interaction between a worker's proximity at work to windows and external weather patterns), the results show a causal effect of worker happiness on sales. These effects are principally driven by employees making more calls per hour, adhering more closely to their workflow schedule, and converting more calls into sales when they are feeling happier.

Across the chapters of this dissertation, I find that affective SWB can shape behavior and outcomes in the workplace and the labor market. Happier workers are more productive, while happier workplaces are likely to find it easier to retain workers as we well as attract them in the first place. Taken together, the findings suggest that improving the wellbeing of the workforce can have a positive effect on organizational performance. But while these findings point in the general direction of there being a so-called "business case" for promoting employee happiness, the extent to which this is true will ultimately depend on a number of factors – including the extent to which it is possible for firms to raise the happiness of their employees, and, importantly, at what cost.

In order to account for the important issue of costs, one approach is to look at the relationship between wellbeing and performance at the firm level. Using very rough proxies for wellbeing, prior work has demonstrated positive correlations between the two. Edmans (2012) shows, for example, that firms appearing on "Best Places to Work" lists outperform their peers in the stock market (see also Chamberlain, 2015; Edmans, 2011; Symitsi et al., 2018). This work is highly indicative but potentially limited, however, by only looking at the very best firms rather than the whole distribution. Alternatively, recent work has used "star ratings" on *Glassdoor* as a very rough proxy for company-level job satisfaction. This has the benefit of using data across a larger range of firms but the drawback of using only a very rough proxy for the concept being studied. This work generally finds positive correlations of the star rating not only with stock market performance but also with other important organizational outcomes like profitability (see, e.g., Green et al., 2019; Melián-González et al., 2015). Future work should look to use more direct measures of affective and evaluative employee SWB in order to assess the firm-level relationship more clearly, as well as look for natural experiments that may help to identify the macro-level relationship in a causal manner. In this dissertation, the focus is on what one might instead call the micro-foundations of the macro relationship. That is, across the three chapters I study three channels, or mechanisms, though which an organization may go from having a happier workforce to better performance — in terms of productivity, retention, and recruitment.[5]

But can firms reasonably do anything anyway? As will be shown in Chapter 1, using data on nearly 30,000 firms in the USA, there is significant variation in employee happiness across companies even within tightly-defined industries and locations – that is, across companies facing the same business environment. Future research is required to fully understand what sorts of management and organizational practices lie behind this variation. In the survey of managers, noted above, managers appear to recognize that their decisions and actions clearly affect their workers' happiness. Only a very small number (around 5%) said they believed that organizational and management practices have no influence on their workers' happiness (HBR Analytical Services, 2020). I discuss this at more length in Chapter 2, where I am also able to look at the issue from the worker's side – and in doing so assess the beliefs of workers about how much of their happiness at work they see as being in the control of firms versus themselves. More generally, however, even though evidence on the effectiveness of well-intentioned add-ons like employee "wellness programs" are decidedly mixed (Jones et al., 2019), there is a growing body of multi-disciplinary work demonstrating the more fundamental point that the ways in which work is managed and organized by firms, and the cultures they create, has a significant impact on employee SWB (Krekel et al., 2019). A recent field experiment involving airline pilots by Gosnell et al. (2020) shows, for example, strong effects on workplace happiness of various management practices – including monitoring, performance information feedback, personal targets, and pro-social incentives. Moreover, Moen et al. (2016) find that a program that increases supervisor support and

[5]This is, of course, a non-exhaustive list of channels. Further research may look to other mechanisms. One line of research, for example, may look to firms' ability to raise capital, particularly as more and more data on employee wellbeing becomes available to investors via crowd-sourced platforms. Indeed, measures of employee well-being may increasingly come to be seen as a key data point in defining firms' environmental, social, and governance (ESG) impact (see Allas and Schaninger (2020) for a more detailed discussion of ESG goals and, in particular, the ways in which consistently measured company-level happiness data—such as those studied in Chapter 1—might help to fill a large and important data gap related to social impact (the "S") of firms).

employee autonomy among IT workers improves employee SWB, while a field experiment by Bloom et al. (2014) shows that working from home improves the emotional experience and job satisfaction of call center workers. Work should continue to experiment in this direction – and now is a good time to do so.

The multidisciplinary field of industrial (or employment) relations was founded amidst the backdrop of a significant transition from a broadly agrarian economy to an industrial one (Osterman et al., 2001; Piore, 2008).[6] One hundred or so years later, the world of work is again in a state of flux, both with the advent of new technologies as well as the disruption to working arrangements brought about by the recent global pandemic. Employment relations has always been a problem-centric field, and this dissertation begins with a problem: Unfortunately, work remains, on average, a very unhappy experience for most people (Bryson and MacKerron, 2016; Krueger et al., 2009). But the future of work could look different, and there may be the possibility to redesign work in ways that are more conducive to worker wellbeing. Ultimately, doing so is likely to require a mixture of actions on the part of government, labor, business, emerging labor market institutions, and others. At present, despite a recent resurgence of interest in and professed enthusiasm for employee wellbeing, the actions of business—who the pluralist view of employment relationships typically conceptualize as the actors most directly interested in promoting 'efficiency'—suggest that many are yet to be convinced that employee wellbeing is a topic that need be of great importance to them. Taken together, however, the three chapters of this dissertation provide a modest sliver of evidence to suggest, at the very least, that workplace happiness is far from being antithetical to organizational success – and that it may even help to promote it.

[6]See, for example, Webb and Webb (1897) and Commons (1909) for classic examples in the UK and USA respectively.

Chapter 1

Workplace Happiness and Job Search Behavior: Evidence From A Field Experiment on an Online Jobs Platform

1.1 Introduction

The role played by employee happiness in shaping workplace outcomes has long been a topic of intense interest to organizational researchers as well as practitioners. In a large body of literature that stretches back many decades (see, e.g., Hersey, 1932a; Houser, 1927; Kornhauser and Sharp, 1932), one of the most extensively studied aspects of workplace happiness has been its effects on firm financial performance. According to this so-called "holy grail" of organizational research (Landy, 1985; Wright and Cropanzano, 2004), happier workplaces will increase an organization's economic performance since, at the micro level, employees are expected to (i) perform better in their jobs and (ii) be less likely to quit.

In this chapter, I study a third possible reason happier workplaces may perform better: the increased ability of companies with a happier workforce to compete in the labor market and attract workers in the first place. I present evidence from a large-scale field experiment that provides job seekers with information about the aggregate happiness of incumbent workers at companies to which they are considering applying. The experiment took place on a major online jobs platform, *Indeed*, where job seekers can already see a range of crowdsourced information about companies—including extensive details of salaries, employee reviews, and star ratings left by workers. I explore the effects of providing additional information to job seekers about company-level happiness, in an experimental way. By assessing labor supply responses to information about the comparative happiness of firms, while keeping information about salaries constant, I am able to provide a test of the extent to which people are attracted to, and motivated to work at, happier firms.

While much of the early literature on employee happiness operationalized the concept using measures of job satisfaction (Wright and Bonett, 2007), a more recent turn has seen a greater focus on hedonic (or affective) measures (Barsade and Gibson, 2007; Brief and Weiss, 2002). Here, I follow this line of work by studying the behavioral effects of showing job seekers aggregated company-level information on the extent to which employees agree with the statement "*I feel happy at work most of the time.*" *Indeed* has amassed a crowdsourced survey dataset of answers to this question from over 5.5 million workers across the USA, making it, to the best of my knowledge, the largest ever source of data on employee happiness. For the 20,000 or so companies that had individual-level survey responses from 20 or more workers during the period of the experiment, a "work happiness score" was calculated. To help motivate the field experiment, I first show that this measure of happiness varies significantly across companies, even within tightly-defined industries (see Figure 1-3). This begs the question of what effects there may be of this variation for firms. While a long line of work has investigated the downstream effect of employee happiness on productivity and retention (see Walsh et al., 2018), the extent to which it may have implications on a firm's ability to compete in the labor market is not yet clear.

I experimentally display this happiness information prominently to job seekers on the company profile pages of the platform, over a 10-month period. My main finding is that job seekers respond behaviorally to additional information on workplace happiness. Displaying happiness scores to job seekers serves to redirect applications away from low happiness companies to happier ones. I document an asymmetry in this effect for low and high worker happiness firms, whereby information revealing low happiness discourages applications more strongly than equivalent information revealing high happiness encourages them. This redirection effect is thus driven largely by jobseekers of the platform screening out low happiness organizations from their job search. For companies with scores below 60 on the 100-point happiness score scale ($mean = 63.8; SD = 8.8$), I find an estimated treatment effect of around -2.75%. For scores between 60 and 80, there is little discernible effect on applications, and for companies with scores over 80, I find an estimated treatment effect of 2%. In addition to studying effects on application behavior, I am also able to track actual job hires for a subset of the experimental sample. Although additional information on workplace happiness seems to make job seekers more cautious in their applications by screening out low happiness companies, I show that this does not affect the probability that they ultimately get a job.

I distinguish between *information-provision effects* and *score-value effects*. The field experiment allows me to first estimate information-provision effects on labor supply by varying which job seekers see information about a company's happiness. Here, I study what List (2007) refers to as a natural field experiment—where subjects behave naturally in the environment in which they are being studied, without knowing they are taking part in an experiment. This has the benefit of combining the advantages of randomization that come from laboratory or framed field experiments with the realism that comes from studying observational field data.[1] One downside to this approach, however, is that although I am able to vary the provision of information, it is not feasible (without deception) to induce any experimental variation in the value of the score itself—which would allow me to estimate the effect of changes in the score on application behavior.[2]

To isolate score-value effects on labor supply, I turn to observational data from treated

[1]In my setting, real job seekers view jobs at real companies in the US labor market and are provided with truthful information about the happiness of those companies. Moreover, the experiment involves a behavioral and potentially highly consequential outcome: the decision whether or not to apply to a company.

[2]In an ideal laboratory experiment, one would likely vary access to information about the happiness of incumbent workers, but also randomly vary the score itself such that treated and control job seekers are looking at the same company but with different levels of happiness.

job seekers and rely on a variety of quasi-experimental empirical strategies. I make use of the fact that (i) the experiment takes place over a 10-month period and (ii) treated job seekers view multiple companies with different scores, to estimate application equations that include company, time, and job seeker fixed effects. First, in order to identify the effect of the score on applications, I leverage this within-company and within-jobseeker variation in the scores displayed. Going from being defined on the platform as having "Low" happiness (scores of 49 and below) to "Average" (scores of 60 to 69), the probability of a viewing job seeker applying goes up by 13.93%. Again, however, I find an asymmetry in this association, with there being little discernible relationship between the score and the probability of applying above scores of 65. Second, I build on this evidence by exploiting discrete jumps in the score caused by rounding rules (cf. Luca, 2016; Sockin and Sojourner, 2020), and, in doing so, present evidence to suggest that this is a causal relationship.

The remainder of the chapter is structured as follows. Section 1.2 surveys some of the existing literature and discusses theoretical links between information about workplace happiness and job seeker behavior. In Section 1.3, I introduce the institutional setting by describing the survey instrument used to collect data on individual-level happiness and by explaining how company happiness scores are calculated. In Section 1.4, I report results from the natural field experiment that randomly allocates happiness information to job seekers, before focusing in Section 1.5 on the effect of the score itself on applications among treated job seekers. Section 3.5 discusses managerial implications, a number of key limitations, and some possible extensions. Section 3.6 concludes.

1.2 Background and Theory

1.2.1 Subjective wellbeing in the workplace

The philosophical study of human happiness is a long-running concern, but the empirical investigation of what is now more typically termed subjective well-being (SWB) has increased dramatically over the past three decades (Clark, 2018; Diener et al., 2017). SWB is often referred to loosely and collectively as happiness, but is actually typically thought of as having two separate components that measure (i) how people think about the state of their lives and jobs and (ii) how people feel moment to moment (Krueger and Stone, 2014). The former is usually thought of as evaluative SWB and is measured via concepts such as life or job satisfaction, while the latter is often referred to as hedonic or affective SWB. A third dimension is increasingly also recognized and refers to eudaimonic well-being or purpose (see, e.g., Cassar and Meier, 2018; Chadi et al., 2017; Gartenberg et al., 2019).

SWB varies significantly across people and across organizations, even within industries and locations – that is, across otherwise observationally similar firms facing the same business environment. On the one hand, a large literature has sought to understand the underlying causes of this variation, for example by studying the effects of different management practices as well as various aspects of organizational cultures (e.g., Bloom et al., 2014; Breza et al., 2018; Card et al., 2012; Clark, 2010; Gosnell et al., 2020; Jencks et al., 1988; Moen et al., 2016). On the other hand, work has sought to understand the downstream effects of this variation. At the level of the firm, happier workplaces have been shown to have higher financial performance (e.g. Edmans, 2011, 2012). Trying to understand the micro-foundations of this relationship, a large stream of research has focused on the so-called "happy-productive worker thesis" (Tenney et al., 2016), both in terms of job satisfaction (e.g., Judge et al., 2001; **?**) as well as affect (e.g., Amabile et al., 2005; Estrada et al., 1997; Oswald et al., 2015; Rothbard and Wilk, 2011; Staw and Barsade, 1993). A smaller body of research has also employed panel data to show a predictive link between happiness and subsequent employee quits (Clark, 2001; Green, 2010; Levy-Garboua et al., 2007). However, although the link between workplace happiness and recruitment has occasionally be discussed in theoretical terms, the extent to which happier workplaces are better able to compete in the labor market and recruit talent is an open empirical question.

1.2.2 Information problems in the labor market

In canonical economic models of job search, job seekers typically receive job offers sequentially—with a wage offer from a known distribution (e.g., McCall, 1970; Mortensen, 1970). The job seeker makes the decision in each case whether to accept the offer or to keep looking. In the most simple versions of such models, workers have a reservation wage and accept the first offer that satisfies that wage level. This basic model has been developed and enriched in a number of different directions, including the incorporation of factors such as matching, time horizons, and risk preferences (e.g., Lippman and McCall, 1976; Mortensen and Pissarides, 1994). A recent behavioral literature has also sought to build on it and explain a number of anomalies in job seeker behavior that are difficult to account for using the standard model. Examples of this line of work include the study of, among other things, the roles of present bias (DellaVigna and Paserman, 2005), reference dependence (DellaVigna et al., 2017), and locus of control (McGee, 2015) in explaining search effort and reservation wages.

Like almost all markets, the labor market is one in which participants face (often large) information problems (e.g. Carmichael, 1984; Jäger et al., 2021; Sockin and Sojourner, 2020). Although a large body of work has studied the employer's information problem, such as the extent to which they can observe the ability of potential workers and deal with that problem by using performance-contingent contracts, referrals, and so on (see, e.g., Lazear and Shaw, 2007), a great deal less is generally known about the job seeker's information problem. A recent series of papers has begun to address this by experimentally providing job seekers with information about the labor market in general and aspects of the search process.[3] However, much less attention has been afforded to the job seekers information problem in terms of knowing what working is like at different companies.

[3]For example, Altmann et al. (2018) run a field experiment that gives unemployed job seekers a brochure including information on job search strategies (see also Belot et al., 2019). Gee (2019) reports on a natural field experiment on *LinkedIn* that informs jobseekers how many applications a job already has, and Bhole et al. (2021) show using a field experiment on *Jobs on Facebook* that providing information about competition for a vacancy on the platform serves to redirect applications to jobs with fewer prior applications. Coffman et al. (2017) also show in a field experiment that information on job take-up rates from prior years of the *Teach For America* program can have large effects on which jobs people choose.

This strand of research has been productive in better understanding job seeker behavior, but so far has focused on providing information about the labor market and job search process in general, rather than information about prospective companies. Moreover, the focus in this strand is typically the wage level of different jobs and the extent to which job seekers know and understand the wage offer distribution—and, ultimately, behave in ways that maximize expected utility proxied by wage.

Carmichael (1984) points toward employer reputation as a means through which job seekers may be able to deal with this information problem, at least partially. However, information on reputation was, for a long time, mostly limited to companies at the tails of the reputation distribution. Brown and Matsa (2016) show, for example, significant hiring effects for firms that are known publicly to be in severe financial distress. But while information on a wide cross-section of firms was long unavailable, recent digitization of labor market institutions means that detailed information on a broad distribution of organizations is now more freely and easily available to job seekers.

1.2.3 Digitization of labor market institutions

Growing digitization of the labor market has the potential to significantly alter the ways in which people look for and choose between employers and jobs. Platforms such as *Indeed*, *Glassdoor*, and *LinkedIn* have rapidly become the dominant way that people look for work, at least in high-income countries such as the USA (see, e.g., Horton and Tambe, 2015). A nascent body of work has begun to use data from digital platforms to understand a range of issues, such as how job seekers behave over the course of a period of unemployment (Faberman and Kudlyak, 2019; Marinescu, 2017) or respond to unemployment insurance (Baker and Fradkin, 2017; Marinescu and Skandalis, 2021). But in addition to hosting job adverts and helping to facilitate applications, such platforms also serve another—often overlooked—key function: to crowd-source and display information about companies.

In a small but growing literature beginning to try to better understand the informationaggregation function of digital labor market platforms, Sockin and Sojourner (2020) use data from *Glassdoor* to better understand the reasons why people give reviews of companies and, in particular, why they might conceal parts of their identity in situations where retaliation is more likely (see also Sockin et al., 2021). In addition, the authors exploit rounding rules in the overall star rating of companies on the platform, to show effects on labor supply, at least for small companies. In addition to this, Benson et al. (2020) also use a field experiment on *Amazon Mechanical Turk* and find effects of employer reputation on labor supply in the online market for gig work – and in doing so also develop of formal model to demonstrate the important point that online employer reputation systems have the potential to discipline firms who mistreat workers in different ways.

A larger applicant pool is usually a positive outcome for an organization given that it typically allows firms to choose from a larger pool – and in doing so be able to pick more talented and productive workers, as well as workers with a broader array of diversity along various demographic characteristics. To the extent that attracting a larger application pool is beneficial to the firm, a causal relationship between firm reputation and the ability to attract applications means that there is an incentive to ensure a good reputation as an employer.

1.2.4 Social Information and Economic Behavior

An established stream of research has shown that word-of-mouth and referrals within social networks can affect job search behavior (e.g. Dustmann et al., 2016; Fernandez et al., 2000). However, the extent to which this translates to aggregate information about the experiences of strangers is far from clear. Indeed, even if job seekers are motivated to pursue happiness in their search, they may see crowdsourced information on digital platforms as only a very weak signal of their own expected happiness at a particular company. This is particularly the case to the extent that happiness in a job is the result of better matching between workers and employers, rather than a general quality of the employment relationship for all employees at a company.[4] Moreover, platforms such as *Indeed* already show large amounts of information about salaries, reviews, star ratings, and so on. Job seekers may thus see little *additional* signal to happiness information.

A series of studies has shown the impact of social information on economic behavior in general. Much of this work has probed relatively low-stakes decisions such as contributing to a movie ratings website (Chen et al., 2010), the market for wedding services (Tucker and Zhang, 2011), and donation behavior (Frey and Meier, 2004; Gee and Schreck, 2018). A rapidly growing body of work is studying the effects of crowdsourced ratings on digital platforms (Dellarocas, 2003). While much of this research has up to now been focused almost entirely on product markets, using data from platforms such as *Yelp*, *TripAdvisor*, and *eBay* (for prominent examples, see, e.g., Chevalier and Mayzlin, 2006; Helmers et al., 2019; Luca, 2016; Reimers and Waldfogel, 2021), the extent to which this translates to the labor market is much less clear. Jobs are, after all, different to consumer goods. Decisions made in the labor market of where to work tend to have larger and much longer-lasting impacts than decisions about which products to buy or which restaurant to eat in—both because people spend a large percentage of their waking hours working at the company they choose and because the commitment is a longer-lasting one.

1.3 Empirical Setting

Indeed is a large online jobs platform, with over 250 million unique visitors worldwide each month.[5] The website hosts job adverts, which job seekers can search and browse. In addition, each company that is listed on the platform also has its own set of "company pages." Here, the platform displays a wide range of information about each company, much of which is crowd-

[4] If happiness were a general quality, then any representative worker's experience would give useful information since everyone's criteria for assessing the company would be the same. However, if everyone has different preferences for different types of organizational and managerial practices, then I would be much more likely to find the recommendation of someone in my social network to be information given that she is more likely to know my preferences (or have similar preferences to me, which is typically true within a social network).

[5] This figure is based on unique visitors calculated by Google Analytics, February 2020.

sourced from other jobseekers of the website. This includes job seeker-written text reviews, a headline star rating based on the question "Overall, how would you rate this company?", questions about the interview process at the company, a large amount of salary information collected from current and former employees, and a more flexible question-and-answer section. In this chapter, I study the effects of adding information on workplace happiness to this large existing corpus of information displayed to job seekers about companies.

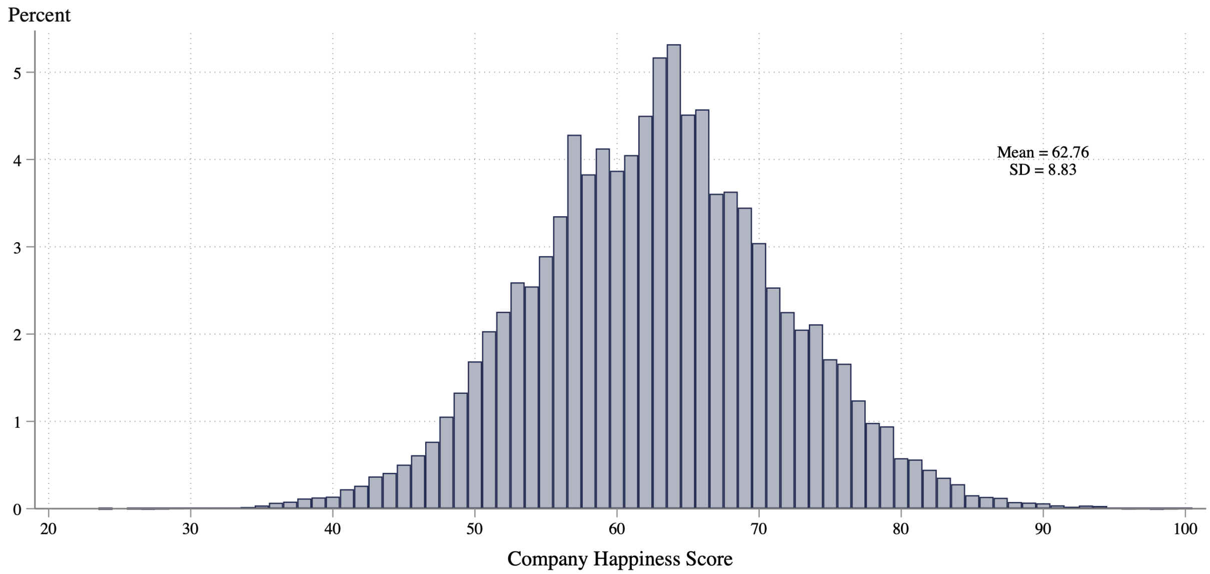
1.3.1 Crowdsourced Happiness Survey

Beginning in October 2019, visitors to the website were invited as part of the platform's general data collection efforts to answer survey questions on workplace happiness at the firms they were currently working at (or had previously been employed by). The data collection process is ongoing, and in this chapter, I use data up to March 17, 2021, giving me a total of 5,338,631 individual survey responses in the USA.

The survey question was based on standard academic definitions of happiness. The headline question of the survey asks respondents the extent to which they agree, on a 1 to 5 ("strongly disagree" to "strongly agree") scale, with the statement "*I feel happy at work most of the time*." This is a hedonic well-being question in that it asks about feelings of happiness. However, it also a global judgment in that it asks not about a specific time frame but about workers' experience at the company as a whole. The validity and reliability of measures such as this have been the subject of decades of academic research (for a full discussion of the issues surrounding the measurement of SWB, as well as a detailed overview of the ways in which the validity and reliability of such measures have been tested, see Krueger and Schkade (2008); Krueger and Stone (2014)).

Respondents were assured their responses would be anonymous and told that their honest responses would help other job seekers. A further 12 questions follow this main question and ask the respondent about different sub-dimensions of workplace well-being. The exact question wording is provided in an appendix, but the sub-dimensions include: achievement,

Figure 1-1: Company-Level Work Happiness Scores



Note: Histogram shows the distribution of the work happiness scores that are displayed to job seekers during the experiment. Bin width=1.

appreciation, belonging, energy, flexibility, inclusivity, learning, management quality, fair pay, purpose, support, and trust (see A.1).[6]

1.3.2 Company-Level Happiness Scores

Once a company has 20 or more completed surveys, it is eligible to have a workplace happiness score shown on its company pages. Company-level happiness scores are calculated to be on a scale out of 100. In practice, this entails taking the mean of happiness responses on the 1–5 agreement scale and multiplying by 20—giving a score that is shown to job seekers as an integer between 20 and 100. Between October 2019 and April 2020, individual-level happiness surveys were conducted, but no company-level happiness scores were yet shown on the website. At the end of April, the company-level scores began to be experimentally

[6] More recent data collection has included questions on job satisfaction and stress. This gives a fuller picture of SWB that is in line with academic understanding of the concept, which, as noted above, includes evaluative, hedonic, and eudaimonic dimensions (Diener et al., 1999). Data on stress and satisfaction were not shown to any job seekers during the course of the experimental period I study in this chapter.

shown to job seekers on the site.

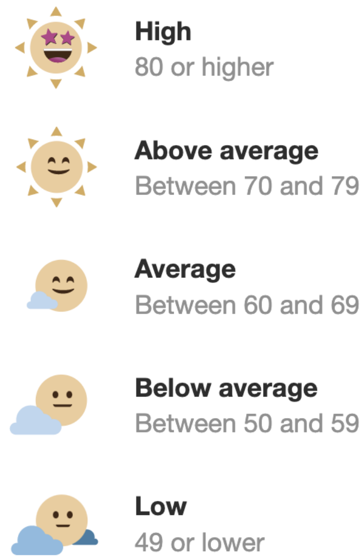
Figure 1-1 plots the distribution of company-level happiness scores. Scores are centered around a score of 63. Very few companies have a score below 40 or above 90, and only a small number have a score below 50 or above 80. In Figure 1-3 I show the distribution of scores within 2-digit industry codes, and find that there is substantial variation across organizations. Further, in Figure A-1 I show substantial variation in company-level happiness, even after partialling out fixed effects for 4-digit industry codes, suggesting that organizations vary in the happiness of their workforce, even with tightly-defined industries.

Clicking on a company's page on the website, the job seeker arrives at the company's landing page. There are a number of tabs, which can take the job seeker to other areas of the company's pages (but always with a banner of tabs at the top such that the job seeker can navigate back to any of the company's pages). The main landing page includes basic information about the company, such as industry, company size, year founded, and CEO name. Other pages the job seeker can navigate to include reviews written by job seekers, crowdsourced salary information, and job listings.

The happiness score is shown prominently as the first item on the company's main landing page (for more details, see Section 1.4.1). Given that the score does not have any natural units, whenever the score is posted on the website, it is accompanied by an emoji and a piece of text explaining that the score is high, above average, average, below average, or low (see Figure 1-2). This information is based on cutoffs that were fixed throughout the course of the experiment and were the same for all companies.

Although happiness is the headline piece of information, a further 12 questions were also asked on sub-dimensions of workplace well-being. The happiness score and its emoji are always shown first and most prominently. Two sub-dimensions are shown alongside the happiness score. These are the two sub-dimensions with the highest two scores for that company. The job seekers can choose to expand to a "full report" to see the scores of the remaining sub-dimensions if they wish to.

Figure 1-2: Emojis Displayed to Job Seekers Alongside The Happiness Score



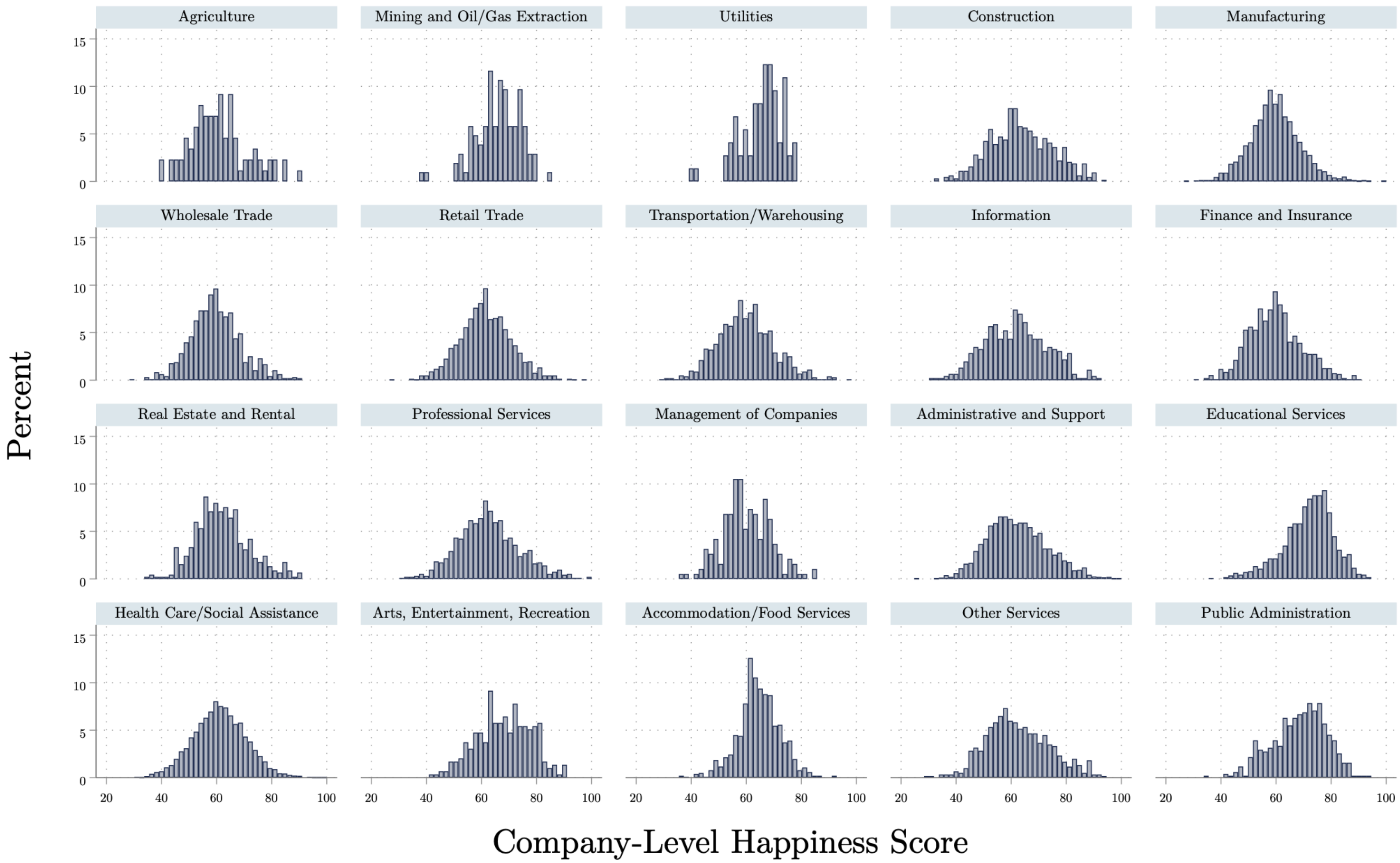
Note: Screenshot of the accompanying emojis and distributional "signposting" statements. When the happiness scores are displayed to job seekers on the site, they are accompanied by an emoji and a statement to help people situate the score they are seeing within the overall distribution.

1.4 Field Experiment: Information-Provision Effects

1.4.1 Experimental Design

Job seekers are randomized on the basis of their internet cookie to be in either the treatment or the control group. Treated job seekers will see the happiness score for all eligible companies (i.e., companies that have at least 20 responses to the happiness survey, so that a happiness score can be calculated) if they navigate to that company's page. Control job seekers visiting the same company's pages will see those pages as normal, but without the happiness information. Figure 1-4 shows what the two experimental conditions look like, for a treated and control jobseeker who visit the page of an eligible company (see Figure A-2 for what the treatment looks like on a tablet computer or smartphone). This is a somewhat conservative test of the the importance of happiness to job seekers. First, I am estimating the experimental effect of showing happiness information, over and above all of the information about workers' subjective experiences at the firm that are already contained in star

Figure 1-3: Company-Level Happiness by Industry in the USA



Note: Plotted are the distributions of the company-level happiness score on the final day of the experiment, within 2-digit NAICS industries.

ratings, reviews, and so on. Second, the experiment is conditional on a job seeker having already navigated to the company's page, and thus does not include any effects of happiness attracting job seekers to view the page in the first place.

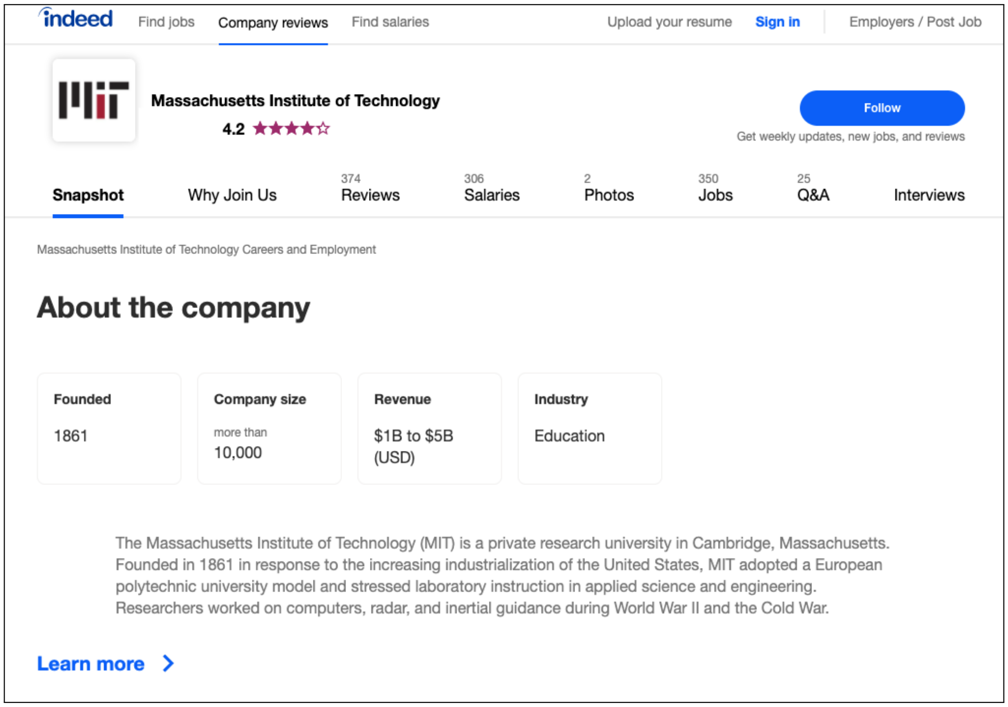
The experiment ran for 10 months beginning in May 2020 and was based on a roll-out design (for pre-registration details, see AEA RCT Registry AEARCTR-0007170). Happiness scores were shown to the majority of job seekers when the score was launched as a product on the platform. A hold-out group comprising a randomly assigned 5% of job seekers, which I refer to as the control group, did not see any happiness scores. I am able to analyze the data at the level of jobseeker-company-day triads. The sample is restricted to jobseeker-company-days on which the company had a happiness score (since treatment and control will look the same if the company does not have the 20 or more happiness surveys required to calculate a score to display). Further, I restrict to jobseeker-company-day where the job seekers navigates to the company's page.[7] For the main analyses of the chapter, I further limit the sample to include only the first observation per jobseeker–company pair. For robustness, I also limit the sample to include only the first jobseeker-company-day for each job seeker, given that there may be path dependence in job search strategy arising from the treatment assignment.[8] The main outcome is whether or not the job seeker applied to the company, from anywhere on the site, during that day. This means that the job seeker clicks on the "Apply Now" button of at least one job listed by the company.[9]

[7]The happiness score appears at the top of the page, so that I assume that a job seeker who navigates to the page sees the score (or would have seen it if they were not in the control group).

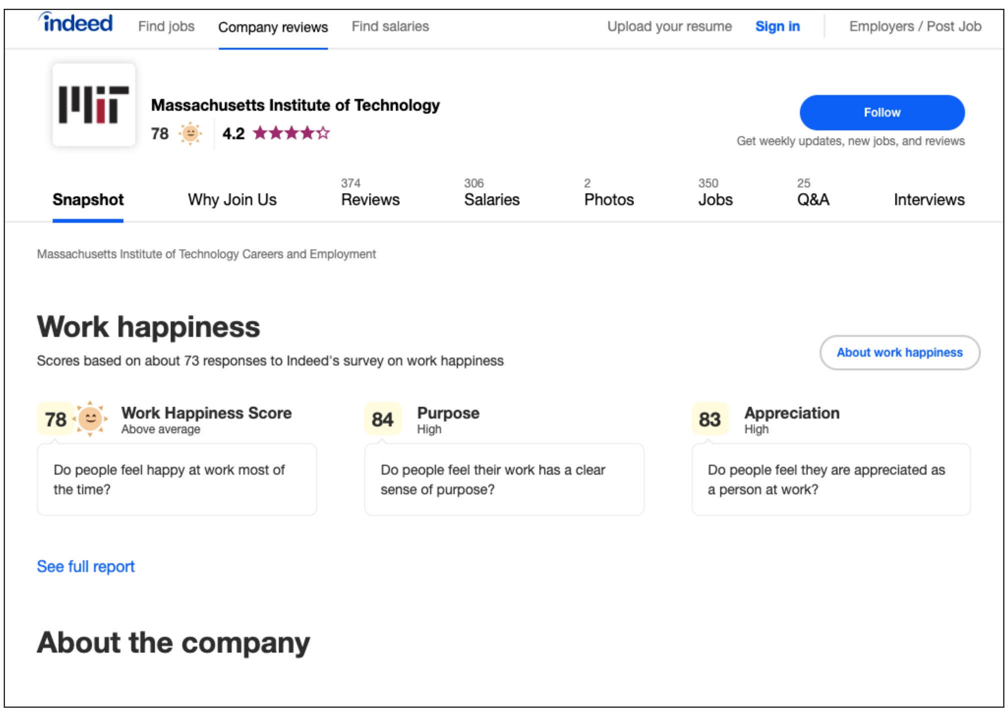
[8]That is, the initial treatment on the first company-day that a job seeker sees with an eligible score could have a knock-on effect on things such as search intensity or which companies and jobs the job seeker subsequently searches for.

[9]Note, however, that clicking on the apply button does not necessarily mean that the job seeker goes through with the whole application. In many cases, the apply button will direct the job seeker to a third-party website, usually the company's own website, to make the application. In these instances, I am not able to track anything beyond this initial click. Nevertheless, "application click" remains a good behavioral proxy for applying for a job, and is itself often used as a behavioral check in survey experiments that ask job seekers to state their hypothetical interest in different types of jobs and companies.

Figure 1-4: Experimental Conditions



(a) Control



(b) Treatment

Note: Screenshots from the desktop version of the platform, showing treatment and control conditions for an eligible organization. Job seekers who are randomized into the treatment condition visiting any company page that has 20+ individual-level happiness surveys see the full Work Happiness Score widget, as depicted in Panel (b). Job seekers in the control condition see all company pages as usual—with no happiness information—as depicted in Panel (a).

1.4.2 Summary and Balance Statistics

Overall, the sample includes 23,376,519 job seekers and 37,373,151 jobseeker-company-day observations. Only relatively basic information about the characteristics of job seekers is collected routinely by the website—I am not, for example, able to observe characteristics such as age, race, or gender. Nevertheless, I am able to observe the age of the job seeker's cookie, whether or not the job seekers is registered with the website, whether or not the job seekers is viewing on a desktop computer, and the job seeker's county (see Table A.1 for summary statistics).[10] Randomization was carried out using the firm's in-house A/B experimental software. The job seekers in the treatment and control groups are similar in terms of these observable characteristics, and the intended proportion of job seekers in both treatment and control is as intended (see Table A.1). There is a marginally statistically significant difference in the proportion of desktop job seekers ($p = 0.058$); however, the magnitude of the difference is extremely small (62.6% versus 62.5%) and a joint test for all covariates is not close to being statistically significant.

At the company level, I am able to observe a number of characteristics, such as whether the company is in the Fortune 500, the number of employees it has, the number of jobs it has listed on the site, as well as other crowdsourced information routinely shown to job seekers, such as reviews and star ratings. Table A.2 shows balance across treatment and control groups in these characteristics. In the experiment, job seekers view the happiness score of companies. This is on average 62.8 and is well balanced across treatment and control conditions.

[10] As part of the website's normal business practices, around 25% of the job seekers I observe also upload their resume to the site, which entails filling out details such as previous education and employment.

1.4.3 Empirical Strategy

I begin by estimating a simple model whereby

$$A_{ijt} = \beta T_i + \varepsilon_{ijt}, \tag{1.1}$$

where A_{ijt} is an indicator variable equal to 1 if the job seeker i clicks to apply for at least one job at company j on calendar day t, and 0 otherwise. T_i is a treatment indicator equal to 1 if the job seeker is in the treatment group, 0 otherwise. ε_{ijt} is an error term that is adjusted for clustering on cookies (the level of randomization).[11]

In the main analyses, I estimate linear probability models (LPMs). I multiply the outcome variable by 100, such that it is equal to either 0 or 100. This does not change anything other than making the coefficients easier to interpret as percentage point changes (essentially, multiplying them by 100). The use of LPMs allows for more readily interpretable estimates, particularly once interactions are included in the equation. They also more easily allow for the inclusion of high-dimensional fixed effects. In an appendix, however, I also show that results are similar when estimating logistic regressions (see Table A.3).

One concern is that treatment may interact with various characteristics of job seekers and companies—most pertinently, the happiness score that is displayed on the page. To this simple equation, I add date fixed effects as well as a series of observable job seeker and company characteristics. In a more restrictive specification, I omit the observable company characteristics and date fixed effects and instead introduce a company-by-day fixed effect, τ_{jt}, such that I am comparing job seekers seeing the same company on the same day. It is not possible to introduce a jobseeker fixed effect, since job seekers are always in the treatment or control group, but I do introduce into the equation a vector X_i' that controls for the job seeker's commuting zone, the age of their cookie, and whether or not they are a desktop

[11] I also experiment with clustering on companies, company-days, and two-way clustering on job seekers and companies as well as job seekers and company-days. Results are robust to all of these specifications and generally have slightly smaller standard errors than those reported in the main analysis.

jobseeker and a registered *Indeed* job seeker.

The sample consists of observations that correspond to jobseeker-company-days. For the main analyses, I limit the sample such that each jobseeker has only one observation per company, which is the first day they visit that company's page.[12] Equation (1.1) provides a causal estimate of showing the score, since T_i is randomly allocated by design across job seekers. However, the perhaps more interesting and important question is the extent to which this effect may vary according to the score itself. Here, I follow two broadly different strategies. One is to split the sample according to five bins of the happiness score that correspond to the emojis shown (as in Figure 1-2). The other is to interact the treatment dummy with a linear term for the happiness score of the company H_{jt}, such that,

$$A_{ijt} = \beta_1 T_i + \beta_2 (T_i \times H_{jt}) + X_i' + \tau_{jt} + \varepsilon_{ijt}. \tag{1.2}$$

1.4.4 Main Experimental Results

Showing happiness scores has, on average, a small negative effect on applications. Column (1) of Table 1.1 reports a pooled treatment effect estimate of -0.273 [95% confidence interval (CI): -0.350, -0.197]. In a fuller model of the pooled effect, reported in column (2), which includes controls for job seeker observables as well as company-by-date fixed effects, I find an estimate of -0.276 [95% CI: -0.351, -0.202]. The mean of the outcome variable, which is equal to 100 if the job seeker applied for the job and 0 otherwise, is 20.05 in the control group.

The average effect masks significant heterogeneity according to the company's score. This is to be expected since the treatment is very different depending on whether positive or negative information is being shown to the job seeker. In Figure A-3, I simply plot the application rate at different happiness levels, for treatment and control job seekers. Within

[12]I do this because the initial treatment may itself determine whether the job seeker returns to the page or not. In robustness tests, I instead limit the sample to include only one observation per job seeker, in which case I include only the first company-day for each job seekers. Alternatively, I extend to the whole sample, such that, if a job seeker returns to that company's page, I track them on multiple days.

Table 1.1: Effect of Showing Happiness Score on Application Behavior

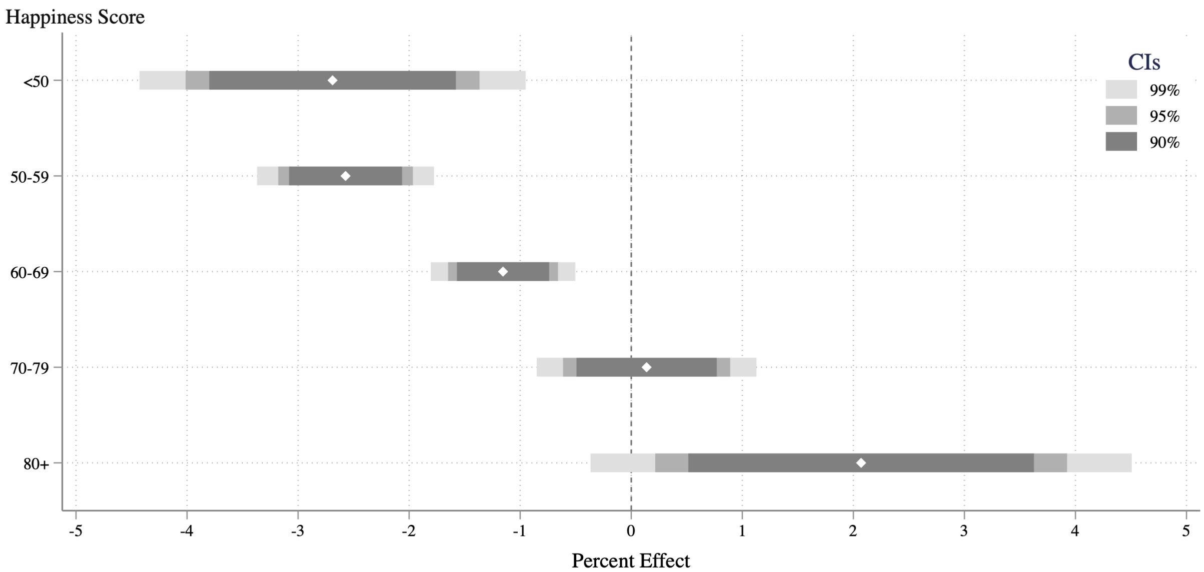
	Applied = 100			
	(1)	(2)	(3)	(4)
Main Effect				
Treated	-0.273***	-0.269***	-0.271***	-0.509***
	(0.039)	(0.037)	(0.037)	(0.125)
Interactions: Treated				
× Happiness (z-score)			0.203***	
			(0.031)	
× score is 50-59				-0.003
				(0.135)
× score is 60-69				0.266**
				(0.132)
× score is 70-79				0.537***
				(0.142)
× score is 80-100				0.890***
				(0.213)
Observations	37,309,899	37,309,899	37,309,899	37,309,899
User Controls		✓	✓	✓
Company-by-Date FEs		✓	✓	✓

Notes: Robust standard errors are in parentheses, adjusted for clustering on individuals. Linear probability models are estimated, in which the outcome is an indicator variable equal to 100 if the job seeker applies to that company. jobseeker controls: logged-in, desktop job seeker, commuting zone fixed effects, cookie age. In column (4), the omitted happiness score category is 20–49. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

the regression framework, I interact the treatment dummy with the happiness score, which is z-scored to have a mean of 0 and standard deviation of 1 across the sample. Here, I find in column (3) of Table 1.1 that the interaction term is oppositely signed and well-defined statistically. This suggests a negative effect that is larger at lower levels of happiness, and a less strong effect at higher levels.

To explore this heterogeneity across scores more fully, in column (4) of Table 1.1, in my preferred specification I allow the treatment effect to vary more flexibly by the happiness score shown. Here, I divide the score into five bins, according to the cutoffs for which emojis are shown. These are below 50, 50–60, 60–70, 70–80, and 80 or higher. Below 40 and above 90, there are very few observations, as can be seen in Figure 1-1, which is why the lowest and highest bins have a wider range. The main effect of treatment suggests that for companies

Figure 1-5: Percent Effects of Showing Happiness Score



Note: The graph illustrates the experimental effects of showing company-level happiness information to job seekers looking at a firm's Indeed page on the probability they will apply to that company. Percent effects are calculated using the control group mean of Pr(Apply) in each model. Results are derived from separate LPMs, with the sample split according to the happiness score of the company in question. Controls are included in all models for jobseeker characteristics as well as fixed effects for commuting zone and company-by-day. Standard errors are adjusted for clustering on job seekers. Full results are shown in Table A.7.

with a score below 50 (the omitted category), the treatment effect is -0.523 [95% CI: -0.773, -0.273]. The main effects of the happiness score bins are not estimated in this regression, since it includes company-by-date fixed effects. However, the first reported interaction term suggests that for companies with a score slightly higher—between 50 and 59—this treatment effect is not discernibly different to when the score is below 50. The interaction effects for the three score bins above 60 are oppositely signed to the main treatment effect and statistically different from zero. Once the score is over 70, the negative effect is canceled out, and once the score is over 80, the treatment effect becomes positive and significantly different from zero (see Figure A-7 for a plot of the implied treatment effect at each level of happiness from this interaction).

To make the results more intuitive and interpretable, I instead split the sample according to these five happiness score bins. In this case, I report the percent effect, using the control group mean of the outcome variable in each regression. Figure 1-5 reports these estimates (for full details of these results, see Table A.7). For companies with a happiness score below 50, showing this “bad news” to jobseekers reduced the likelihood of applying to that company by a little over 3%. For companies with scores between 50 and 59, the effect size is again similar. For average happiness companies, namely those with scores between 60 and 69, there is a negative effect of showing the score, albeit a smaller one of around 0.25 percentage points (column (3) of Table A.7). For scores that are above average, between 70 and 79, there is no effect of treatment. For very high happiness companies, with scores over 80, there is a positive effect of displaying this “good news” to job seekers, of around 2%.

The magnitude of these effects is likely to be conservative, in that I am experimentally estimating the effects of showing happiness information (i) over and above the existing plethora of information contained in star ratings, reviews, and so on and (ii) conditional on jobseekers navigating to the company’s page in the first place. The sharp asymmetry in effect accords with a long line of literature in psychology demonstrating that negative stimuli produce stronger responses than equivalently positive ones (Baumeister et al., 2001). Moreover, the asymmetry is perhaps to be expected, given the way in which jobseekers are likely to use the site. A job seeker who is interested in a job at a particular company, or is interested in the company in general, clicks through to the company’s page on *Indeed*—at which point, they can mostly only be discouraged. In fact, another way to interpret the point at which the information no longer has any impact is the level of the job seeker’s subjective beliefs about the happiness of the company in question. Consistent with recent work by Jäger et al. (2021), I find that workers are not perfectly informed about how they’d fare at other firms (see also Reynolds, 1951, for a classic study of the extent to which workers hold correct beliefs about the wage and non-wage conditions of employment at other firms).

1.4.5 Robustness and Replications

In Table A.4, I show that the effects are little changed by including differing combinations of controls and fixed effects. Instead of introducing company-by-day fixed effects into the equation, in Table A.4, I use a series of company and day fixed effects and find consistent results—regardless of whether controls for jobseeker observables are also included. Given that the outcome variable is binary, a nonlinear model may be more appropriate to the data. I show, however, that the use of LPMs gives similar results to when using a logistic regression to estimate the treatment effects (see Table A.3 and Figure A-4). In the main specification, I include in the sample only the first day that I am able to observe in the experiment for each jobseeker–company pair. I can instead use all observations, such that each jobseeker–company pair can appear twice in the data if the job seeker returns to the company's page on subsequent days. In the other direction, I can restrict the sample such that I include only the first company-day that I observe for each job seeker, to reduce any concerns relating to path dependence of the treatment. In both cases, I find consistent results, as can be seen in Table A.4 for the main effect and Table A.7 for heterogeneous effects by score.

A further concern is that by randomizing across cookies, there is a possibility that I may observe the same jobseeker twice if she visits from another device. This is particularly a concern if the job seeker is then in both control and treatment group. It is worth noting that such a situation would likely bias any estimates toward zero by adding noise; nevertheless, I can restrict the analysis to the subset of jobseekers who are registered. In Table A.9, I add into the main application equation a dummy for if the job seeker is registered and logged in to the platform, and interact it with the main treatment dummy. As before, for ease of interpretation, I split the sample into five, according to the buckets of the score that determine the emoji shown, such that I am essentially looking at a three-way interaction between treatment, being logged in, and the score displayed as part of treatment. I find that at lower levels of happiness (below 60), the magnitude of the negative treatment effect is significantly stronger for registered job seekers.[13]

I also replicate the field experiment in Canada and the UK.[14] The experiments were the same as in the USA, with scores experimentally displayed on eligible company pages. Randomization was carried out across jobseekers (1,542,329 job seekers in the UK and 1,270,398 in Canada). Some 25% of jobseekers were randomized into the treatment group, with the remaining 75% of jobseekers receiving business-as-usual. The studies took place between March and August 2021. Results are reported in Tables A.5 and A.6 and depicted for ease of interpretation in Figure A-6. Similar to the USA, treated jobseekers make fewer job applications when faced with negative information. When the score is below 60, there is around a 3–4% reduction in applications in the UK, and a 4–5% reduction in Canada. In the UK, there is little effect of treatment for scores beyond this. In Canada, scores between 60 and 70 also see a reduction in applications (and no effect beyond this), suggesting potentially higher expectations about the happiness levels of companies in Canada than in the UK.

1.4.6 Other Outcomes

Although applying to the firm is the main outcome of interest, it is also possible to track various other behavioral outcomes. First, I am able to observe whether or not the job seeker clicks to "follow" the company on the platform, and in doing so be more informed in the future about job openings the company have. Second, I look whether or not the job seeker clicks onto the "reviews" tab of the company's page, to test the hypothesis that negative news in particular will lead to further information searching on the part of the job seeker, who has been prompted to be more wary of applying to a firm with a poor reputation for workplace happiness. Third, rather than look at the number of jobs the job seeker applies

[13] However, it is worth noting that the base rate of applications (in the control group) is different: Among control group non-registered job seekers, the mean of the dependent variable is 13.49, compared with 25.05 among control group registered job seekers

[14] Micro-level data collection of happiness surveys began in August 2020, and by August 2021, there were 123,140 completed surveys in the UK and 86,255 in Canada. Given the generally smaller amount of data compared with the USA, companies are eligible to have a happiness score shown when they have responses from 10 employees. In the UK, this meant 1,196 companies with a score to be shown during the experiment, and 792 in Canada.

to, I look at the number of job clicks, i.e. how many jobs the job seeker clicks to view the description of at that company.

Results are reported in the three panels of Figure A-5. At the lowest levels of happiness (below 50), I find a small negative effect on the number of job clicks. There is no effect between 50 and 70, but showing higher levels of happiness leads jobseekers to click on more jobs that the company has to offer. The provision of negative information reduces follows, but this is not a significant effect. At higher levels of happiness, there is a small increase in follows, but again this is not statistically significant. At the highest and lowest levels of happiness, there is no effect of showing the happiness levels to jobseeker on whether they visit the reviews tab; however, at scores between 60 and 80, there is a small negative effect.

A final outcome, of particular interest in this experiment, is job hires. While I am able to observe at a granular level the process of job searching in the applications phase, it is much more difficult to track what happens after that. Nevertheless, the website does track jobseekers through surveys (of both job seekers and employers) to see who actually gets hired. The data available to me here is much smaller and is likely to be a non-random subset of jobseekers who are traceable. Nevertheless, given the very large sample size and the fact that there is little reason to believe that attrition will vary because of treatment status, it is still possible to observe job hires as an outcome in the experiment.

Although the overall effect on applications is negative, there is no discernible difference in job hires across treatment and control groups. As can be seen in Table 1.2, I observe 13,942 hired among the control group of jobseekers in the sample, and 267,992 in the treatment group. This is 1.19% and 1.20% of job seekers, respectively. If anything, the treated jobseekers in the experiment have around a 1% higher number of hires, but a t-test suggests that the difference is not statistically different from zero ($p = 0.25$).

Table 1.2: Effect on Job Hires

	Treatment	Control	Difference
Number of Users	22,275,728	1,170,755	
Tracked Hires	267,934	13,942	
Users with Tracked Hire	1.203%	1.191%	0.012%

Notes: The table shows the number of jobseekers in the experiment who can be successfully tracked by the platform into actual job hires.

1.4.7 Heterogeneity

The experiment randomly varies, across job seekers, the provision of information on workplace happiness. The headline number is always happiness (see Figure 1-4). However, a further two sub-dimension scores are shown by default, with the full set of sub-dimensions available if the job seeker clicks "see full report." In Table A.8, I interact the treatment dummy with the value of the happiness score (as above in the main results) as well as the 12 sub-dimensions. The interaction effect with happiness is oppositely signed and statistically well defined. However, none of the other interactions is statistically different from zero. Overall, this suggests that jobseekers respond principally to the headline score in their application behavior. Further research may look to better understand the extent to which job seekers might use this information in different stages of job search, as well as the extent to which the provision of a range of well-being measures helps workers to match with employers that will maximize their happiness (i.e., obtain jobs at firms with high levels of the workplace characteristics they value most) rather than take jobs with companies that have the highest aggregate levels of happiness.

Guided by theoretical considerations, I also investigate differences in the treatment effect across subgroups of jobseekers and companies. I first look at differences between registered and non-registered job seekers, since we may expect that jobseekers who are more actively looking for a job—as opposed to just browsing—will be more responsive to happiness information. When browsing through jobs less seriously, the potential downsides of working at a low happiness company may not be as salient. In Table A.9, discussed above, I find that the negative effect of showing low-happiness information is stronger among logged-in job seekers. For this subset of registered job seekers, I am able to observe various piece of information about their characteristics.[15], which in Table A.10, I interact with the treatment effect. I first test for any differential effect by cookie age, since in the job search literature that focuses on wages, it is often found that people's reservation wage drops the longer that they are unemployed—as the job seeker's liquidity constraints become more binding (Krueger and Mueller, 2016; Mortensen, 1986). A similar dynamic may be at play here with people's "reservation happiness"—that is, the longer someone has been looking, the more they are willing to compromise on the desire to work at a high happiness company, given that they are more in need of a job. However, I find no evidence of significant differences by cookie age, which is a proxy for length of time looking for a job (Table A.10).

We may expect that someone who is currently employed will be less urgently in need of a job, and so has more ability to "shop around" for a higher happiness company.[16] In this sense, the job seeker has more power—given that they are more able to pick and choose. However, I find no significant differences in the effect for employed versus unemployed job seekers. Job seekers who have higher education and/or more work experience may also have more power in the labor market. Equally, they may have different preferences, particularly if higher education increases expectations about workplace treatment. The direction of the interaction effect for both education (those with a bachelor's degree or more) and work experience (cumulative employed months inputted to the job seeker's resume) is negative at lower levels of displayed happiness, as expected, suggesting that such jobseekers may be more selective. However, the interaction terms are not statistically different from zero (Table A.10). Labor market tightness may also play a role in relation to power: In areas where

[15] These are mostly gleaned from the resume section of the platform, where jobseekers input information about their past work experience and so on. Unfortunately this information is limited, and does not include characteristics such as race and gender.

[16] Although a different setting, Marinescu and Skandalis (2021) show that job seekers decrease the quality of jobs they target around the time that their unemployment benefits will run out, both in terms of wage as well as other characteristics such as educational and experience requirements and contract type. This highlights the link between the extent to which a job seeker is in need of a job and their selectivity.

unemployment is high, job seekers are less likely to have the ability to shop around, given that they know competition is more fierce for jobs. At low levels of happiness, I estimate an interaction term that is positive, suggesting that in high unemployment areas, the negative effect of showing low happiness scores is less pronounced; however, this interaction effect is not discernibly different from zero (Table A.10).

Testing for treatment effect heterogeneity across types of organization, I first look at firm size. A great deal of the theoretical link between happiness information and job seeker behavior outlined above relies on the extent to which this new information helps to solve the worker's information problem. Firm size is likely to be important here, since people tend to already possess information on larger firms. Thus, the hypothesis here is that if jobseekers are already familiar with very large companies, additional information is not new information—and is not likely to sway job seeker behavior, since they already know what they are getting. A similar dynamic is found for restaurants on *Yelp*, where the effects of star ratings are non-significant for chain restaurants (Luca, 2016). This is also in line with the idea that young organizations face the "liability of newness" (Stinchcombe, 1965), in that they have a credibility problem that makes it more difficult for them to trade with other organizations, compared with larger, more established firms (see also Benson et al., 2020). In column (2) of Table A.10, I find that the effect of showing low-happiness information is less negative for firms with a larger number of jobs listed, which may be thought of roughly as a proxy for company size. I also find a positive interaction with a dummy variable equal to 1 if the company has 10,000 or more employees; however, in this case, it is not statistically different from zero. Finally, I am also able to observe the star rating that a company already has on the page, which is displayed to both control and treatment job seekers. This star rating already signals what working at a company may be like. It may be expected that only *additional* information will have a stronger effect on jobseeker behavior. At low levels of happiness (column (2) of Table A.10), I find a positive interaction effect with the company's displayed star rating. In situations where the experiment shows jobseekers a low happiness score, the effect is more negative when the star rating shown alongside is higher, such that the negative happiness information is more surprising. However, I do not find significant differences across star ratings for higher happiness companies.

1.5 Score-Value Effects

Thus far, we have seen that job seekers respond to information about the happiness of prospective workplaces. This suggests that employers have incentives to improve the well-being of their workforce if they want to attract workers as information about workplace well-being continues to become more widely available. There is, however, a subtle difference between the effect of showing jobseekers scores—even across companies with different levels of happiness—and the effect of a company's score *per se*. An ideal laboratory experiment would most likely not only vary the provision of information, but also randomize the signal contained in that information as well—for example, by adding a noise parameter to the score each time it is shown. While this was not possible in the field experiment—for various legal, ethical, and logistical reasons—in this section, I nevertheless attempt to make progress in identifying the effect of the score's value that is displayed to treated job seekers.

1.5.1 Fixed-Effect Estimates

Identification Strategy

Individual-level data collection from jobseekers on their workplace happiness was ongoing throughout the course of the 10-month experiment, and company-level scores change in real time. As a company receives surveys, the score is updated in real time, providing me with within-company variation in the score over the 10-month period I study. In addition to this within-company variation in the score over time, I also benefit from having micro data on job seekers. In particular, I am able to observe jobseekers as they browse various companies, such that I have within-jobseeker variation in the scores they see. Around a quarter of the treated jobseekers (5,441,500) that I observe during the experiment view two or more company pages with a happiness score. On average, these jobseekers view the pages of 3.5 companies, giving a total of 18,666,892 jobseeker-company-day observations that are usable when I rely on jobseeker fixed effects.

Restricting the sample to treated jobseekers who saw more than one happiness score, I estimate equations of the following form:

$$A_{ijt} = \beta H_{jt} + U_i + C_j + T_t + X'_{jt} + \varepsilon_{ijt} \tag{1.3}$$

where A_{ijt} is an indicator variable equal to 100 if jobseeker i applies to a job at company j on calendar day t, or 0 otherwise; H_{jt} is the happiness score of company j on day t; U_i is a jobseeker fixed effect, C_j is a company fixed effect, and T_t is a date fixed effect; ε_{ijt} is an error term that is adjusted for two-way clustering on jobseeker and company.

Whereas in equation (1.1), the assumptions for identifying the causal effect of the main variable of interest—the treatment dummy—were relatively straightforward, given that the treatment was randomly assigned, here, the identifying assumption when estimating the causal effect β is that the happiness score H_{jt} is exogenous to application decisions, conditional on these job seeker, date, and company fixed effects. One major concern is that some time-varying third variable—which is presumably reputational in nature—may affect the company's score as well as applications. Usefully in this context, the website shows not only happiness scores, but also reviews and a star rating of the company. I am able to include in the equation a time-varying control for the company's star rating on the website, which should help to control for any general reputational shocks. I also include controls in the vector X'_{jt} for the number of jobs the company has on the site on that day and the number of happiness surveys that make up the firm's happiness score.

Table 1.3: Effect of a Company's Score on Application Behavior

	Applied = 100				
	(1)	(2)	(3)	(4)	(5)
Linear					
Happiness	0.084***			0.077***	0.077***
	(0.016)			(0.016)	(0.020)
Piecewise Linear					
Spline: below mean		0.144***			
		(0.018)			
Spline: above mean		0.012			
		(0.028)			
Piecewise Linear					
Spline: 20-39			0.390***		
			(0.128)		
Spline: 40-49			0.143**		
			(0.056)		
Spline: 50-59			0.134***		
			(0.021)		
Spline: 60-69			0.070***		
			(0.026)		
Spline: 70-79			-0.061*		
			(0.037)		
Spline: 80-89			0.093		
			(0.084)		
Spline: 90-100			-0.115		
			(0.343)		
Polynomial					
Happiness^2				-0.003***	-0.003***
				(0.001)	(0.001)
Happiness^3					0.000
					(0.000)
Company FEs	✓	✓	✓	✓	✓
Job Seeker FEs	✓	✓	✓	✓	✓
Date FEs	✓	✓	✓	✓	✓
Observations	18,660,412	18,660,412	18,660,412	18,660,412	18,660,412
R2	0.423	0.423	0.423	0.423	0.423

Notes: Robust standard errors are in parentheses, adjusted for clustering on jobseekers and company. Outcome in all models is equal to 100 if the job seeker applied to a job at that company, 0 otherwise. Time-varying company controls are included in all models for number of happiness surveys, company's star rating displayed on company pages, and number of jobs listed by the company. Linear probability models are reported. Happiness score is re-centered around 0 in models (4) and (5). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

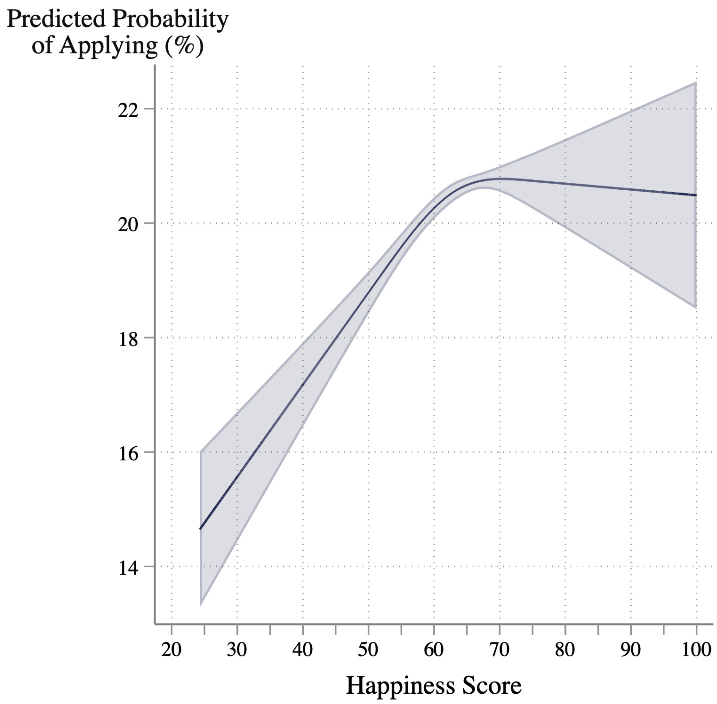
Results

Column (1) of Table 1.3 suggests that a one-unit increase in the happiness score—which lies between 20 and 100—is associated with a 0.084 percentage point increase in the probability of applying for a job [95% CI: 0.053, 0.115], from a base of 20.167. This suggests that a one-point increase in the score has around a 0.41% positive effect on application behavior. A one standard deviation increase in the score increases application probability by 3.65%, on average. However, given the findings of the prior field experiment, there is reason to suspect that any effects the score may have are likely to be nonlinear since jobseekers were more sensitive in their application behavior to being exposed to negative information than they were to positive information about the happiness of prospective companies. In Figure A-9, I introduce into the regression, indicator variables for a series of equally spaced happiness score bins (leaving out the 60–69 indicator, which is the bin that contains the mean value of the score) instead of the linear happiness term H_{jt}. Compared with scores between 60 and 69, lower scores are significantly associated with lower application probabilities. Improving the score beyond the 60–69 bin, however, does not have much of an effect. Although the point estimates for scores of 70 and above are positive, they are not statistically different from zero.

Building on this, I estimate piecewise linear regressions using splines of the happiness score variable. I first split happiness above and below the mean. In column (2) of Table 1.3, I find a significant positive effect of increasing the happiness score up to the mean, with a coefficient of 0.145 [95% CI: 0.107, 0.179], which is around 1.7 times greater than the simple linear coefficient shown in column (1). Below the mean of 63, a one standard deviation increase in the happiness score increases the application probability by around 6.2%. Above the mean, however, I find a positive coefficient of 0.012 [95% CI: −0.042, 0.0675], which is not statistically different from zero.

However, splitting above and below the mean is relatively crude. In column (3) of Table 1.3, I go further in separating the happiness score into seven linear splines, with equally

Figure 1-6: Non-Linear Effects of Happiness on Application Behavior



Note: The figure plots a linear probability model in which the outcome is an indicator variable equal to 100 if the job seeker applies to that company. The happiness score is a restricted cubic spline with 3 knots located at percentiles suggested by Harrell (2015). The regression includes company fixed effects, job seeker fixed effects, date fixed effects, and a set of time-varying company controls. 95% confidence intervals are reported, using standard errors that are adjusted for two-way clustering on companies and job seekers.

spaced cut-points between 40 and 90. The slope is steepest at the lowest levels of happiness and becomes steadily less steep as the score increases. For the four splines where the score is below 70, the piecewise linear effect is in each case positive and statistically different from zero. As the score increases, the magnitude of this effect declines. Over 70, the slope becomes indistinguishable from zero.

In addition to this piecewise linear strategy, I use a restricted cubic spline (see Figure 1-6).[17] This provides similar results to the piecewise linear estimates, with application probabilities increasing as the score goes up to around 65, whereupon the relationship flattens

[17]In this model, I create a restricted cubic spline of the score with three knots, located at percentiles suggested by Harrell (2015), and introduce it into the application equation. In Figure 1-6, I report an alternative model where I specify the knots to be located at 50, 60, 70, and 80.

out. Using this restricted cubic spline model, I find that for observations with scores below 50 ("Low") the average predicted probability of applying is 18.07 whereas this rises to 20.58 for observations with scores between 60 and 69 ("Average"). This represents a 13.9% increase in the probability of applying, going from Low to Average. Going above Average, however, there is little effect on the probability of applying.[18]

1.5.2 Local Randomization Regression Discontinuity Estimates

In this section, I build on the fixed-effects estimates by making use of discrete jumps in the score that lead to a different emoji being displayed to the job seeker. This, with appropriate assumptions, provides an approximation of what we might call a framing field experiment (cf., List, 2007). For an alternative causal inference strategy, with a different set of identifying assumptions, see Appendix A.8, where I instrument the happiness score using plausibly exogenous variation in the source of individual-level happiness scores across companies.

Identification Strategy

The happiness score is shown to the nearest integer on a company's page. Recall that accompanying the score is an emoji and piece of text describing where in the overall distribution the company's score is (Figure 1-2). The emojis and distributional signposts (low, below average, average, above average, and high) are fixed during the whole period and are assigned according to five bins of the score—with integer cutoffs at 50, 60, 70, and 80.

For example, once a company goes from a mean happiness level of 69.4 to 69.5, three things happen. First, the score displayed goes up by one unit, from 69 to 70. Not only does it go up by one integer overall, but so too does the left-hand digit. This is important since there is a well-documented "left digit bias" in consumer behavior, which suggests that people's judgments are disproportionately influenced by the leftmost digit in a number (see,

[18]Introducing into the equation a quadratic of the happiness score also confirms these findings. This regression is reported in column (4) of Table 1.3. One benefit of the quadratic term is that, while being a reasonably good approximation of the functional form established in the spline and categorical analyses, its properties are well understood in two-stage least squares (2SLS) settings using instrumental variables.

e.g., Bhattacharya et al., 2012; Lacetera et al., 2012). Second, the neutral-face emoji that appears next to the score is replaced by one with a happier face. Third, the text next to this emoji changes from "Average" to "Above Average."

I restrict the sample to observations where the happiness score is within a window of 0.2 either side of the four cutoffs (49.5, 59.5, 69.5, and 79.5). On the original 1–5 scale on which happiness is measured, this corresponds to a mean score that is 0.01 either side. That is, if the score were shown on the 1–5 scale and the threshold were 4, this restriction rule would limit the sample to include only instances where the score is in the window between 3.99 and 4.01. I make the identifying assumption that within this tight window, the treatment of having the score rounded up can be thought of as being "as-good-as-randomly assigned." This local randomization regression discontinuity (RD) design implies that within the window, we can think of the latent score as being the same—and see the rounding-up treatment in a similar way to it being an experiment (for a more formal discussion of this local randomization RD approach, see Cattaneo et al. (2015) and Cattaneo et al. (2017); see Sockin and Sojourner (2020) for a recent example of this approach using star ratings on *Glassdoor*).

Results

In Table 1.4, I find a positive effect of rounding up of the score. In the simplest model—reported in column (1)—I take all jobseeker-company-day observations that fall up to 0.2 either side of the four cutoffs, pool them, and regress the application outcome on a dummy equal to 1 if the score is equal to or above the cutoff (and a set of fixed effects for the cutoffs). The coefficient is positive and statistically significant [95% CI: 0.071, 0.332], suggesting that tipping over the threshold makes jobseekers more likely to apply to the company. In the remaining columns, I add into the equation a series of fixed effects and further control variables, which do little to affect the result—suggesting that within this tight window, being either side of the threshold is not likely to be influenced by the characteristics of companies or job seekers. In Figure A-13, I vary the window size around each of the thresholds, up to a

Table 1.4: Discontinuity Estimates of the Effect of Happiness Score on Application Behavior

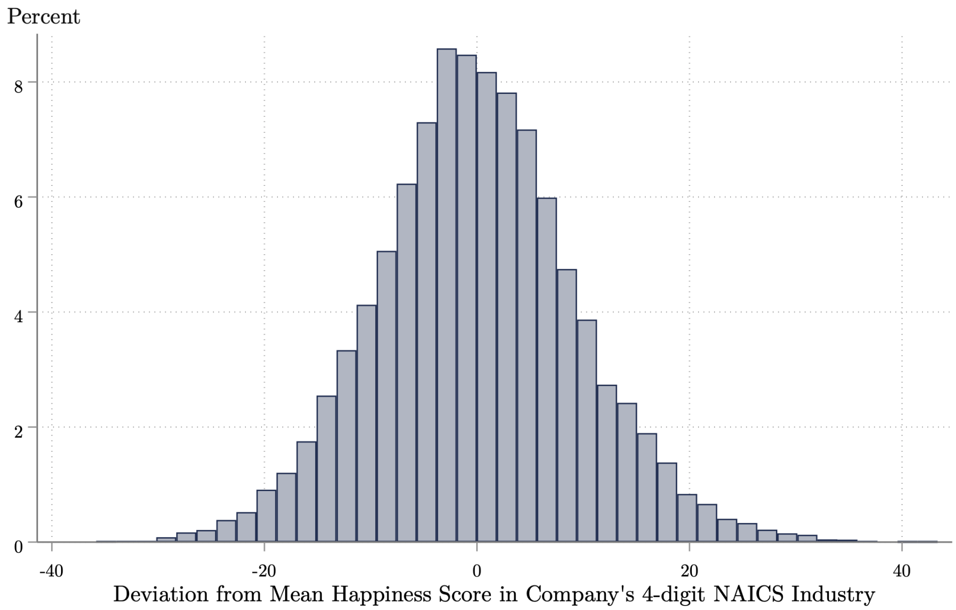
	Applied = 100			
	(1)	(2)	(3)	(4)
Rounded Up	0.202***	0.233***	0.203**	0.211**
	(0.067)	(0.087)	(0.086)	(0.088)
Observations	1,453,747	1,453,747	1,453,747	1,453,747
Cutoff Point FEs	✓	✓	✓	✓
Date FEs		✓	✓	✓
Company FEs		✓	✓	✓
User Observables			✓	✓
Extra Controls				✓

Notes: Robust standard errors are in parentheses, adjusted for clustering on job seekers. Outcome in all models is equal to 100 if the job seeker applied to a job at that company, 0 otherwise. The sample is all jobseeker-company-days within 0.2 points of 49.5, 59.5, 69.5, and 79.5. Rounded up is equal to 1 if the score is above the threshold and thus rounded up to the nearest integer, which triggers the display of a different emoji. jobseeker observables: logged in, desktop job seeker, commuting zone, cookie age. Extra controls: number of happiness surveys, company star rating, number of jobs listed by company. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

maximum of 0.5 either side. This does not significantly alter the pattern of results. Finally, as an additional sensitivity check, I follow the randomization inference approach of Cattaneo et al. (2015) and again find similar results (Table A.12).

The issue of whether or not changes in a company's score over time have causal effects on applications is an important one, since it directly addresses the incentives that companies face to invest in management and organizational practices that are conducive to happier workers. In this section, I have followed three strategies: fixed effects, regression discontinuity, and instrumental variables. Each approach has advantages and disadvantages, with varying degrees of plausibility to their identifying assumptions—and each should be interpreted with caution since there is no explicit randomization. Nevertheless, taken together, the three different approaches converge on the conclusion that improvements to a company's score have a positive effect on job seekers' probability of applying to that company.

Figure 1-7: Company Happiness Score Variation Within Industries



Note: Plotted are the residuals from a regression of company-level happiness score on the final day of the experiment on a set of industry fixed effects.

1.6 Discussion

When surveyed, around 87% of managers in the USA subscribe to the belief that workplace well-being can provide their firm with a competitive advantage (HBR Analytical Services, 2020). One explanation for this is a well-established link between happiness and productivity (see, e.g., Oswald et al., 2015; Walsh et al., 2018). Nevertheless, this is not the only reason firms may routinely measure happiness, make it a strategic priority, and invest in organizational practices conducive to a happier workforce. 94% of managers believe that greater employee happiness would make it easier to attract workers (HBR Analytical Services, 2020). However, although this is seen as a key main mechanism linking workplace happiness to firm performance, empirical evidence is currently lacking. In this chapter, I demonstrate that job seekers respond to information about the happiness of workplaces to which they are considering applying.

1.6.1 Policy and Managerial Implications

My main finding is that workers place a positive value on workplace happiness over and above other key aspects of jobs such as salary and security, which are held constant. This suggests that companies face incentives to look after the happiness of their employees if they want to attract a larger pool of applicants. This is particularly the case as information about the well-being of workers at different companies becomes more widely accessible because of the growing digitization of the labor market.

Taking the average happiness score of the 20,000 or so companies in the study and regressing this on a series of 4-digit industry codes, Figure A-1 plots the residuals from this regression. In doing so, it makes clear that there is significant variation in happiness across companies even within tightly-defined industries – that is to say, companies that are observationally similar and facing the same business environment. This chapter shows that this variation has downstream effects on behavior in the labor market (at least, once it is revealed). This begs a number of follow-up questions for firms and policymakers, such as i) what drives workplace happiness, ii) whether it can be influenced by managerial decisions, and iii) what job seekers are encoding when they see workplace happiness scores.

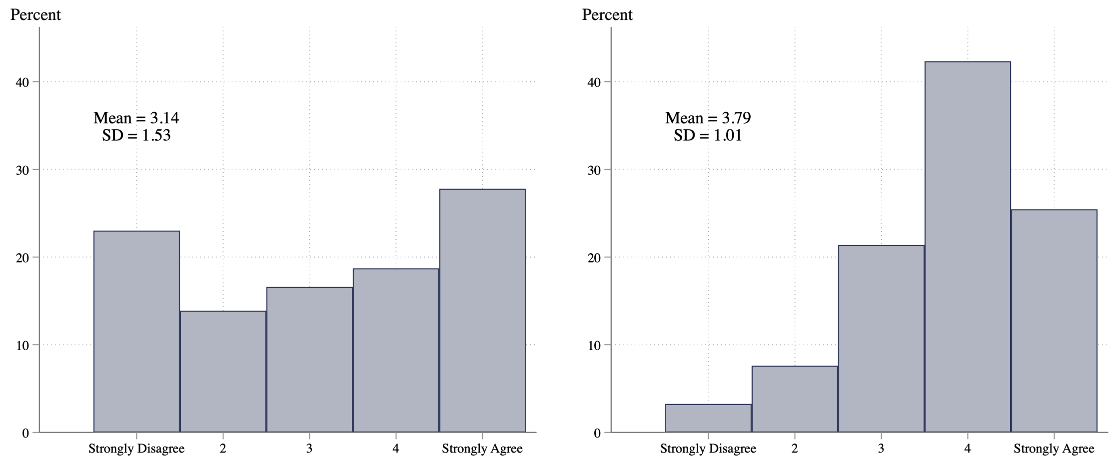
Can firms influence the happiness of their workers?

A growing body of work on happiness in the workplace suggests that employee subjective wellbeing—including measures of job satisfaction, happiness, and purpose—is at least determined by structural factors related to how work is organized as well as cultural factors and relationships within organizations. One strand of this literature regresses measures of subjective wellbeing a range of workplace characteristics, in order to derive relative importance weights (see, e.g., Clark, 2010; Krekel et al., 2019). Results from this approach suggest that pay is an important predictor of workplace wellbeing, but that it is frequently not the variable that explains most of the variance in the happiness of workers.[19] Interpersonal relationships—both horizontally between co-workers and vertically between managers and staff—are highly important, as are things such as flexibility and work arrangements that are conducive to work–life balance.

This initial approach is potentially limited by only being able to make only cross-sectional comparisons between workers. Nevertheless, a more recent line of work is field experimental in nature – with researchers experimentally trying different managerial and organizational practices, and tracking measures of employee happiness. For example, Gosnell et al. (2020) ran a field experiment for airline pilots and found effects on workplace happiness of various management practices, including monitoring, performance information feedback, personal targets, and pro-social incentives. Equally,Breza et al. (2018) show that pay inequality can have large effects on happiness (see also Cullen and Perez-Truglia, 2020), while Bloom et al. (2014) find that working from home improves emotional experience as well as satisfaction of call center workers. Also in relation to autonomy, Moen et al. (2016) find that a program designed to increase supervisor support and employee independence improved various measures of worker subjective wellbeing.

This burgeoning literature on the causal determinants of subjective wellbeing demonstrates at least three things: i) organizational practices can be changed, ii) employee happiness is at least partly determined by these managerial practices, and iii) that there are specific areas that may be targeted based on this research. Nevertheless, more research is needed in order more fully understand the determinants of worker happiness, particularly how they may differ across different industries, occupations, and demographic characteristics of workers such as age, race, and gender.

Figure 1-8: Individual-level responses across survey modes: "*I feel happy at work most of the time*"



(a) Crowdsourced Survey (b) Nationally Representative Survey

Note: Panel (a) shows the distribution of 2,357,522 happiness responses from individuals in the USA who responded to the online survey between October 2019 and March 2021 in relation to a company that had at least 20 responses over the period. Panel (b) shows the distribution of 4,033 happiness responses collected in an online survey in March 2021 from a sample of US individuals that is representative of the US labor force. See Chapter 2 for more details.

1.6.2 Crowdsourced Happiness Data

Subjective wellbeing in the workplace may be measured in a number of ways, such as surveys and, more recently, approaches like gleaning emotional states using natural language processing from social media posts or emails. One innovation of this chapter is to use crowdsourced data. This has the key benefit of being able to provide data measured in a consistent way across a very large number of firms. However, the underlying sample is not random and not likely to be representative. Job seekers select into visiting the website, and then select into answering the happiness survey. The data is thus subject to a number of caveats that are common when working with such crowdsourced data. Indeed, one may expect that respondents selecting into filling in the survey may be likely to be either very happy or very unhappy, either of which may motivate them to respond.

[19]Indeed, even within the context of pay, it is typically found that happiness is more strongly predicted by relative pay rather than absolute pay, suggesting a key need for fairness in pay practices (see Breza et al., 2018; Card et al., 2012).

In this chapter, this is largely a moot point. I am interested in the ways in which job seekers *respond* to aggregated company-level summaries of crowdsourced information about employee happiness. The extent to which the underlying data may be noisy or biased in different ways falls outside the scope of this chapter and should be subject to further research.

Nevertheless, in order to shed light on the issue, I included in a nationally representative survey (to be discussed in greater depth in Chapter 2) the same questions that are asked on the platform. Panel (a) of Figure 1-8 shows the individual-level crowdsourced responses on *Indeed*. The two most common responses are 1 and 5—that is, strongly disagree and strongly agree. This sort of bimodal distribution is consistent with the idea that people answering these surveys are likely to be either very happy or very unhappy with their employer, leading them to answer a survey. In contrast, the representative sample shown in Panel (b) of Figure 1-8 is more normally distributed. While the disproportionately high 1 and 5 responses in the crowdsourced data to some extent cancel each other out when it comes to the overall mean, the very-unhappy-responder bias dominates, such that a given company's score is likely to be lower than its "true" level of worker happiness. However, making comparisons between aggregate measures of company happiness (assuming a relatively high threshold for the number of surveys needed in order to calculate the score), this will only be problematic to the extent that bias varies systematically across companies. Ultimately, additional research is required in order to more fully understand the this new source of large-scale data on workplace data; but the initial evidence suggests that it may provides an exciting "worker's eye view" of thousands of organizations across the USA (and increasingly also other countries around the world) that could allow for the testing of a range of interesting hypotheses across the social and behavioral sciences.

1.6.3 Magnitude of Experimental Effects

During the experiment, jobseekers continued to see employee reviews, companies' overall star ratings based on employee feedback, as well as information about salaries and other question-and-answer content. It is reasonable to expect that the magnitude of the effect of the provision of additional information about workplace happiness would be larger if it were provided by itself versus nothing. The effects estimated in this experiment are the *additional* effects, conditional on all of this prior information. It is also the case that the effects presented here are conditional on a jobseeker navigating to a company's page to see (or not) the score—that is, I study companies presumably already in job seekers' choice sets. This is likely to underestimate the size of the happiness effect if companies' scores also serve to attract job seekers to their pages in the first place.

The strength of the treatment effect is greater when negative information is shown, as compared with equivalently (in terms of distance from the mean score) positive information.[20] Job seekers appear to use the happiness score information to screen out miserable workplaces from their consideration. This accords with a long line of research in psychology that shows that "bad is stronger than good" (Baumeister et al., 2001). An alternative explanation is that workers are generally risk averse, such that they place a stronger value on negative information about companies they are considering working for (cf. Sockin and Sojourner, 2020). In addition, the asymmetric effect also makes sense insofar as jobseekers navigate the company pages of the platform when they are already interested in applying to that company. Having seen a job listed for the company elsewhere on the site and become interested in applying, jobseekers may then go to their page. At this point, they can really only be discouraged from applying.

One puzzle is that jobseekers have a negative (albeit small) reaction to happiness scores that are around the mean. The explanation for this should be subject to further research. One interpretation is that jobseekers are attentive and fully understand the scale, but are

[20]For companies with scores below 60, I find an estimated treatment effect of around −2.75%. Over the course of the experiment—taking all jobseeker-company-day observations—2,591,043 treated jobseekers applied to companies with scores below 60. In the absence of treatment, we would expect 71,254 more applications over the 10-month period. For companies at the other end of the spectrum—with scores over 80 (of which there are far fewer)—job seekers in the treatment group made 210,749 applications. Using the estimated treatment effect of 2%, the provision of the score led to around 4,215 more applications to these high happiness firms.

motivated to avoid "average" happiness companies as this is below what they want. Another interpretation is that a score of 65 is seen as low, particularly in the USA, where scores out of 100 are often intuitively understood on the basis of things such as educational grade scores in which 65 would be relatively low. This may be particularly the case if jobseekers are inattentive—that is, even though there is text next to the score saying that 65 is average, job seekers may not fully take in that piece of information and instead focus on the more pertinent piece of information—the number. Another interpretation may relate to so-called grade inflation. Filippas et al. (2019) suggest that ratings have become inflated over time on many platforms such as *Amazon*, *eBay*, and *Airbnb*, since raters feel pressure to give above-average scores. Given this, it might be that jobseekers have become so accustomed to skewed distributions when it comes to online reputation systems, where scores below 4 out of 5 are now frequently dismissed by jobseekers as very low, that their behavior is affected even in situations where scores are more normally distributed.

1.6.4 Generalizability

How generalizable are these findings? Here, I follow List (2020) in discussing the so-called SANS (selection–attrition–naturalness scaling) conditions. In terms of selection, the field experiment is not a random selection of the labor force, but it is nevertheless the population of online job seekers who navigate to company pages on *Indeed*. Having access to administrative data on jobseeker behavior on the site, attrition is not a major concern. The field experiment is at-scale and is natural in that subjects do not know they are taking part. Altogether, the experiment has high generalizability. Nevertheless, the study does have a number of limitations.

1.6.5 Limitations and Extensions

The experiment was not able to randomly vary the score itself or the salary of the jobs offered. There is a trade-off here between being able to run a large-scale natural field experiment that is highly generalizable and the level of experimental control. To add to the experimental evidence of showing the score, versus not, I exploit discrete jumps in the score owing to rounding rules.

Although I have been able to quantify the extent to which improving workplace happiness raises the size of the applicant pool, which I interpret as meaning that there are incentives for companies to improve their score, two caveats are worth noting. First, one needs to know the marginal cost of improving workplace happiness, on the one hand, as well as the comparative benefits of a larger applicant pool on the other. In both cases, the strength of the incentives faced by firms to improve employee happiness will depend on a number of factors. Second, instead of providing incentives to improve workers' happiness, an alternative interpretation is that firms instead simply face an incentive to improve their score on crowd-sourcing websites, which may well involve manipulation of these scores rather than any changes in employment or management practices. Sockin and Sojourner (2020) discuss the issue of retaliation by employers against employees who give negative reviews of companies on online jobs platforms such as *Glassdoor*. I cannot observe any manipulation (either overtly or through the threat of retaliation) directly in the data. However, it is worth noting that the website screens surveys for data quality and abnormal activity—for example by bots. Further research should investigate these dynamics further to understand more deeply the incentives faced by organizations.

Several further extensions of the experiment would add to the findings. For example, a very useful extension would be gather follow-up data that would allow not only to track the extent to which treated and control jobseekers get hired, but also the extent to which there are potential differences in terms of match quality. This may be in terms of how happy they are with the jobs they end up, or a more revealed preference approach of how long they stay in that job. Further, it would also be useful to display happiness information on job adverts themselves, rather than on company pages. The majority of job seekers on the website search for jobs using the search function, rather than navigating to company pages and deciding whether or not to apply. This would also allow for more fine-grained heterogeneity analysis, across occupations and types of job, rather than at the coarser level of company characteristics. A further extension would investigate more fully the other dimensions of workplace experience in the scores shown on the website. Experimentally manipulating which of these are displayed, as well as their scores, would shed light more fully on which workplace characteristics people value the most and the extent to which the provision of information on fine-grained aspects of work life might enable for better matching between workers and firms, assuming that workers likely have heterogeneous preferences for things such as flexibility, purpose, and appreciation.

Finally, the experiment took place in North America and the UK. The extent to which preferences for workplace happiness are culturally specific is not clear and cannot be estimated in these data. Further research is required to test the replicability of the findings in different contexts. Moreover, while it was not possible in this instance to observe demographic characteristics such as race and gender in the field experiment, further work is required in order to investigate the extent to which demographic and socioeconomic characteristics might moderate the main effects.

1.7 Conclusion

Happiness is seen as a key outcome of interest by a growing number of governments and public policymakers. For example, a number of national statistical offices are routinely collecting data on the SWB of citizens, and various governments are now using such data as a means of evaluating and informing their policy priorities and decisions (Graham et al., 2018; Krueger and Stone, 2014). A key question thus arises as to what sorts of labor market policies and institutions might serve to help improve well-being. The rapidly increasing digitization of the labor market, which has the ability to ease the flow of information to workers, is one potential—but unstudied—avenue through which to discipline opportunistic firms whose managerial and organizational practices typically induce low levels of happiness among their workforce. The evidence presented in this chapter suggests that job seekers respond behaviorally to the provision of information about the happiness of incumbent workers, in a real online labor market. The net effect of this is to redistribute applications away from low- toward high-happiness organizations – suggesting that firms beginning to measure and invest in workplace happiness may well have a business case for doing so.[21]

[21] I am very grateful to *Indeed* for their generous data sharing and collaboration. All views expressed in this chapter are solely mine and should not be attributed to any other entity. For pre-registration details of the field experiment, see AEA RCT Registry AEARCTR-0007359. The study was approved by the MIT Committee on the Use of Humans as Experimental Subjects (Approval Number E-2700).

Chapter 2

Do Workers Value Happiness? Evidence from a Discrete-Choice Survey Experiment

2.1 Introduction

A long-running discussion on human motivation centers around the question: What do people want? A fundamental feature of many theories of motivation is that humans want to be happy (Lawler, 1973; Myers et al., 1993). Testing this hypothesis empirically can prove to be a difficult task, however. In this Chapter, I attempt to shed light on the extent to which people value happiness in the context of the workplace. I use a discrete choice experiment embedded within a large-scale survey of the US labor force, the purpose of which is to elicit people's willingness to trade off salary in order to work at organizations with differing levels of employee happiness.

The past few years have seen a resurgence of interest in workplace wellbeing, and a growing number of companies are at least claiming to care about the happiness of their employees. In a recent large-scale survey of U.S. executives, for example, around 87% agreed that workplace happiness can provide their firm with a competitive advantage (HBR Analytical Services, 2020).[1] Nevertheless, work is still far from a happy experience for most people.

[1]For more details of the survey of managers, which included responses from a wide range of sectors and organization sizes, but was not a randomly drawn sample, see HBR Analytical Services (2020).

Time-use data show that paid work is ranked lower in terms of happiness than almost any other activity that people engage in (the only thing worse is being sick in bed), and that the most unhappy moments of people's day are when they are with their boss or supervisor (Bryson and MacKerron, 2016; Krueger et al., 2009). But despite the apparent widespread professed agreement among managers about the benefits of employee happiness, the large majority remain hesitant to do anything to actively improve it. Indeed, only a third of organizations see employee happiness as a strategic priority, and fewer than 20% actually have any sort of employee wellbeing strategy in place (HBR Analytical Services, 2020).

For the relatively small number of firms who do manage to foster worker happiness, is there a "business case" to do so? This depends, at least partly, on the extent to which workers actually value how happy they feel at work. This is for two main reasons. First, if it is the case that workplace happiness is highly valued, then happier workers will be less likely to quit and look for opportunities elsewhere. Second, to the extent that people looking for work value happiness in the workplace, they are likely to be more willing to apply to and work for companies with higher happiness levels at any given wage level, particularly as more and more information on affective aspects of organizational life are made available through online labor market institutions.

Although it has long been recognized that workers want more out of a job than just a paycheck (see, e.g., Rosen, 1986) and would be willing to trade off wages for a wide range of job amenities such as flexibility, team work, autonomy, and purpose (e.g., Burbano, 2016; Jencks et al., 1988; Maestas et al., 2018; Mas and Pallais, 2017; Stern, 2004), the extent to which they value the more *general* concept of happiness in the workplace is an open question. In fact, happiness might be seen as a frivolous distraction from the things that really matter to people. Indeed, the fact that only 20% of companies report having any sort of strategy to look after or improve employee happiness suggests that this is likely the actual view of the majority of firms.

In this chapter, I offer survey respondents hypothetical choices between jobs, where I am able to vary both the wage offered and the level of happiness of incumbent workers at different companies. This focus on happiness is different from a long tradition of scholarly work on employee wellbeing—across various disciplines in the social and behavioral sciences—that has focused instead on the concept of job satisfaction. Here I move away from the focus on evaluative measures of wellbeing, which are designed identify how people think about and evaluate their jobs overall (or their attitudes towards their jobs), to instead use affective measures of wellbeing that identify how people experience their jobs or how they feel in their day-to-day work lives. This follows a general turn in the literature, particularly in psychology and organizational behavior, that has investigated affect in organizational contexts (Barsade and Gibson, 2007; Brief and Weiss, 2002; Walsh et al., 2018) as well as the affective tone of organizations more generally (Knight et al., 2018).

I find that, on average, survey respondents are willing to trade off wages in order to work companies with higher incumbent happiness. Participants are introduced to a company-level happiness scale out of 100 and told that a typical company has a happiness level between 60 and 70. In all cases, people are offered two jobs. One of the jobs is at a company with a happiness level of 65 and the same wage as their existing job. The other job offered varies both wage (% higher or lower than current salary) and happiness level (45, 55, 75, or 85). In order to work at company with a happiness level of 75 as opposed to 65, people are willing, on average, to forgo around 13% of their current salary. In order to work at an even happier company, with a score of 85, this increases to around 17%.

The results add to the huge body of work in psychology, economics, organizational behavior, sociology, management, and elsewhere on happiness in the workplace (Barsade and Gibson, 2007; Brief and Weiss, 2002; De Neve and Ward, 2017). While a great deal is already known about the effects of happiness on organizational outcomes likes productivity and, to a lesser extent, turnover (e.g. Estrada et al., 1997; Oswald et al., 2015; Walsh et al., 2018), little is known about the extent to which workers themselves value it. Filling this gap in our knowledge is not only important in terms of improving our academic understanding of the concept of workplace happiness, but it is also of practical importance to firms as they consider how to attract and retain a workforce.

Although the question of what workers care about in a job has been studied for many decades, the literature has been re-enlivened in recent years with the creative use of surveys and experiments. For example, Maestas et al. (2018) pose hypothetical job choices to survey respondents while experimentally varying the wage and different bundles of job characteristics (see also Wiswall and Zafar, 2018). Mas and Pallais (2017) embed a survey experiment that elicits preferences for flexibility within a recruitment drive for a call center, and Flory et al. (2015) run a natural field experiment where interested job seekers on an internet jobs board were shown the same job with differing compensation regimes.[2]

In addition to this research on specific job amenities, other studies has also investigated preferences for the characteristics of the organizations as a whole – such as a firm's commitment to corporate social responsibility (CSR) (Carnahan et al., 2017; Hedblom et al., 2019; List and Momeni, 2021) or pro-social work (Ashraf et al., 2020). Burbano (2016) shows, for example, that providing positive information about company-level CSR on job adverts attracts applications from job seekers. When studying the mechanisms behind the link between CSR and application behavior more closely, she finds evidence that people respond to such information largely because they interpret it as a signal that the company is likely to treat workers well, thus raising the expected utility of working there. In line with this, happiness is typically thought of as an *overall* measure that is ultimately a function of various aspects of organizational life including pay, security, flexibility, interpersonal relationships, autonomy, purpose, and so on. In this paper, I thus take a more general approach than the literature studying specific amenities, and instead estimate the effects of providing information more directly on the experienced utility (cf. Kahneman et al., 1997) of incumbent workers at different firms.

[2] In related work, Bandiera et al. (2010) find using a field experiment that workers are motivated to work together with friends (see also Nagaraj and Piezunka, 2020), while Benson et al. (2020) and Breza et al. (2018) show, broadly speaking, that workers have a strong preference for fair treatment.

The remainder of the chapter is structured as follows. Sections 2.2 and 2.3 introduce and describe the survey. Section 2.4 outlines the research design used to elicit people's preferences for happier workplaces, which is then used to derive the results that presented in Section 2.5. Section 2.6 discusses managerial implications, a number of key limitations, and some possible extensions. Section 2.7 concludes.

2.2 Data Collection

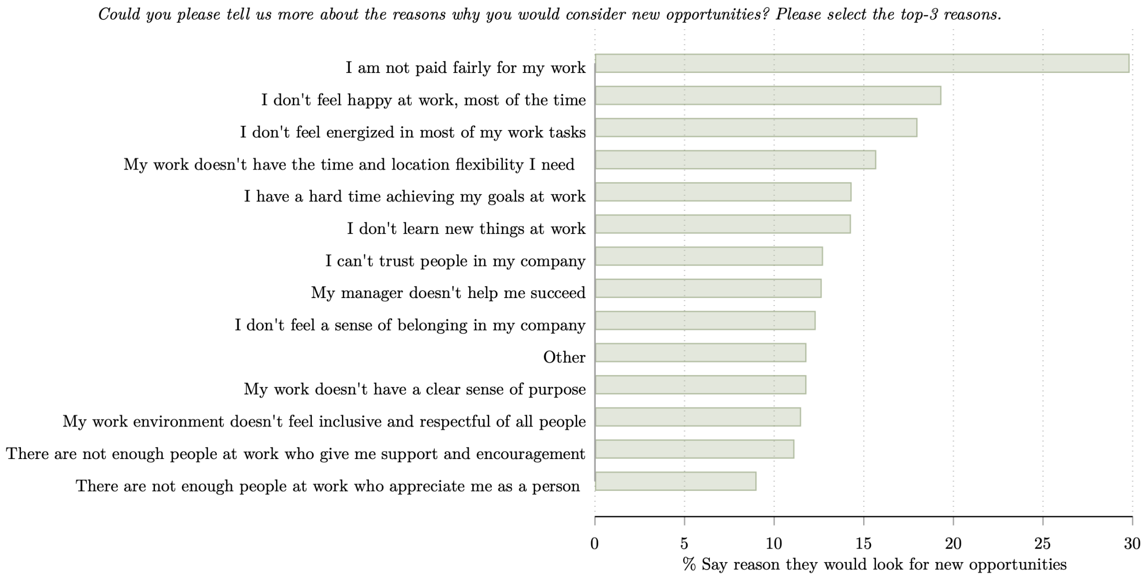
The survey was conducted using an online sample, collected in March 2021. The survey sampled adults aged 18+ who reported i) currently working full-time, ii) currently working part-time, or iii) currently not working but are looking for work. In total, 4,033 US adults were surveyed. Smaller samples were also collected in the United Kingdom (N=1,534) and Canada (N=1,532).

The sample was drawn from an existing online panel of potential survey respondents, using *Forrester*. The incidence rate for the sample, given the inclusion criteria, was around 75% of these panelists. To ensure a reasonably representative sample, quotas were set by age, education, gender, region, and income.[3] While in a "traditional" (i.e. not online panel-based) it is typical to report a response rate. However, in this setting this is less relevant since the panel vendor does not sent out study-specific invitations. Rather, they collect information on demographics and other characteristics, and then distribute the potential respondents to different studies based on those characteristics.

The survey asked the same battery of workplace experience questions as those that are asked on the job search platform, *Indeed*. These questions include the same happiness question that was used in Chapter 1, which asked respondents the extent to which they agree (on a 1-5 scale) with the statement "I feel happy at work most of the time." A further 12 questions followed this main question (as on the online platform) and asked the respon-

[3]These quotas were chosen using Census data. Toward the end of the field period, the collection allowed for some flexibility in the quotas in order to conclude the data collection in a timely fashion. However, the final deviations from the quotas are very minimal, as can be seen in Appendix B.1.

Figure 2-1: Reasons for Job Search



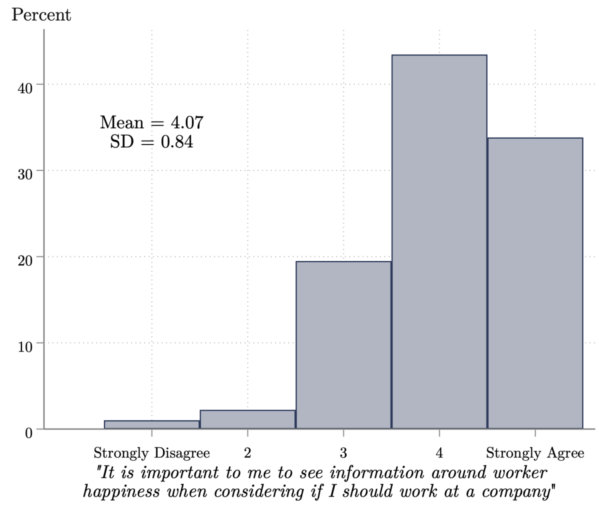
dent about different aspects of the subjective work experience at their company, including: achievement, appreciation, belonging, energy, flexibility, inclusivity, learning, management quality, fair pay, purpose, support, and trust (see Appendix A.1).

In addition to these main items on the workplace experience, I also include in the survey a range of questions designed to measure people's preferences for happiness in the workplace, their reasons for job search, and their general beliefs about wellbeing at work. Finally, I also embedded a survey experiment into the data collection, with the intention of eliciting people's willingness to trade off salary in order to work at happier workplaces.

2.3 Descriptive Evidence

One initial way to try to understand the extent to which employees value workplace happiness is to simply ask them what reasons they might have for looking for new or different jobs than the one they currently have. To this end, Figure 2-1 shows the main reasons people give for looking for new opportunities – given the options of the 13 workplace experience items asked above and on *Indeed* platform.

Figure 2-2: Preferences for Company Happiness Information



Note: The figure shows the distribution of responses on a 1–5 agreement scale to the statement shown. The sample is 4,033 survey responses from an online survey representative of the US labor force.

Compensation is hugely important, as one might expect, particularly given the weight given to the role of wages in standard models of job search (e.g., McCall, 1970; Mortensen, 1970) as well as work identifyign what is important to people in a 'good job' (Jencks et al., 1988; Krekel et al., 2019). But it is far from the only reason people may think of looking for other jobs. The most frequently cited reason is that the respondent is not being paid fairly for their work, but more affective motivations are also clearly present – for example, nearly 20% say that not feeling happy (or energized) at work would be a main reason to look for another job. More eudaimonic reasons are also frequently given, such as not learning new things or achieving goals.

This initial question relies on information an employee has about their own current experience – what one might call "push" factors. To build on this, I also ask respondents to think about the types of information they would like to have available to them when looking for new opportunities – in other words, what might "pull" them toward other companies.

To this end, respondents were asked to:

Imagine that a website has asked millions of workers across the country about *their happiness at work, in order to help job seekers make more informed decisions about where to work. After surveying 4.5 million people working at different companies, this company has published this new data for each company.*

Following this introduction to the concept of the happiness score, respondents were asked a relatively simple, non-behavioral question about the usefulness of this information for making job choice decisions. Figure 2-2 reports the distribution of responses to this question, showing that 77% of respondents agree or strongly agree that it would be important to them to see information about worker happiness when considering whether to work at a company. However, this descriptive evidence should be treated with caution, particularly as there are no constraints. For this reason I build on it in the following section using a discrete choice exercise designed to elicit people's willingness to trade off salary for happiness.

2.4 Survey Experiment

2.4.1 Experimental Design

To go beyond the initial descriptive suggestive evidence, I attempt in this section to elicit respondents' willingness to pay for (or accept) workplace happiness. Respondents are randomized to be in either a willingness-to-pay (WTP) survey exercise or a willingness-to-accept (WTA) one. Half the sample are randomly assigned to be in the group that is asked two questions about jobs with higher levels of happiness but lower pay, and the other half is asked two questions about lower levels of happiness but higher pay – giving a total of around 8,000 choice scenarios.

In each case, respondents are asked to choose between hypothetical job offers with differing levels of happiness and wage. In all cases there is a base job ("Position 1") with a happiness level of 65 and the same level of wage as the respondent's current job (or their last job, if unemployed). In the WTP exercise, respondents are offered a second job ("Position 2") with higher happiness but lower pay. In the WTA exercise, respondents are offered Position

B with higher pay but lower happiness. The design of both procedures relies on an iterative bidding procedure, in which respondents are offered Position 2 at a given happiness level with a sequential series of wage offers until they decide to switch away from Position 1. The wordings and procedure is outlined in more detail below.

Willingness-to-pay

For the half of respondents randomized into the WTP exercise, they are presented with the following:

> *Imagine you are looking for a job on the website we have described. You are comparing two positions, both of which are in the same industry and location as your [current/last] job. In this instance, assume the positions, job description and companies are the same outside of pay and happiness levels. Which position would you be more likely to choose?*
>
> 1. *Position A: The workplace happiness score of the company is 65. This position pays you the same as your current job.*
> 2. *Position B: The workplace happiness score of the company is [75,85]. The position pays you [35, 20, 10, 5, 2]% less than your current job.*

Each respondent is asked this question twice – once with position B being a happiness level of 75 and once with a happiness level of 85. In each case, they are first asked whether they would take position A or B when B pays 35% less than their current wage. If they choose A, they are then asked if they would take B when B pays 20% less – and so on in an iterative process. If the respondent continues to choose position A when offered a 2% lower salary for B, this is then coded as "none."

Willingness-to-accept

For the half of respondents randomized into the WTA exercise, they are instead presented with the following:

> *Imagine you are looking for a job on the website we have described. You are comparing two positions, both of which are in the same industry and location as your [current/last] job. In this instance, assume the positions, job description and companies are the same outside of pay and happiness levels. Which position would you be more likely to choose?*
>
> 1. *Position A: The workplace happiness score of the company is 65. This position pays you the same as your current job.*
> 2. *Position B: The workplace happiness score of the company is [45,55]. The position pays you [2, 5, 10, 20, 35]% more than your current job.*

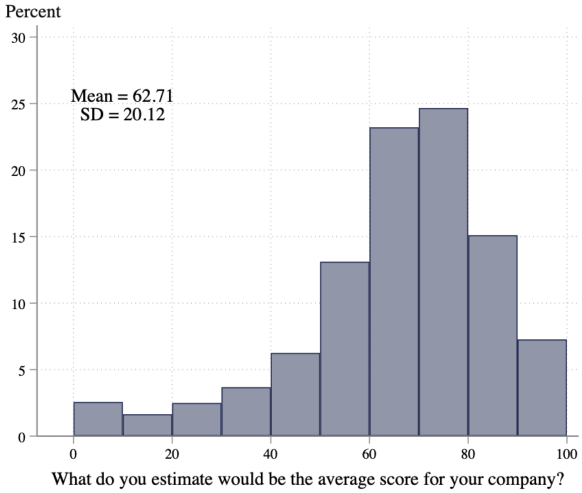
Each respondent is asked this question twice – once with position B being a happiness level of 45 and once with a happiness level of 55. In each case, they are first asked whether they would take position A or B when B pays 2% more than their current wage. If they choose A, they are then asked if they would take B when B pays 5% more – and so on. If the respondent continues to choose position A when offered a 35% higher salary for B, this is then coded as "none."

2.4.2 Calibration of Experiment

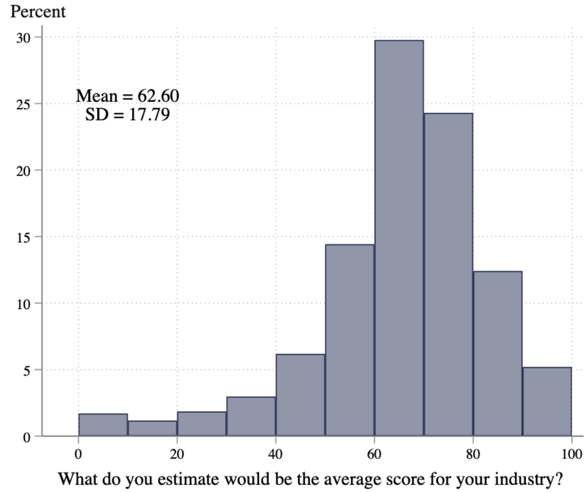
Understanding of the Scale

In all cases, the default job is 65. One possibility is that the choice of baseline job may bias any estimates. Encouragingly, the broadly similar approach of Maestas et al. (2018) found little evidence of sensitivity of estimates according to the baseline job they used. Nevertheless, in order to calibrate the experiment, I orient the respondents more clearly to the scores as well as conduct an understanding check. I first ask respondents to:

Figure 2-3: Understanding Check on Score Ranges



(a) Own Company



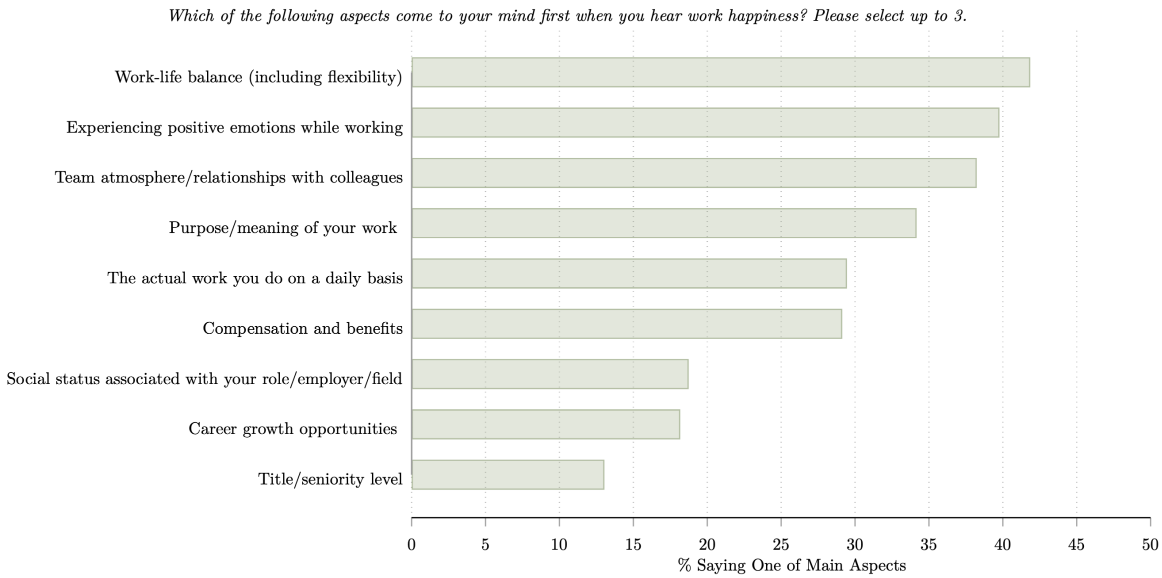
(b) Own Industry

Note: Figure plots the distribution of responses from 4,033 survey respondents in the USA. Survey was conducted in March 2021 – see text for more details.

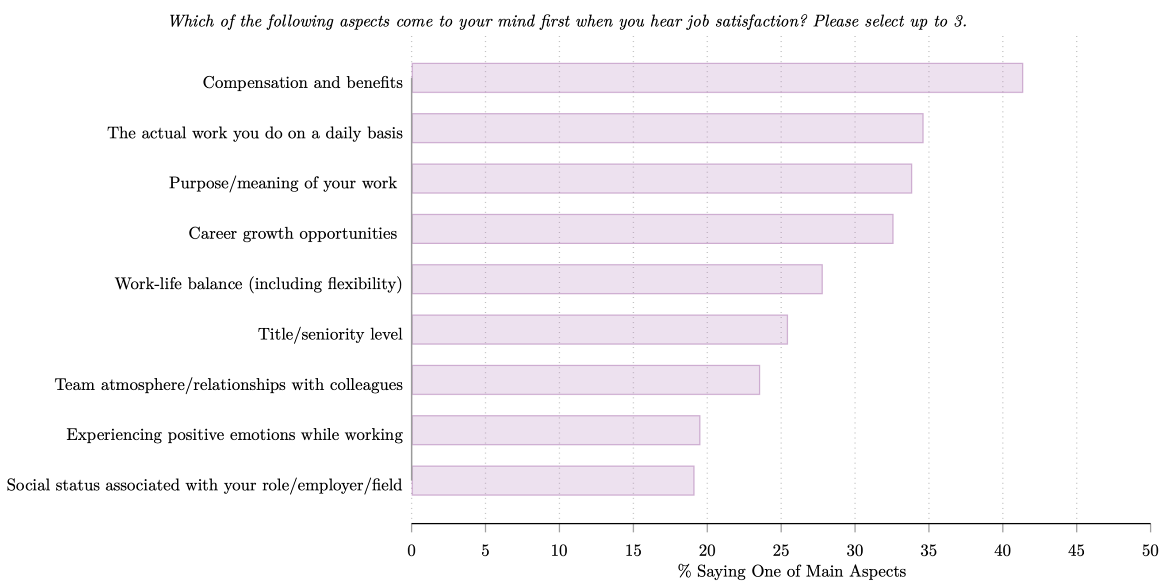
> *Imagine that employers are all given a score on happiness out of 100 points based on employees' answers to the question "I feel happy at work most of the time." Across all employers on the website, the average happiness score is typically in the range between 60 and 70.*

I then ask respondents to estimate the happiness scores of their own company and industry. Prior to being given the choice scenarios, respondents are asked to guess what they think the score would be of their current company or of companies in their industry more generally. As can be seen in Figure 2-3, the mean response is 62.7 for own company and 62.6 for own industry (compared with the actual mean across companies in the *Indeed* data, studied in Chapter 1 of 62.8). However, despite this, it is worth noting that the spread of responses is greater: the standard deviation when asked to guess the happiness level of their company and industry is 20.1 and 17.8, respectively (compared with the actual standard deviation in the *Indeed* data of 8.8).

Figure 2-4: How do people interpret "workplace happiness"?



(a) Affective Happiness



(b) Job Satisfaction

Note: N=4,033 individuals surveyed in the USA in March 2021.

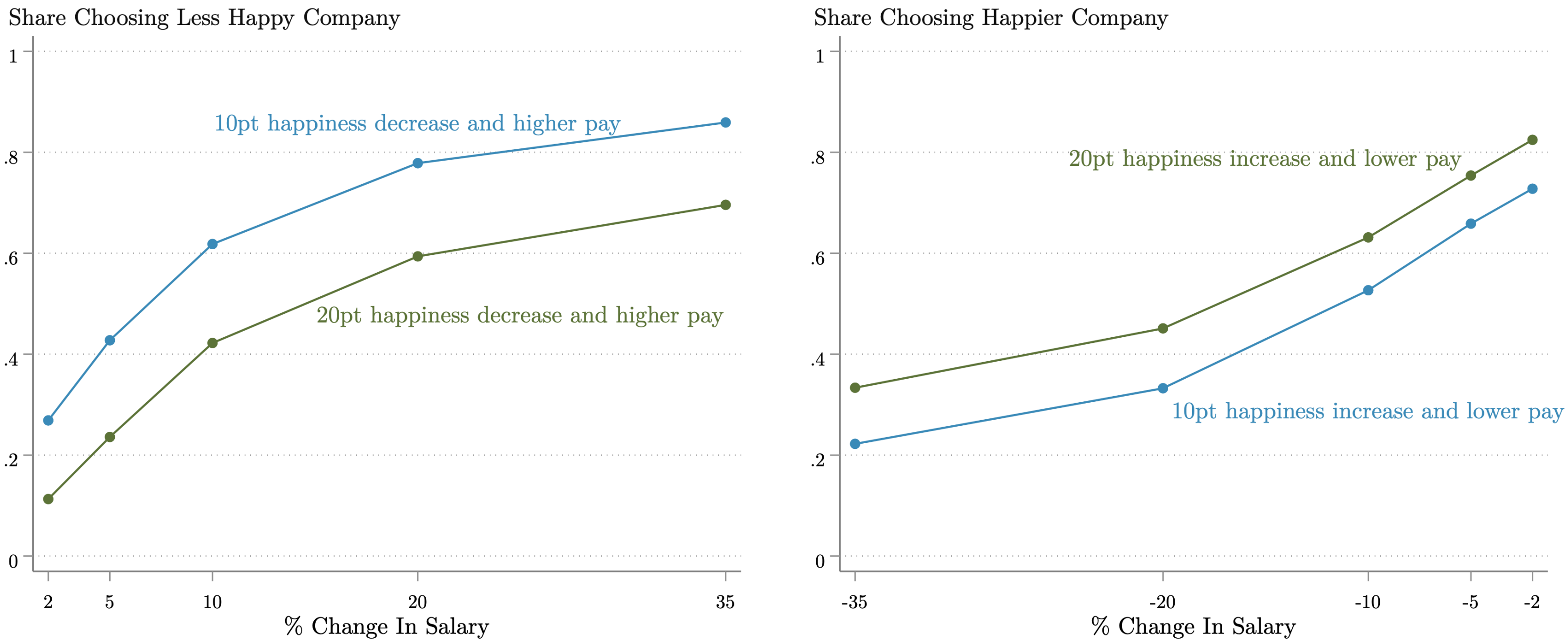
What do job seekers understand by workplace happiness?

Another potential concern with the experimental design is that it may be unclear what workplace happiness means to people. As discussed above, subjective wellbeing is typically thought of as having two main components – evaluative and affective (Diener et al., 1999). Evaluative measure of workplace wellbeing refer to how people think about their jobs overall, and are typically measured using concepts like job satisfaction. Affective measures, on the other hand, refer not to how people *think* about their jobs but rather to how they *experience* them, in other words how they feel while at work.

When a respondent sees a high or low work happiness score, what is she likely to be inferring from that? I ask people the following relatively straightforward question: "Which of the following aspects come to your mind first when you hear work happiness?" Respondents are asked to select up to 3. As can be seen in Figure 2-4, the most common response is work-life balance, followed by social relationships, and enjoyment while at work as well as purpose derived from it. At least part of workplace happiness is understood to be about compensations and benefits, though less so than the aforementioned aspects of work.

Respondents appear to largely understand the demarcation, typically theorized and described in the academic literature, between affect and evaluation. Evaluative measures of wellbeing are more typically used in labor economics and are most frequently linked with aspects of jobs like compensation and career growth. Indeed, when I ask the same question in relation to job satisfaction, the most frequently-cited aspect of a job that comes to mind is compensation and benefits. Ultimately, more research—preferably using more in-depth qualitative methods—is required in order to better understand what job seekers are encoding when viewing happiness scores. However, this relatively simple survey exercise suggests that people do understand that happiness scores are tapping into the affective nature of jobs – i.e. how people experience them and not just how they think about them.

Figure 2-5: Survey Evidence of Willingness To Pay for Workplace Happiness in the USA



(a) Higher Pay, Lower Happiness (WTA)

(b) Lower Pay, Higher Happiness (WTP)

Note: In all cases, the base job is at a company with a happiness score of 65. Panel (a) reports the share of respondents preferring the company with lower happiness, at differing levels of pay increase. Panel (b) reports the share of respondents preferring the company with higher happiness, at differing levels of pay decrease.

2.5 Results

2.5.1 Main Findings

Results from the survey experiment in the USA are reported in Figure 2-5. In Panel (a), I plot the cumulative share of respondents choosing Position B at differing levels of salary, in the instances where Position B involves a move down from a company with a score of 65 to a company with a score of either 45 or 55. When comparing between staying at a company with a score of 65 to a much less happy company with a score of 45, when offered a 35% increase in pay, around 70% would switch—but 30% still say they would not go to the company with a much lower happiness score. Offered a 20% increase in pay, 60% would switch. Around 40% of those offered a 10% pay raise would take the job in the much less happy company. A similar pattern is found when offering respondents hypothetical positions at a company with a happiness level of 55, though willingness to switch to such a job is lower, as would be expected, given the smaller jump downward in happiness.

Panel (b) of Figure 2-5 reports the extent to which respondents are willing to accept a lower level of pay to move to a company that has a generally happy workforce. The pattern of results is similar. Overall, the data suggest that workers value happiness as a workplace amenity and would be willing to trade off salary to get it. I also conducted the same analysis in Canada and the UK. The distribution of responses is similar in all three countries, as can be seen in Figures B-1 and B-2.

While Figure 2-5 provides the distribution of respondents' WTP and WTA, it is also informative to calculate the mean level in each of the four cases (a 10-point increase in happiness, 20-point increase, 10-point decrease, and 20-point decrease).[4] Figure 2-6 reports the mean pay changes across the sample at which respondents would be willing to choose

[4]Given the design of the survey instrument, in each case, there are some respondents who chose Position A regardless of the wage offered for Position B. Here, I make the conservative coding choice of taking the penultimate wage offer (i.e. the one before "none") as being the one they would take. However, I can also test the sensitivity of estimates to this choice and code alternatives (such as 50% or 0% in the WTA and WTP exercises, respectively).

Position B—i.e., switch away from the job that has a happiness score of 65. In the USA, to take a job at a company with a score of 75 (compared with 65), respondents would be willing to pay, on average, around 13% of their of their salary. For a job with a happiness score of 85 (compared with 65), they would be willing to pay around 17% of their salary – suggesting a diminishing marginal return. As is frequently found in the valuation literature across a range of contexts (e.g. Johnston et al., 2017; Kim et al., 2015),[5] the willingness-to-pay estimates are generally lower than willingness-to-accept estimates in this exercise. In the USA, to take a job at a company with a score of 45 (compared with 65), respondents would require, on average, around a 14% pay raise compared to their current salary. For a job with a very low level of happiness (45 compared with 65), respondents would only take it, on average, if they were to be offered around a 20% raise.

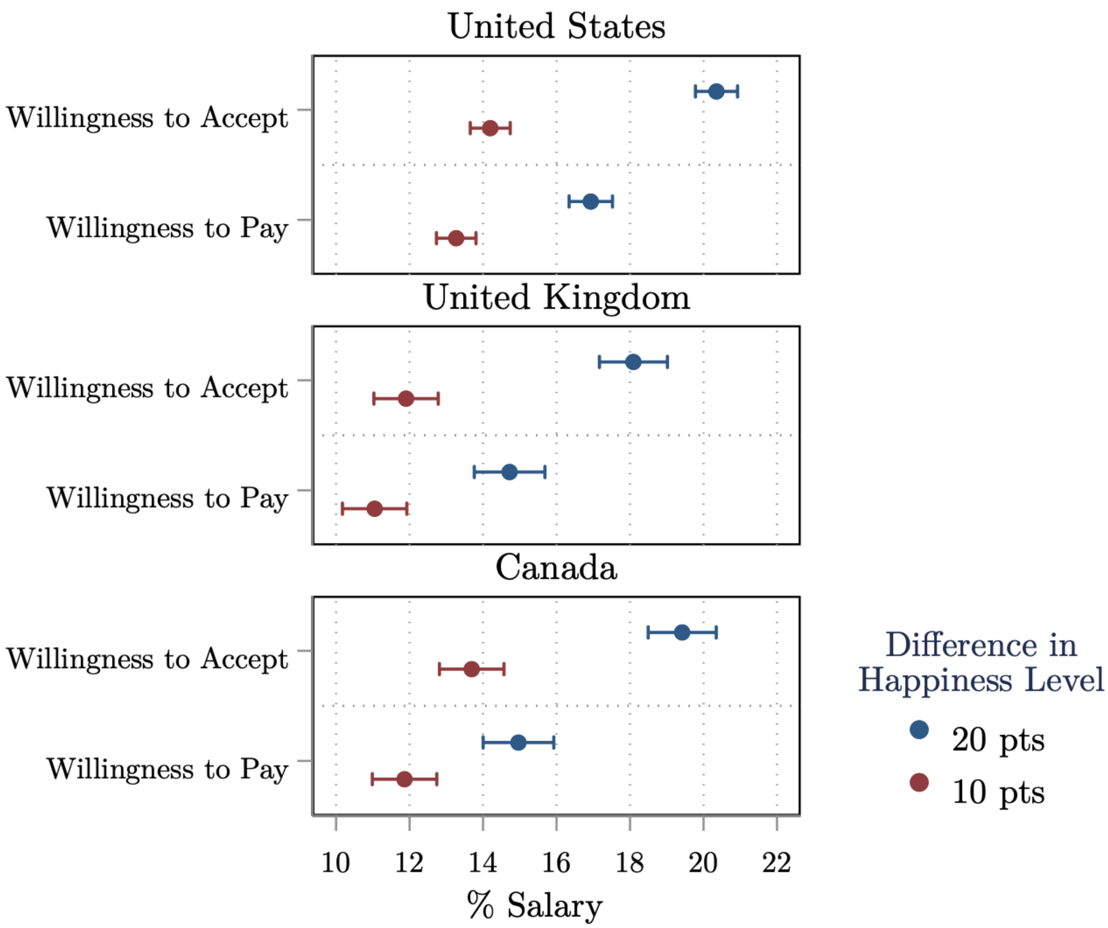
2.5.2 Heterogeneity

The mean WTP and WTA figures mask a large amount of heterogeneity. Indeed, even though just over 30% of people would take a 35% pay cut to work at a company with a high happiness score (85 compared to 65), nearly 20% would not even take a 2% pay cut to do so. What lies behind this variation? What drives people's differing preferences and/or constraints?

In Table 2.1, I begin to assess this heterogeneity by simply regressing the wage at which the respondent switches to position B on a range of observables including race, gender, income, and education. As above when calculating means, I code the "none" option as the last wage offer that the respondent was given (but refused). However, in further analyses, I can also relax this assumption and instead estimate ordered logistic regressions – where the [2,5,10,20,35,none] set of wage responses are an ordinal dependent variable in the WTP analyses (or [35,20,10,5,2,none] in the case of WTA). These regressions are reported in Ta-

[5] Brown (2005) find, for example, that WTP estimates typically lower owing to a general reluctance to suffer a net loss from any transaction and tendency to consider a sale much below any assumed market price as a loss.

Figure 2-6: Valuation of Happier Workplaces



Note: Mean estimates are shown with 95% confidence intervals.

bles B.6-B.9, and here I find similar patterns of findings as the more straightforward linear regression analyses in Table 2.1.

Consistent differences are found, whether considering WTP or WTA, across countries – with respondents in the UK being less willing to trade off salary for workplace happiness (see Table 2.1). There are also consistent differences according to respondents' current level of happiness in their job. Workers who currently enjoy higher levels of workplace happiness are more willing to trade off wage for happiness. Currently-happier workers require a higher wage raise to work at a low happiness firm and are willing to take a larger wage cut to work at a happier company. Marginal effects from the ordered logit regression, for example, suggest

Table 2.1: Heterogeneity

	WTA: ln(% Wage Raise Needed to Work at Less Happy Company, With Score):		WTP: ln(% Wage Cut Willing to Take to Work at Happier Company, With Score):	
	(1)	(2)	(3)	(4)
	45 v. 65	55 v. 65	75 v. 65	85 v. 65
Demographics				
Men	-0.052	0.043	0.087**	0.064
	(0.066)	(0.041)	(0.042)	(0.042)
Age	0.028	0.035***	0.004	-0.003
	(0.016)	(0.011)	(0.012)	(0.012)
Age^2	-0.000	-0.000***	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Married	-0.011	-0.064	0.045	0.021
	(0.101)	(0.050)	(0.050)	(0.052)
Has Children	0.049	-0.011	0.161***	0.201***
	(0.023)	(0.047)	(0.047)	(0.048)
Has BA	0.094**	0.041	0.036	0.074
	(0.020)	(0.046)	(0.046)	(0.047)
Black (v. white non-Hispanic)	-0.106**	-0.122	0.190**	0.188**
	(0.016)	(0.089)	(0.090)	(0.091)
Hispanic	-0.145	-0.110	0.222*	0.224*
	(0.111)	(0.112)	(0.116)	(0.115)
Asian	-0.060	-0.016	0.008	0.041
	(0.079)	(0.067)	(0.068)	(0.071)
Other race	-0.033	-0.085	0.097	0.106
	(0.070)	(0.099)	(0.098)	(0.102)
Current Job				
Work Happiness	0.105**	0.086***	0.028	0.042**
	(0.012)	(0.021)	(0.021)	(0.021)
Income (log)	0.100***	0.119***	-0.021	0.008
	(0.010)	(0.036)	(0.036)	(0.037)
Tenure (company)	-0.007	-0.008**	-0.004	-0.007*
	(0.003)	(0.004)	(0.004)	(0.004)
Current Weekly Work Hours	-0.000	-0.001	0.004	0.006**
	(0.001)	(0.002)	(0.002)	(0.002)
County (v. USA)				
UK	-0.098**	-0.178***	-0.146***	-0.134**
	(0.015)	(0.052)	(0.052)	(0.053)
Canada	-0.040**	-0.046	-0.059	-0.073
	(0.009)	(0.051)	(0.052)	(0.053)
Observations	2926	2926	2928	2928
R^2	0.035	0.032	0.040	0.064

Notes: Robust standard errors are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

that a one point increase in the 1-5 current happiness scale increases the likelihood that the respondent will say there is no wage at which they would take the job with happiness level 45 (vs. 65) by 0.041 (see column 7 of Table B.6).

Heterogeneity along other characteristics is identifiable, but frequently with differences between the WTP and WTA estimates. Levels of accumulated human capital, for example, are not very significantly related to people's willingness to pay, in terms of wage, in order to work at a higher happiness company. However, both income and education are strongly and positively associated with people's willingness-to-accept – that is, the wage raise at which they would decide to take a job at a less happy company. Marginal effects from an ordered logit regression suggest that a one point increase in logged income increases the likelihood that the respondent will say there is no wage at which they would take the job with happiness level 45 (vs. 65) by 0.033 (see column 7 of Table B.6). To aid interpretation, this relationship between current income and WTA is shown graphically in Figure B-3.

Willingness-to-pay estimates are higher for parents than non-parents. Equally, both Black and Hispanic respondents, compared to white respondents, are willing to pay more in terms of wage to work at happier workplaces – holding constant factors like income, education, and age. There are no significant differences between Asian and white respondents, however. Marginal effects from an ordered logit regression suggest that being Black (as compared to white), increases the likelihood of being willing to take a 35% pay cut to work at a company with a score of 85 (vs. 65) by 0.08 – a figure that rises to 0.11 for Hispanic respondents (see column (7) of Table B.9). Future research will be required in order to better understand this heterogeneity. One very tentative hypothesis may be that for Black and Hispanic workers, there is more at stake when it comes to company-levels of happiness. This may be because these groups of workers frequently suffer from a range of discrimination, microaggressions, and so on, in their day-to-day work lives. The hypothesis that there is more at stake is also consistent with the finding that willingness-to-pay estimates are higher for employees who currently work longer hours (see column (4) of Table 2.1). This is potentially reflective of the fact there is more at stake in terms of having a happy vs. unhappy place to work, the larger the proportion of one's life that is spent at work.

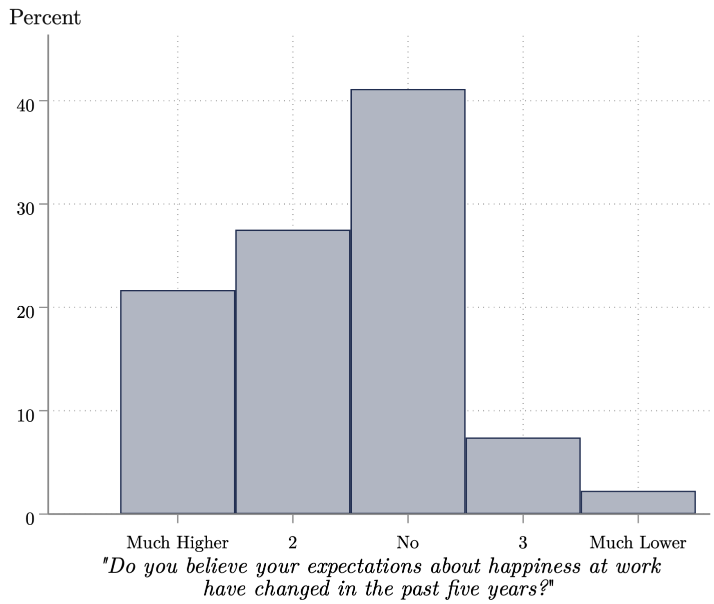
2.6 Discussion

The past decade has been one of rapid change in the workplace, a process that has only been heightened during the recent global pandemic. Amid transformations in the ways in which work is organized, managed, and carried out, the relationship between employers and employees is also in flux. Half of respondents in the survey carried out in this chapter report that their expectations about workplace happiness have risen over the past 5 years (see Figure 2-7). The majority of managers in the US also believe that expectations among employees about workplace happiness have gone up (HBR Analytical Services, 2020). But despite this conjecture, the extent to which workers really care about happiness is an open empirical question – with potentially important implications for how firms may think about going about attracting and retaining their workforce.

2.6.1 Managerial and Policy Implications

Even if people value workplace happiness, does this necessarily have anything to do with firms? It may not even be something within their control, after all. One initial approach to answering the question of whether firms can reasonably be expected to do anything is to ask workers themselves. To begin to shed light on this, I include in the survey a question about who is responsible for employees' happiness. Here, I find that people say that individuals themselves take on at least some responsibility. But they also see a large role for management—from the CEO down to line managers—to influence the well-being of workers (Figure 2-8). Similarly, from the 'other side,' in a survey of US managers run by HBR Analytical Services (2020), respondents were asked how much control they believed their organization has when it comes to influencing the happiness of its workforce. Some

Figure 2-7: Worker Expectations

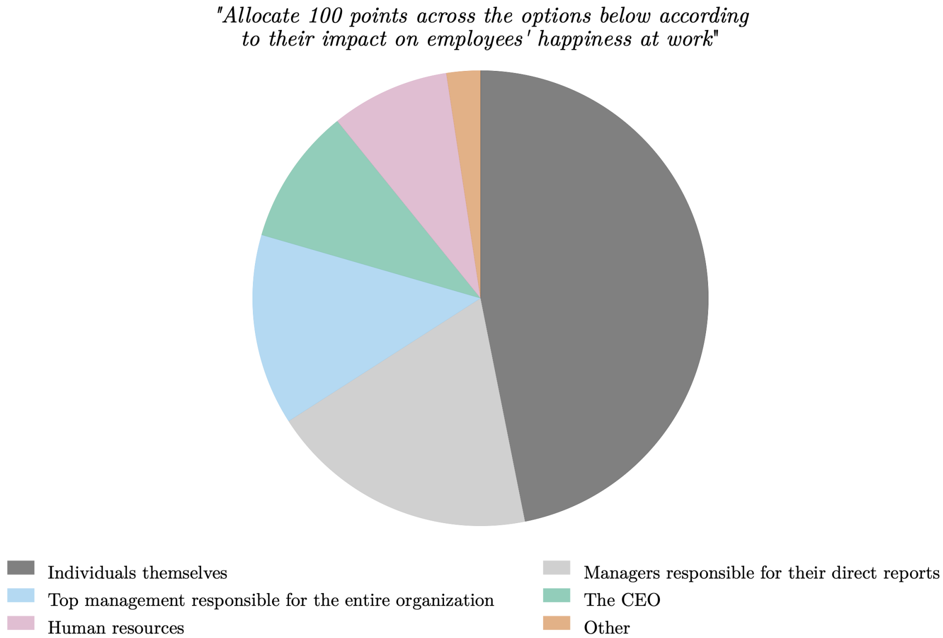


Note: N=4,033 individuals surveyed in the USA in March 2021.

38% said a high degree of control, and 57% said some control, with only a small number saying that organizational and management practices are irrelevant to people's happiness (HBR Analytical Services, 2020).

In addition to this initial evidence, a growing body of work across the social and behavioral sciences shows that a range of managerial and organizational practices are causally linked to worker levels of happiness. In a field experiment on airline pilots by Gosnell et al. (2020), for example, the authors find effects on workplace happiness of various management practices – including monitoring, performance information feedback, personal targets, and pro-social incentives. Equally, Bloom et al. (2014) find that working from home improves emotional experience (as well as satisfaction) of call center workers in a Chinese travel agency, while Moen et al. (2016) find that a program designed to increase supervisor support and employee independence improved various measures of worker subjective wellbeing. While more research is required in this direction, the general finding is a fundamental one: the ways in which work is managed and organized—which can be changed—have significant effects on

Figure 2-8: Who is responsible for worker happiness?



Note: N=4,033 individuals surveyed in the USA in March 2021.

worker subjective wellbeing.

2.6.2 Limitations and Future Research

The survey in this Chapter has the benefit of taking place using a broadly representative sample of US workers. This was achieved by the survey company introducing quotas for age, education, gender, region, and income. Nevertheless, the sample is drawn from an online panel and the representativeness could be improved upon by using more elaborate (and expensive) survey collection techniques. Additional quotas for important demographic categories such as race and ethnicity, for example, would be beneficial in improving data quality and generalizability.

Replications were conducted in the UK and Canada, which improves the generalizability of the findings. However, these two countries are both also liberal market economies (cf. Hall and Soskice, 2001) and are not hugely culturally distant to the USA. Further research across a broader range of cultures and economies would allow for more in-depth heterogeneity analysis along institutional and/or cultural lines. One might expect, for example, that in countries with stronger labor market regulation workers may be less willing to trade off wage for company happiness since there is less at stake if the "floor" is likely to be higher in terms of employer treatment.

Naturally, the discrete choice exercise is somewhat abstract and should be interpreted with the appropriate amount of caution. People are making choices between hypothetical jobs that differ along two characteristics, with no consequential outcome to the respondent. Nevertheless, people's preferences in the abstract world of hypothetical choice scenarios are consistent with behavior observed in the natural field experiment described in Chapter 1, where, across all three countries, job seekers are viewing real companies and making real-stakes decisions on which companies to apply to.

The current Chapter attempted to elicit both willingness-to-pay as well as willingness-to-accept estimates, in order to triangulate on people's valuation of happiness. This is based on the assumption that no valuation methodology is perfect (or research methodology more generally, for that matter). Further work may look to use different research designs in order to further triangulate and minimize bias in these initial valuation estimates. For example, in this preference elicitation exercise, I use here an iterative bidding procedure in which respondents are given sequential wage offers at differing levels of happiness. This has the benefit of conceptual simplicity; however, it may introduce bias to the extent that people may anchor on the initial wage offered. Further research should look to randomize the initial wage that is offered, before proceeding to sequentially offer higher or lower wages. Alternatively, further work should run a between-subject survey experiment in which respondents are asked (in a single step) to choose between two jobs, but where the wage and happiness level are both randomly allocated across subjects.

Finally, it may be argued that the elicitation of people's preferences about happiness is of little interest since such information is never actually available to workers at scale in the real work. The rapid digitization of the labor market, however, is making this less true – and information about company-level happiness is now routinely shown to job seekers on websites such as *Indeed*. Given that this is something that workers appear to value, digitization of labor market institutions—which increases knowledge diffusion—has the potential to increase worker power, and in doing so provide incentives to employers to improve worker happiness.

2.7 Conclusion

Employee wellbeing is increasingly discussed by at least some companies as a priority, a development that has only been accelerated by the COVID-19 pandemic in which large-scale workplace change was a key feature for many workers. But do organizations actually have any incentives to invest in managerial and organizational practices that might improve the experience of work? Much depends on the extent to which workers value happiness. The greater the value is that people place on workplace happiness, the more that company-level happiness is likely to affect turnover and recruitment. The evidence presented in this chapter suggests that people do value workplace happiness and would be willing to trade off salary in order to work at happier companies.[6]

[6] I am extremely grateful to *Indeed* for generously funding the survey and to *Forrester* for helping to design and, ultimately, carry out the survey experiment. In particular, Janeane Tolomeo, Katie Stevenson, Rae Albee of *Indeed* as well as Aliona Madjar of *Forrester* generously provided excellent advice and assistance with the experimental design. For pre-registration details of the experiment, see AEA RCT Registry AEARCTR-0007359. The study was approved by the MIT Committee on the Use of Humans as Experimental Subjects (Approval Number E-3045).

Chapter 3

Does Happiness Have an Impact on Productivity? Quasi-Experimental Estimates in the Field

By *Clément Bellet, Jan-Emmanuel De Neve, and George Ward*

3.1 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 in their jobs. While this recent focus on employee happiness may seem a relatively new development, the study of employee well-being and performance does in fact have a long history that stretches back almost a century.[1] However, isolating the direct contribution of happiness to productivity—particularly in field settings—has remained an empirical challenge, leading to mixed results (Tenney et al., 2016).

[1]In reaction to early Scientific Management theories, which imported tools from engineering in order to study labor productivity, the Human Relations movement of the early 1930s sought to place human factors such as motivation and attitudes at the center of their studies of factory production (Vroom, 1964). Beginning with the Hawthorne studies of the late 1920s, this movement—and subsequent scholarship in related fields like human resource management (HRM) and organizational behavior (OB)—placed an emphasis on employee satisfaction and studied its relationship with productivity (see, e.g., Hersey, 1932b; Kornhauser and Sharp, 1932, for early examples).

In this chapter, we study the causal productivity effects of employee affective states. We use highly detailed data on the behavior and performance of the universe of telesales workers at one of the United Kingdom's largest private employers, British Telecom (BT). We link this administrative data with a survey instrument that we designed to measure the week-to-week happiness of employees over a six month period. Central to our study is the capacity to identify and leverage exogenous mood shocks. To that end, we turn to data on weather patterns local to each of the 11 call centers. We hypothesize that only the most psychologically-relevant aspects of weather should affect mood, and that any mood effect should be contingent on workers' visual exposure to weather. We confirm this intuition and show that while visually gloomy weather has a strong negative impact on worker mood, this effect is dependent on window coverage of each call center. Exploiting the plausibly exogenous variation in happiness arising out of these differences in the interaction between architecture and weather, we provide quasi-experimental estimates of the effect of happiness on sales.

Call centers provide a particularly good setting in which to study productivity – that is, the residual variation in output that cannot be fully explained by observable inputs (Syverson, 2011). Workers in our setting are paid an hourly wage to do the same job, which is to take incoming calls from new and existing customers, and seek to sell them various products such as broadband internet contracts, cell and landline phone deals, and television packages.[2] To do so, they use the same phones and computer system, with the firm routinely collecting very detailed data on a large number of labor inputs like workers' daily attendance, sickness, working minutes while at work, and fine-grained break-taking behavior. Rather than having to rely on subjective outcome measures—such as managerial performance evaluations—by studying telesales workers we are able to use administrative data on a clear and objectively measurable output: sales.[3] To investigate channels through

[2]Workers are paid a fixed hourly wage. On top of this, they may receive a bonus amounting from 10% to 30% of their base salary if they perform well. Note that this is not an explicit piece-rate but rather a slow-moving bonus scheme.

[3]A large number of existing papers have relied on employee self-reports of productivity or on managerial

which any effect of happiness may translate into sales, we also study fine-grained measures of behavior such as working speed (calls per hour), the percentage of calls converted to sales, and the extent to which workers "adhere" to their computer-scheduled workflow.[4]

We add to a long history of work on the determinants of labor productivity (Syverson, 2011). While answers tended historically to point towards factors like human capital (Abowd et al., 2005) and monetary aspects of work like incentive pay (Lazear, 2000), an alternative stream of research has cast the net wider in order to focus on further aspects of work like intrinsic motivation, team incentives, social preferences and connections, fairness concerns, employee recognition and meaningfulness of tasks (e.g. Bandiera et al., 2009; Bradler et al., 2016; Cassar and Meier, 2018; Delfgaauw et al., 2022; DellaVigna and Pope, 2017; Gneezy and Rustichini, 2000). In this chapter, we build upon a lengthy tradition of work on the extent to which employee subjective well-being (SWB) affects employee performance.

SWB is typically thought of as having both cognitive and affective components (Krueger et al., 2009). A large initial stream of research examined the association between job satisfaction and worker performance, and has generally found a modest positive (partial) correlation between the two (see Judge et al., 2001, for a review). A more recent stream of research has moved away from cognitive assessments of job satisfaction, in order to study affective (or hedonic) SWB in the workplace.[5] Among others, two key benefits of studying affective states are that they a) vary more frequently over time and b) are also manipulable in laboratory settings.

While much of the earlier work on SWB and performance was based on cross-sectional

evaluations (e.g. Staw and Barsade, 1993; Zelenski et al., 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 large amounts of measurement error. For example, there may be 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.

[4]Although metrics like speed are sometimes seen as measures of productivity in themselves, we prefer to remain agnostic as to whether making many faster calls is any more productive than making fewer longer calls. Indeed we show some evidence that speed is only weakly related to sales, with the direction of the relationship suggesting that longer calls are more conducive to sales.

[5]For an extensive review of the evidence arising from both – affective and evaluative – streams of research on employee well-being and performance, see Walsh et al. (2018).

between-worker comparisons, more recent studies have been able to leverage within-worker variation in longitudinal research designs that assuage concerns related to unobserved heterogeneity between workers (see, e.g. Koys, 2001; Miner and Glomb, 2010; Rothbard and Wilk, 2011; Staw and Barsade, 1993; Staw et al., 1994).[6] While these studies have provided a great advance in the field, there remains the possibility that omitted time-varying confounding variables could still bias any estimated effects. The fact that affective states are not only more variable but also more manipulable, however, has allowed for more clearly causal research designs, at least in laboratory settings. A recent pioneering paper in this respect induces happiness in the lab to show a robust causal effect on a stylized, piece-rate productivity task (Oswald et al., 2015).[7] However, the extent to which these affect-induced productivity effects i) translate into real-world employment settings and ii) are persistent beyond a very short-lived mood induction procedure in the lab, remains an open question.

Despite this lack of causal field evidence, an emerging field-experimental literature shows that management practices can have simultaneously positive impacts on i) employee happiness and satisfaction as well ii) as productivity (Gosnell et al., 2020). From pay inequality (Breza et al., 2018; Cullen and Perez-Truglia, 2020) to gift exchange (DellaVigna et al., 2020) and work autonomy (Bloom et al., 2014), this line of research suggests employee happiness as one possible channel through which workplace organization may feed through to productivity – but has been unable to isolate this channel in a causal manner.[8] In this chapter, we more clearly identify the happiness-productivity causal channel.

Beyond productivity, affective states are increasingly seen by academics as a potentially important factor in driving human behavior more generally (see, e.g., Loewenstein, 2000).

[6]A small number of studies have also demonstrated that prior affective states predict subsequent performance. 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.

[7]For earlier experimental work in psychology showing the effects of positive mood on task performance in a non-incentivized setting, see Erez and Isen (2002), as well as processes like motivation, cognitive flexibility, negotiation, and problem solving skills (see Isen, 2001, for a review). Recent related experimental work inducing stress states in the lab have also shown causal effects on a piece-rate task, at least for women Cahlíková et al. (2020).

[8]A number of papers also investigate the impact of work arrangements on job satisfaction (Card et al., 2012; Clark et al., 2009) or happiness (Moen et al., 2016), without measuring productivity itself.

Laboratory experiments show such effects, for example, on outcomes including time preferences (Ifcher and Zarghamee, 2011), reciprocity (Kirchsteiger et al., 2006) and behavior in ultimatum and trust games (Capra, 2004). We therefore join a growing literature that has attempted to extend this work to field settings. This research typically proceeds by estimating the effects of various proxies for mood, such as weather or sports results, on economic outcomes like stock returns, consumption and real estate transactions (e.g. Agarwal et al., 2020; Edmans et al., 2007; Hirshleifer and Shumway, 2003; Hu and Lee, 2020; Li et al., 2017; Saunders, 1993). We label this the reduced-form approach to mood effects in the field. Indeed, implicit to this methodology is an instrumental variables (IV) set up, whereby i) weather has an impact on mood and ii) it affects behavior solely through that mood mechanism.

One source of skepticism with the existing reduced-form approach to studying mood effects using weather variation is that evidence for the implied first-stage impact of weather is not necessarily a given (Feddersen et al., 2016; **?**), since the empirical relationship between weather and affective variables like feelings of happiness can be unstable (Denissen et al., 2008).[9] 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, only when people are experimentally assigned to be outdoors.[10] 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 et al., 2002). Given this, we note two things. First, in order to convincingly demonstrate mood effects in the field it is important to measure mood and

[9] Relatedly, weather has been used as an IV in a variety of empirical settings. Unless one clearly specifies the conditions under which weather might affect the variable of interest within a specific context, this multiplicity of usage may cast doubt on the exclusion restrictions of weather as an IV (Gallen, 2020; Mellon, 2020).

[10] A related literature on shows that visually pleasant weather improves mood and prosocial behavior in studies with outdoor settings (Cunningham, 1979), though this finding has been challenged using time-series data on tipping in an indoor restaurant (Flynn and Greenberg, 2012). Rind (1996) study an indoor setting but leverage variation in the presence of windows in hotel rooms to help estimate the effects of pleasant weather on tipping behavior. Overall, this literature suggests that the salience or visibility of weather is key to any relationship with mood.

verify that there is a strong first-stage relationship in the study setting. Second, in order to convincingly use weather as a mood proxy, it is (a) useful to have variation in weather that is *visual* in nature (e.g. gloomy weather instead of cold weather) 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, such as physical sickness.

Using administrative data from weather stations in the vicinity of our 11 call centers, we create an index of gloomy weather.[11] Usefully for our purposes, the call centers are widely dispersed geographically across 3 nations – from the south coast of England to the north of Scotland – in the United Kingdom, where weather is infamously variable. Equally useful for our purposes, the call centers happen to vary significantly in terms of their architecture – from high rise buildings covered in windows at one extreme to warehouse-style buildings with almost no windows at all on the other. We are thus able to re-weight our gloomy weather index to account for the percentage of each call center's wall coverage that is glass window, such that our preferred proxy for mood reflects workers' *visual exposure to gloomy weather*. We show a negative impact of gloomy weather on sales. Moreover, we show that this effect is strongly dependent on visual exposure. Gloomy weather has no significant impact on sales for employees working in call centers with few windows. For employees working in call centers with many windows (above median window share), the effect of a one unit increase (on a 0-to-10 scale) of gloomy weather decreases sales by 1.5%. We interpret this as strong yet suggestive causal field evidence for an effect of happiness on productivity. We see this piece of reduced-form evidence as strong since the significant interaction with proximity to windows points towards a psychological interpretation. But it is only suggestive in that, without actually measuring worker affect, it is impossible to be sure that exposure to weather is driving any mood differentials among workers.

We designed a weekly survey in order to measure the psychological affective state of

[11]This is the sum of the daily incidences of fog, rain and snow during any given week. Unfortunately, the U.K. Met Office does not collect sunshine or cloud-cover data in their weather station data.

telesales workers at BT. We study a broad measure of affect: weekly feelings of overall happiness. Although happiness is a specific concept in psychology, we can think of the within-person, week-to-week variation in overall feelings of happiness as being a good proxy for the broader psychological concept of mood. One key benefit of studying happiness is that we are able to use a survey measure that has been widely investigated over multiple decades, whose properties are well understood, and whose validity and reliability have been well documented (see, e.g. Krueger and Schkade, 2008; Krueger and Stone, 2014). Using this psychological survey data, we are able to show a strong ("first-stage") impact of gloomy weather on workers' week-to-week happiness. Again, we find this effect is much more strongly negative for workers who are more exposed to the visual aspects of weather by being able to see out of windows while at work.

Putting this all together, we are then able to more directly provide quasi-experimental estimates of the effect of happiness on productivity. Our key identifying assumption here is that visual exposure to gloomy weather has an effect on sales only through this variation in measured happiness. We explore a number of threats to this assumption. One concern, for example, is that weather could have direct effects on productivity by affecting the physical health of employees (cf. Adda, 2016). Another is that weather might affect attendance by altering the opportunity cost of leisure since being outside becomes more or less desirable (cf. Connolly, 2008). We are able to show empirically that there are no weather effects on attendance or sickness, as well as control for fine-grained inputs like selling time in our main equations. But, perhaps more importantly, we also show that the impact of gloomy weather on employee happiness is contingent on exposure while at work. That is, our key piece of identifying variation is not weather itself, but rather the interaction between weather and architecture. In call centers with almost no windows, the effects of gloomy weather on happiness are negligible; in call centers with lots of access to natural light, gloomy weather strongly depresses mood. This is consistent with the idea that our instrument affects solely the *psychological* aspect of any effects variation in week-to-week weather may have on employees. A third concern is that mood effects are difficult to discern within two-sided markets such as ours, since there is a danger that the weather variation could affect the happiness of customers as well as employees – and drive sales through changes in demand (Li et al., 2017). The fact that calls within a given call center originate from all parts of the country, and the use of data from multiple, geographically dispersed call centers means that we are able to study the effects of local weather conditional on time (and individual) fixed effects, such that we identify these effects off of differences across call centers within any given week (and not from seasonal movements in national weather).

Using detailed data on worker behaviors, we are able study three potential channels behind the effect of happiness on sales. First, workers could be better at organizing their time while at work, hence adhere more closely to their schedule. Second, they could work faster, i.e. answer a higher number of calls per hour. Third, they could be more efficient at converting calls into sales. Though we find evidence for all three 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 also investigate whether happiness has any causal effect on labor supply beyond productivity itself. We find null effects on attendance, sickness leave, overtime, or length of breaks. Since in our setting employees enjoy very little discretion over their day-to-day work hours, we do not interpret this as strong evidence that happiness could not have an effect on labor supply elsewhere (positive or negative). Instead, we note that the well-defined output, together with the restrictiveness in terms of labor supply, makes for an ideal setting to cleanly estimate productivity effects.

Depending on the specification, we find a point estimate for the effect of a one unit increase (on a standard 0-to-10 scale) in happiness on sales in our IV analysis of somewhere between .08 to .15 log points. Using the point estimate from our main specification, we find an average marginal effect of around 3 additional weekly sales, over a base of 25. These findings are consistent with evidence from laboratory experiments (e.g. Oswald et al., 2015) as well as the measured productivity effects of management practices that simultaneously impacted employee happiness and productivity in randomized field experiments (e.g. Bloom et al., 2014). Our results also tally with those of a paper similar in spirit to ours, which studies the mood and performance of a sample of 29 workers in an insurance firm, finding a positive within-worker relationship between affect and task performance (Rothbard and Wilk, 2011). However, our results stand in apparent contradiction to Coviello et al. (2018), who find negative effects when studying the self-reported work engagement of (predominantly non-sales) call center worker in the USA. We discuss in detail the similarities and differences of our estimated effects as compared to the prior literature.

The chapter proceeds as follows. In Section 3.2, we introduce the institutional setting and describe the novel survey instrument we implemented to measure worker mood. We describe our empirical strategy in Section 3.3 before presenting our main results in Section 3.4. In Section 3.5 we explore psychological mechanisms before outlining and discussing managerial implications, a number of key limitations, and some possible extensions. Section 3.6 concludes.

3.2 Data and Institutional Setting

We use data from BT, a large multinational telecommunications company based in the United Kingdom. We focus our analysis on the firm's 11 call centers—located across England, Scotland and Wales—that have a large number of workers concentrating predominantly on sales.

3.2.1 Administrative Firm Data

We use detailed individual-level administrative data from the firm. We focus our attention on sales workers across the 11 call centers, whose job it is to take incoming calls and sell BT products. These predominantly consist of landline and cellphone contracts, broadband internet contracts, and television subscriptions. 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). We rely on the definition of productivity as being "the component of the production function unrelated to observable labor, capital, and intermediate inputs" (Syverson, 2011). In the case of a sales workers, productivity is the residual variation in sales that cannot be explained from variation in inputs like the number of hours worked, break-taking, or technology.

In our sample of sales workers, 70% work under some form of performance pay: they receive a bonus amounting to up to 30% of their base salary if they meet their target. 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.

Workers are observed in the data on a daily basis and are identified by their unique personal ID as well as a time-varying team ID. Workers sit individually at desks with a computer terminal and a telephone headset, and are clustered physically in the workspace by their team. Although workers are organized into teams and share of a line manager, the job of selling in this context is almost exclusively an independent task with very little to no teamwork involved.

3.2.2 Productivity and Labor Supply Data

At the worker-day level we observe the number of sales. This includes new sales, whether to a new or existing customer, as well as instances where the worker is able to retain a customer through re-contracting. The distribution of sales in our sample can be seen in panel (b) of Figure C-1. As is typically the case with sales data, the distribution is right-skewed. We also look more closely at mechanisms through which any relationship between affect and sales may take place, and in doing so consider other work outcomes like call duration, adherence to workflow scheduling, call-to-sale conversion, and customer satisfaction (see Table C.1 for summary statistics).

In addition, we observe high-frequency measures of labor supply through workers' full and highly detailed time schedule. We are able to track the number of hours the employee is scheduled to work, their actual attendance, the number and duration of breaks they take, as well as the length of time they spend working on sales (which is factored for call waiting time). Finally, using data from the Human Resources department of the firm, we observe a limited number of time-invariant characteristics of workers such as their gender, age, and tenure.

This very detailed dataset provides an exhaustive picture of the determinants of sales workers' productivity in a call center setting. In particular, we confirm that conditional on labor supply, the number of calls per hour is a weak predictor of sales (if anything, more calls per hour reduces the number of sales per worker). We come back to this in more detail when investigating potential channels through which workers' affective states may translate into sales.

3.2.3 Weather and Architecture Data

We link our data with a weather measure drawn from the National Oceanic and Atmospheric Administration's (NOAA) Global Surface Summary of the Day database. We determine the latitude and longitude of each call center, and match each center to the closest weather station available in the data, which is on average 14km away.[12] Each station reports on a daily basis a separate indicator variable for whether there has been fog, snow and rain on that day.

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

[12] See `https://data.noaa.gov/dataset/dataset/global-surface-summary-of-the-day-gsod` for more details of the data.

total number of daily incidences of fog, rain and snow during the week for which individuals reported their happiness. For the weekly index, for example, if it were to rain, snow and be foggy every day of the week, the index would be 15, if it rains for 2 days and is foggy for 1 day, the index would be 3, and so on. As argued earlier, the link between weather and mood arises from variation in weather that is visual in nature.[13] For robustness, we also collect weekly mean temperature data from the same source.

We combine the gloomy weather data with architectural data on the 11 buildings in which the call centers are housed. The type of buildings varies from warehouses, from workplaces with almost no natural light to glass tower buildings. Here we first collect all wall pictures from each 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 an image processing software. The proportion of windows ranges from 3% to 59% (for full details, see C.9).

3.2.4 Affective Well-Being Survey

SWB is typically divided into two separate components that measure i) how people feel and ii) how they think about their lives (Krueger and Stone, 2014). While much of the existing literature on well-being in the workplace focuses on evaluative measures of SWB (such as job satisfaction),[14] we focus in this paper on an affective measure: workers' overall feelings of happiness as they vary from week to week.

A happiness survey instrument was designed following the OECD's guidelines on the measurement of SWB (see OECD, 2013). The validity and reliability of measures such as ours are well informed by decades of academic research on the topic (for a full discussion of the issues surrounding the measurement of SWB, as well as a detailed account of the various

[13]The word "gloomy" is used in the English language to characterize a dark or dim weather (in opposition to bright), and more generally a depressing and unhappy event.

[14]Such measures lend themselves to between-worker comparisons, or within-worker comparisons in infrequent panels such as those using annual survey data from a firm's engagement survey. Though work engagement surveys are sometimes conducted by firms at higher frequencies, they typically capture a variety of psychological dispositions – from extraversion to exhaustion, enthusiasm, hard work or self-efficacy Internally conducted surveys have also been criticized by the management literature as suffering from acquiescence bias, which results in an exaggeration of the extent of engagement (for a review, see Purcell, 2014).

ways in which the validity and reliability of such measures have been tested, see Krueger and Schkade (2008); Krueger and Stone (2014)). Employees were asked "*Overall, how happy did you feel this week?*" on weekly basis for six months, beginning in July 2017.[15] 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 on Thursday afternoon, and is shown in Figure 3-1. We measured happiness toward the end of the week and asked employees about their happiness during that particular week. This is key, since it allows us to link happiness reports and work outcomes over the same corresponding time period. An alternative option would be to ask employees at the start of the week, which would give a clearer temporal ordering between happiness and subsequent outcomes. However, this would come at the expense of not knowing workers' happiness during the period where they are actually working on those sales. Our use of happiness and sales measures that are temporally concurrent has significant benefits for the analysis; however, the contemporaneous nature of our happiness and sales measures underlines the need for an instrumental variables strategy in order to deal with potential issues of endogeneity.

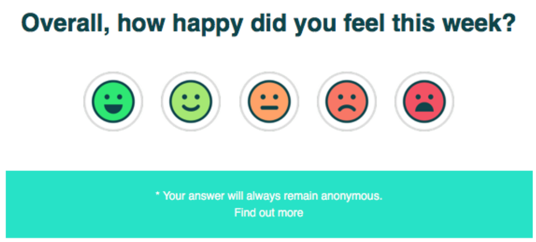
The single-item survey question could be quickly answered within the email. Workers were assured that their individual happiness responses were being collected externally for the sole purpose of academic research, and would not be shared with management.[16] Workers were also offered the opportunity to opt-out of the study at any time, via a simple email click-through.

The happiness responses provide us with a weekly, ordinal measure of affective well-being

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

[16] This is significant given that much existing research in this space uses firm data from internal engagement surveys, where workers are generally prone to acquiescence bias (Purcell, 2014).

Figure 3-1: 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.

or "mood". The distribution of responses is shown in panel (a) of Figure C-1. We use our happiness responses both ordinarily 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, in order to be aligned with the scales used in survey measures in the literature. When doing so, the mean response is around 4 with a within-person standard deviation of 2.4 (a full set of descriptive statistics, at the worker-week level, are shown in Table C.1). Importantly for our identification using individual fixed effect models, panel (c) of Figure C-1 shows that the responses vary significantly within-workers over time during the study, suggesting that we are measuring mood and not simply picking up a more static overall measure of evaluative job satisfaction.

3.2.5 Sample construction and characteristics

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

We observe 1,793 sales workers, distributed across 11 different call centers (for a map of the spatial distribution of these call centers, see Figure 3-2). All of these workers were invited to take part in the study and were sent weekly mood surveys. Participation was high. Indeed, of these employees, 1,438 (around 80%) participated by answering at least one survey over the subsequent 6 months. All of this cohort of workers were sent a weekly email, unless they had since left the organization. We do not follow any workers who subsequently joined the firm after the first week of the study.

Conditional on participating in the study, workers responded to a mean of 10.3 waves (with an SD of 7.1). The weekly response rate of workers who participated was on average around 37%, which rises to 50% if we focus only on workers who did work on Thursday and Friday (recall that the mood survey was sent every Thursday afternoon). In the supplementary material, we assess the extent to which observable demographic characteristics, workplace schedule, performance, and mood-related variables (averaged at the team level for happiness and at the call center level for weather) are able to predict i) participation in the study, on the extensive margin (Table C.2) and ii) the number of survey waves answered if the worker did participate, on the intensive margin (Table C.3). Importantly, neither participation in the study nor frequency of response are significantly related to the average weekly number of sales made by workers during the course of the study, or to the average team happiness and call center characteristics (local weather and window share). Mean hours of selling time is positively predictive of the number of response waves – an extra hour of mean daily sales time is associated with around a 2% increase in the number of waves responded. In other words, the absence of response in certain weeks can be explained by within-worker variation in work schedules, which are set by the firm in advance, such as not working on Thursdays and Fridays.[17]

We drop any participants who responded to only one survey wave, since we rely on within-worker variation over time. Also dropped from the analysis are observations that lead to statistical separation in our individual and week fixed-effect Poisson models. This leaves us

[17]As is common in most call center settings, work schedules are set up in advance by managers and vary within workers over time, so that different workers work on Thursdays or Fridays week-to-week.

with a final sample of 1,157 employees. Summary statistics for this final sample are shown in Table C.1. Around 59% are male, and the modal age category is 26-30 (with over 60% of the sample being between 21 and 35 years old).[18] Mean tenure in the firm is about 5 years, with a large standard deviation of 7 years. Half of the workforce in our sample has been in this position for less than 2 years, and around 7% of the sample experienced turnover during the 6 months – either leaving of their own accord or having their employment terminated.

3.2.6 Non-Response to Survey

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

In Table C.4 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 (weighted 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.[19]

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

[19] 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.

3.3 Empirical Strategy

3.3.1 Baseline Equation

We are interested in whether affect has any causal impact on weekly performance at work. We focus on sales workers so that our preferred performance measure is, at least initially, the worker's total number of weekly sales. We begin by estimating a within-worker productivity equation, such that

$$E[S_{ijt}|A_{ijt}, X_{ijt}] = exp\{\beta A_{ijt} + \gamma X_{ijt} + \nu_i + \tau_t\} \tag{3.1}$$

where S_{ijt} corresponds to the number of weekly sales for worker i in call center j during week t and A_{ijt} is her affect during that same period t. Worker fixed effects ν_i capture any individual-specific characteristic that does not change over time, and τ_t is a time fixed effect partialing out any shocks that may affect both mood and sales. Finally, we include a vector of controls X_{ijt} for two major labor input variables that may vary over time and across workers, namely the (log of the) total number of selling hours during week t, and the fraction of time spent at work in the week on mandatory non-productive activities (internal shrinkage). We adjust the error term to account for two-way clustering on individuals and location-week.[20] We estimate equation (3.1) with a Poisson quasi-maximum likelihood model. The Poisson model is particularly relevant for count data, and also makes the interpretation of β intuitive, as it can be interpreted in terms of a percent change in number of sales.

We begin by using visual exposure to gloomy weather patterns as a proxy for A_{ijt}. We then move on to use a more direct survey measure, which we label H_{ijt}. In line with the extensive SWB literature, and in order to enable easier comparability, we assign a numerical value to each of the responses to our email survey, ranging from 0 to 10. However, we are naturally hesitant to take the scale at face value as a continuous measure. We thus estimate

[20]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.

equation (3.1) using a set of indicator variables for responses (leaving aside the middle response as the omitted category) as a robustness check, as well as using the happiness survey in a continuous fashion in our main models.[21]

3.3.2 Potential Biases in Baseline Equation

When estimating equation (3.1) using the direct survey measure of happiness, there are a number of reasons to be concerned that β 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. In particular, we see two (opposing) major ways in which reverse causality may bias our coefficients. First, more productive workers can get compensated for their higher performance through 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, and conversely, what makes workers happier could lower their productivity, depressing the true β coefficient. One major candidate is the quantity of work itself, as working less may increase happiness. Indeed, 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 downward.

In a customer-facing retail setting such as ours, we expect a strongly downward bias. A long line of work, particularly in sociology, has studied the interactions between workers and customers, and theorized the implications of these interactions on worker feelings and well-being. Research on 'emotional labor,' for example, emphasizes the negative impacts of tasks that involve workers having to enact emotions they are not actually feeling or that force

[21] Making comparisons between groups in terms of self-reported happiness can be problematic. As recently argued by Bond and Lang (2019), the comparison of group happiness requires strong identifying conditions when SWB is used as a dependent variable. In this study, we use reported happiness as an independent variable and rely on within-worker estimates.

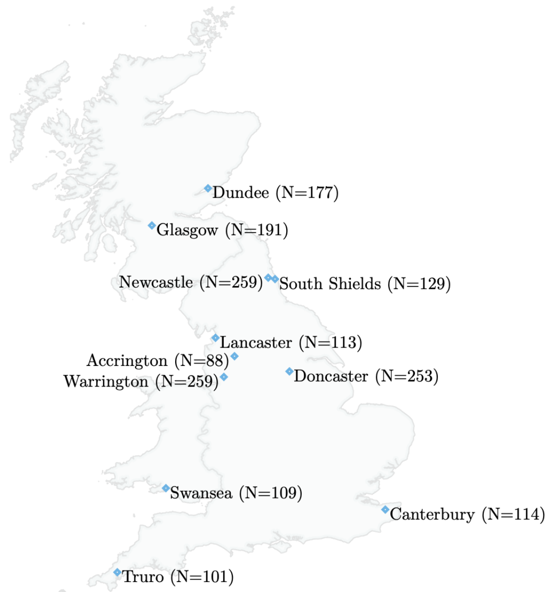
workers to change their emotional states so that they are able to reduce cognitive dissonance (Groth and Grandey, 2012; Hochschild, 1983). This line of work shows that retail jobs, in particular, require more "surface acting" and are, as a result, more emotionally exhausting (Erickson and Wharton, 1997; Hsieh, 2014). In sales call centers, there is a pressure to deal with a constant flow of incoming calls, often initiated by unhappy customers (e.g. recontracting), 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, as noted above, this is a slow moving bonus scheme that makes up only a relatively small amount of total earnings. We thus expect higher work volumes to lead to lower happiness among workers, without any major compensating boost to happiness that might be expected to be provided by a commission-based pay system. We use our data and setting to provide direct evidence on the direction of this bias. In Table C.13, 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 (3.1). 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 (3.1) are thus likely to be strongly biased downward.[22]

3.3.3 IV Strategy

Our main strategy to deal with the endogeneity of self-reported happiness, H_{ijt}, is to rely on an instrumental variable Z_{ijt} that is correlated with the latter but is otherwise independent of sales. In our case, the first stage is:

[22] Note that despite the positive relationship between total number of calls and sales, we still find that happiness improves sales using equation (3.1). 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 3-2: 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.

$$H_{ijt} = \omega Z_{jt} + \Gamma X_{ijt} + \nu_i + \tau_t + \eta_{ijt} \tag{3.2}$$

where Z_{jt} is *visual exposure to gloomy weather*. The standard two-stage least square (2SLS) estimator is appropriate in the case of linear panel models with an instrumental variable. However, when it comes to Poisson models with fixed effects, a fully robust approach is to rely on control function methods. Following Lin and Wooldridge (2019), we estimate the first stage for H_{ijt} via OLS and add the first-stage residuals as a control in equation (3.1). We bootstrap the entire two-stage procedure in order to compute the standard errors. As in equation (3.1), we adjust the standard errors for two-way clustering on individuals as well as on location-week (the level of treatment). In order to do so, we follow the two-way cluster bootstrap procedure outlined in Cameron et al. (2011).[23]

[23] Results are similar when performing a simpler panel bootstrap, clustering on individuals; however, the standard errors are smaller in that instance, and we report here the more conservative two-way clustered errors.

Visual Exposure to Gloomy Weather

For identification, our instrument combines two sources of variation: weather and architecture. We make use of a somewhat unusual combination of factors that are present in our setting. First, we study 11 call centers that are widely dispersed geographically. Our 11 call centers are distributed across three different countries in Great Britain, from the south coast of England to the north of Scotland (see Figure 3-2 for a map). Second, we study a firm based in the United Kingdom, which happens to be a place with very variable weather. This is true across both time and space – even within weeks, locations vary significantly in the weather patterns they experience. Third, the call centers happen to be housed in buildings that are very different in terms of architecture. In particular, they vary significantly in the amount of glass windows they have, ranging from high-rise office building clad with glass to warehouse-style buildings with almost no windows at all.

We are principally interested in the *visual* aspect of poor weather conditions, which is most likely to have purely psychological effects and thus get us closer to true mood shocks. The Gloomy Weather Index is defined by the sum of the daily incidences of fog, rain and snow during any given week. We further interact our index of gloomy weather with the share of windows in the call center. For our preferred instrument, we combine these two pieces of variation into a single variable, which is window-weighted gloomy weather. Specifically,

$$Z_{jt} = Gloom_{jt} \times Windows_j \tag{3.3}$$

where $Gloom_{jt}$ is the index (lying between 0 and 10) of local gloomy weather and $Windows_j$ is the proportion of the call center's wall surface that is covered by glass windows (lying between 0 and 1). We label this visual exposure to gloomy weather. In robustness models, we also disaggregate the weather index in order to look separately at visual exposure to each of the component parts (see Table C.9).

It is worth noting here that the impact of building architecture itself on mood and productivity is captured by worker fixed effects. Identification effectively arises from the *interaction* between gloomy weather across call centers within any given week and workers' ability to observe poor weather while at work. Local exogenous shocks differently affect the mood states of workers depending on whether they are exposed to these shocks or not. Our setting thus differs from popular specifications of combined instruments such as Bartik or shift-share IVs where the same exogenous shock (e.g. national immigrant flows or industry growth rates) applies differently to every observation in the sample.[24] Possible threats to the exclusion restriction are discussed in the next section.

Threats to the Exclusion Restriction

Visual exposure to gloomy weather will only be a valid instrument under the condition that weather has no direct impact on productivity other than through its effect on mood. We discuss a number of main threats to this assumption.

Temperature and Pollution. Rather than visual exposure to gloomy weather having an impact on sales through worker mood, adverse weather could physically be affecting sales 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 et al., 2020).[25] We show in Table C.10 that when controlling for local temperatures, all of our main findings are robust. Moreover, we find no evidence of any reduced form effect of temperature on sales, and though temperature is weakly correlated to mood, the effect is unrelated to visual exposure, so that the weighted instrument has no significant effect on mood (Table C.11). 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, we note that pollution correlates

[24]For a recent discussion on the validity of these approaches, see Borusyak and Hull (2020).

[25]In an empirical strategy somewhat analagous to ours, the authors show null effects in schools that have air conditioning. That is, in both this and our paper, we seek to demonstrate the plausibility of the causal mechanism by being able to "turn it off."

with temperature. There is also little reason to suspect that glass-clad buildings would be more susceptible to air pollution effects than warehouse-style ones.

Sickness. Adverse weather conditions may cause sickness among workers, and impair their ability to work effectively if they attend work while ill.[26] Looking directly in the data, the local gloomy weather index turns out to be unrelated to the local share of workers under sick leave in any given day or week (Table C.7). We also consider the possibility that such a relationship may occur with some lag. Gloomy weather a day (or a week) before remains unrelated to the frequency of sick leaves a day (or week) after. Perhaps more importantly, any sickness argument would apply whether or not the call center had many windows. Given the relationships shown in Figure 3-3, it seems clear that the variation our instrument picks up on is related to psychological rather than physical health.

Opportunity Cost of Leisure. The presence of (in)clement weather outside may change the opportunity cost of leisure time (see Connolly, 2008), leading workers to adjust their labor supply decisions. Relatedly, poor weather could have direct impacts on employee's ability to attend, or arrive on time – for example, if rain affects their ability to commute to work. First, as with sickness or pollution, it is difficult to reconcile the concern with the significant interaction with window coverage: the ability to arrive on time or benefit from weather conditions outside should be similar across call centers with differing window coverage. Though the presence of windows itself may distract workers, there is no reason to expect workers to look out more (vs. less) if the weather is bad (vs. good) outside. Second, it is worth noting that, in our setting, workers have very little discretion over labor supply. Work hours are scheduled by management, and once workers are at their terminal during their scheduled hours they face calls as they come in. Besides, we focus on productivity, so our results are conditional on the number of selling hours done by the employee during

[26]The link between weather conditions and physical sickness is not as clear-cut as it may seem, however, especially for mild symptoms that do not require sick leave. It is generally associated with cold temperatures rather than gloomy weather, and we find null effects of temperature on sales in Table C.11.

the week (as well as other scheduling controls for internal shrinkage). In section 3.4.5, we discuss the possibility of labor supply effects directly – and find null effects of instrumented happiness on various fine-grained measures of labor supply.

Product Demand Effects. One major concern is that gloomy weather may also have a direct impact on customer demand, or, an indirect effect by affecting customers' mood. The inclusion of time fixed effects, τ_t, ensures that our key piece of identifying variation is exposure to weather shocks across call centers j, within any given week t. That is, for identification we rely on variation in weather across call centers within any given week, rather than on movements in national weather conditions from week-to-week. Semi-structured interviews with management of the firm show that calls come from all parts of the country, and are then directed to call centers based on operator availability. Key for our identification is that call centers do not field calls simply because they originate locally (in the United Kingdom, there is no distinction between short- and long-distance phone calls). Thus *local* weather in the vicinity of the focal call center should be independent from customer demand. In addition, we provide direct evidence that customer demand at a particular call center is not affected by local weather. We look at the average number of incoming calls per worker and the extent from which workers may be asked to answer more calls faster as a result of local demand pressures. We show in Table C.8 that none of these variables are affected by daily weather that is local to that particular call center.[27]

Sorting Effects. It is assumed throughout the chapter that no other factors correlated to the share of windows but unrelated to visual exposure itself may explain heterogeneous sensitivity to gloomy 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

[27]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 share. 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 C.9).

people), and if those same workers tend to be systematically working in call centers with more windows.[28] Of course, one should only be worried about such sorting effects if there exists important sources of heterogeneity in worker sensitivity to gloomy weather in the first place. We investigate this possibility directly, looking at whether gloomy weather (unweighted by windows) 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 leaves. In Table C.6, we find no evidence of heterogeneity across any of these dimensions, which reduces the concern that the share of windows would capture higher sensitivity to gloomy weather other than through visual exposure to weather itself.

3.4 Results

3.4.1 Reduced Form Estimates

In column (1) of Table 3.1 Panel A, we estimate the reduced form effects of gloomy weather on sales performance without accounting for whether or not workers are likely to be visually exposed to gloomy weather while at work. We find a negative, though imprecisely estimated coefficient. In column (2), we interact this coefficient with the share of the call center's walls that are covered in glass windows. We find that the negative effect on sales is much stronger in the situation that weather is *visible*. To aid interpretation here, window coverage is z-scored. Thus the main effect of -0.0067 shows the impact of a 1-point increase in the gloomy weather index on sales, at the mean of window coverage. The interaction term suggests that the magnitude of the effect is more than double in call centers that are one standard deviation above the mean in terms of window coverage. Column (3) shows the estimated impact of visual exposure to gloomy weather more succinctly. Rather than add additional

[28]Note that the concern here relates to heterogeneous *sensitivity* to weather, not the fact weather itself may correlates with unmeasured worker characteristics, which are captured by worker fixed effects.

Table 3.1: Impact of Visual Exposure to Gloomy Weather on Sales and Happiness

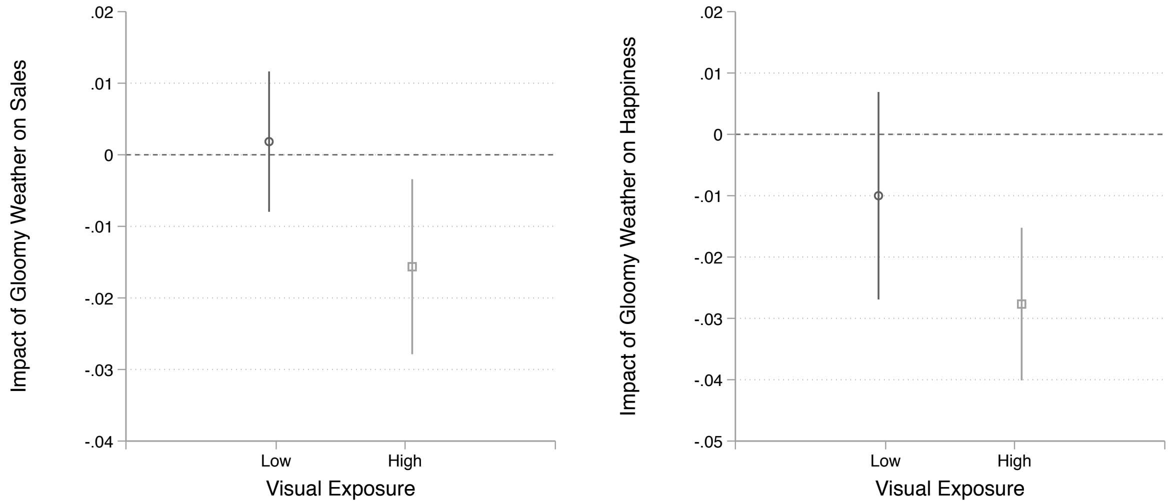
	(1)	(2)	(3)
Panel A: Reduced Form (Dep. Variable: Sales)			
Gloomy Weather	-0.0059	-0.0067*	
	(0.0040)	(0.0040)	
Gloomy Weather × (z-scored) Window Coverage		-0.0095*	
		(0.0054)	
Gloomy Weather (weighted by window coverage)			-0.0413**
			(0.0187)
Observations	12,282	12,282	12,282
Panel B: First Stage (Dep. Variable: Happiness)			
Gloomy Weather	-0.0617***	-0.0592***	
	(0.0204)	(0.0200)	
Gloomy Weather × (z-scored) Window Coverage		-0.0511***	
		(0.0184)	
Gloomy Weather (weighted by window coverage)			-0.3122***
			(0.0654)
Observations	12,282	12,282	12,282
F-Stat of IV(s)	9.17	12.07	22.82

Notes: Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. Panel A: Poisson-FE models reported with weekly sales as dependent variable. Panel B: OLS-FE models reported with weekly happiness as dependent variable. All models include individual and week fixed effects, work schedule controls, and dummies for day of week of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

interactions into the equation, we take the gloomy weather index and re-weight it by the proportion of the call center's walls that are covered by glass windows.

To illustrate what these results mean in terms of magnitude, we regress our gloomy 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%). The results are shown in Figure 3-3a. While the point estimate on sales is close to zero in call centers with very few windows, the impact of gloomy weather is negative in centers with many windows. Here in the above median share group, every 1-point increase on the gloomy weather scale (which varies between 0 and 10, and has a standard deviation of 1.36) lowers weekly sales by 1.5%.

Figure 3-3: Impact of Gloomy 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.

3.4.2 First Stage Impact

While we have argued that visual exposure to gloomy weather satisfies the exclusion restriction, in order to be valid as an instrumental variable, it must also have a sufficiently strong impact on happiness in the first stage. In Panel B of Table 3.1, we regress happiness on our index of gloomy weather conditions. In Column (1), we find that gloomier weather significantly depresses mood. However, once we interact gloomy weather with the share of the call center's walls that are covered in glass windows, we find in column (2) that the effect is much stronger in the situation that weather is visible. The main effect of -0.059 shows the impact of weather on happiness, at the mean of window coverage. The interaction term of -0.051 suggests that the magnitude of the effect is almost double in call centers that are one standard deviation above the mean in terms of window coverage. (The main effect of window coverage is subsumed within the individual fixed effects, and is thus not estimated.) This means two things. First, by using visual exposure to gloomy weather, we are effectively relying on variation in affective well-being rather than any physical effect of weather. Second, we provide evidence that the source of the mood shock is occurring while at work.

Another way to see the contingent relationship of weather on happiness is shown in Figure 3-3b. Here, rather than interact gloomy weather with the share of windows directly, we instead look at the relationship between gloomy weather and happiness between two equally-sized groups of workers: those who work in call centers with few windows (below median window share) and those who work in call centers with many windows (above median window share). As can be seen, similar to what we found for sales, the impact of gloomy weather on happiness is much more negative in call centers where workers can see outside while they are at work. In column (3) of Table 3.1 we estimate the first stage of our IV approach, using our preferred instrument, which we use in our main IV models. In this case, we simply take the gloomy weather index and multiply it by the proportion of windows. Within-workers over time there is a clear negative relationship between adverse window-weighted weather conditions and happiness. An F-statistic from this linear first stage of around 23 suggests the instrument is sufficiently powerful to be valid. It is worth re-iterating here how much we are asking of the data, given that we include time fixed effects in the equation: the coefficient here identifies the effect of visual exposure to weather on happiness across call centers within weeks.

3.4.3 IV Estimates

Our prior examination of the reduced form and first stage equations already provided strong insights on the causal effect of mood on productivity. It also motivates the use of visual exposure to gloomy weather as an instrumental variable, since the IV estimate logically arises from the ratio between the reduced form and first stage estimates presented in models (3) and (6) of Table 3.1.

Table 3.2: Impact of Happiness on Sales Performance

	Sales	
	(1)	(2)
	Non-IV	Poisson-IV
Happiness	0.0141***	0.1331**
	(0.0014)	(0.0619)
Observations	12,282	12,282
1st Stage F-Stat		22.82

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. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

We put all of this together in Table 3.2. The resulting IV estimate and confidence intervals, shown in column (2) of Table 3.2, suggest a strong causal effect of happiness on sales performance. 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. This estimate has a bootstrapped 95% confidence interval lying between .0111 and .2550. We estimate an average marginal effect of 3.3600 additional sales (with a bootstrapped standard error of 1.5413 and 95% confidence interval of .3392 to 6.3809).

The non-IV estimate is much smaller than the quasi-experimental one (though we note here that the non-instrumented estimate falls (just) within the bootstrapped confidence intervals of the IV estimate). We return in more detail to this point in our extended discussion of magnitudes below.

3.4.4 Robustness

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 dummies for different levels of happiness in our main non-IV model. Figure C-6 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).

Analysis using Daily Data. Although our happiness data is measured at the weekly level, our productivity data is reported largely at the daily level. Here we assume that responses to the happiness question, which is asked on Thursday and refers specifically to “this week”, apply equally to each of the weekdays. We again estimate Poisson-IV regressions predicting the number of daily sales, and include a set of individual and date (that is, day*week*year) fixed effects as well as daily controls for selling time and internal shrinkage. Results are very similar when we carry out this exercise (see Table C.16). In table C.17 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 visual exposure to gloomy weather has a negative effect on sales performance. One concern with our findings using the happiness survey is the threat of bias arising from non-response, which we discussed above. Here we are able to make a further check, and show that the reduced-form effect of weather exposure on sales is very similar for worker-days on which we have non-missing and missing happiness data.

Effects on Customer Satisfaction. Even if happier workers are able to extract more sales from customers, this may come at the cost of lower customer satisfaction (and potentially fewer customers in the long-run). Taking customer satisfaction as a proxy for customer surplus, a lower surplus (hence satisfaction) may occur if customers are “forced" to buy. Assuming customers are rational, however, they should only accept offers whose marginal cost does not exceed their marginal utility. To investigate this issue further, we leverage data on customer satisfaction. After each call customers are asked by text or phone to report the extent to which they would recommend BT to others, on a 0 to 10 scale. This measure is somewhat noisy, since each worker on average receives very few feedback ratings per week[29], and suffer from the possibility of selection bias in whether happy or unhappy customers are more or less likely to answer. We thus see these analyses as suggestive. Using the weekly average response for each worker, we show in Table C.12 that, within-workers over time, visual exposure to gloomy weather has a positive but non significant effect on customer satisfaction. We also find little effect of instrumented employee happiness on customer satisfaction.

First Stage Functional Form. In Figure C-4, we show a graphical representation of our first stage regression of happiness on visual exposure to gloomy weather. As can be seen, the relationship looks roughly linear. However, there is some suggestive visual evidence that exposure to particularly gloomy weather may have a stronger (negative) effect on happiness.[30] In order to explore and account for this more fully, we test alternative functional forms for our instrumental variable. In Table C.5 we use the squared value of the exposure to poor weather index as well as a dummy for bad weather (equal to 1 if the index is seven or above). The resultant second-stage coefficients are slightly smaller in magnitude but still within the confidence interval of the IV estimate initially reported above. Lastly, any concerns related to weak instruments should be mitigated by the reduced form evidence provided in this chapter. 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 only identifies a local average treatment effect (LATE) – that is, the effect driven by those whose mood can be most easily manipulated, or in our case, most sensitive

[29] We take only instances where the worker receives two or more in a given week.

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

to visual exposure to gloomy weather.[31] 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 Keueger, 1991), in our case, mood movements are largely "external" to an individual in the sense that one does not have a direct control over them. In other words, heterogeneous treatment effects arising from employees' selection on the productivity gains of good mood are unlikely to occur in our context. Table C.14 shows the first stage of our IV strategy this time interacting *visual* exposure to gloomy weather with each of the six main characteristics described earlier. Again, we find no evidence of heterogeneity across any of these dimensions. This does not preclude the possibility that our estimates arise from a particular sub-population of less resilient "compliers", but we have no ability to identify them in the data.

3.4.5 Evidence on Channels

Having shown a robust causal effect of happiness on overall sales performance, we move on in this section to what behaviors may drive this productivity effect. In order to discuss potential channels, however, it is first necessary to examine what behaviors drive sales in the first place. We see here three main possibilities: the number of calls answered per hour (or speed), the ratio of sales to calls (or call conversion rate, also a measure of quality), and workers' faithfulness to their computer-scheduled workflow (adherence).

First, 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. Even though the firm sets a loose target of 91% adherence each week,

[31] The LATE itself may be 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.

occasional deviance from this workflow may be beneficial if the worker has to stay on a call to complete a sale, for example. We code our outcome variable here as 1 if this target is met. Table C.18 shows that, conditional on the total number of hours spent at work (selling or doing other internally scheduled non-productive activities), adherence appears to have no significant effect on sales.

Second, we observe on a daily basis the total number of minutes spent on incoming calls as well as the number of calls taken. We code the average number of calls per hour during the week. This "speed" measure is what would typically be used as a labor productivity metric in the manufacturing industry. However, in a call center setting, and in the service industry more generally, it is not at all clear that taking more, shorter calls will be beneficial, when the goal is selling (hence the use of overall sales as our main performance metric). 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 C.18 confirms the total number of calls per hour is not a good predictor of productivity. If anything, it is associated with a reduction in the number of weekly sales per worker.

This suggests that any positive productivity effect on sales will likely arise out of a higher ability to convert calls into sales, rather than working faster or adhering more closely to workflow scheduling. We now consider how workers' mood affects each of these three channels. In Table 3.3 Panel A, we present reduced form evidence using our measure of visual exposure to gloomy 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 gloomy weather. Both the reduced form and the second stage give consistent results. First, we find in column (1) that happier workers adhere more closely to the workflow that has been set out for them. Second, and despite a negative link between sales and the number of calls per hour, we show in column (2) that in happier weeks workers do work faster. Finally, we look

Table 3.3: Intensive Margin: Happiness and High-Frequency Labor Productivity

	Adherence (Met Target=1)	Calls Per Hour (Log)	Conversion Rate (Log)	
	(1)	(2)	(3)	(4)
Panel A: Reduced Form				
Exposure to Gloomy Weather	-0.0247**	-0.0191***	-0.0903***	-0.1041***
	(0.0111)	(0.0060)	(0.0265)	(0.0279)
Adherence (Met Target=1)				0.0084
				(0.0102)
Calls per hour (ln)				-0.8782***
				(0.0711)
Observations	12,169	12,100	11,720	11,672
Panel B: 2SLS				
Happiness	0.0822**	0.0632***	0.2832***	0.3254***
	(0.0392)	(0.0241)	(0.1012)	(0.1087)
Adherence (Met Target=1)				-0.0628*
				(0.0360)
Calls per hour (ln)				-0.5623***
				(0.1456)
Observations	12,169	12,100	11,720	11,672
1st Stage F-Stat	19.83	20.87	21.51	19.93

Notes: Second-stage 2SLS models reported, using visual exposure to gloomy 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. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

at the ratio of sales to incoming calls. Column (3) shows that in happier weeks, workers convert more of their calls to sales.

Which of these three channels is most likely to explain our main productivity effect? The point estimate for adherence and calls per hour are relatively small. Though a 1-point increase in happiness is equivalent to a rise in the average number of calls per hour from 5 to 5.3 calls, speed seems unlikely to explain the large increase in sales. As we showed above, if anything, faster calls are negatively related to sales. This suggests that the main productivity effect on sales is occurring through the conversion rate channel. This 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. This further supports the benefit of using sales as a non-ambiguous

Table 3.4: Extensive Margin: 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					
Exposure to Gloomy Weather	0.0297	0.0084	0.0106	0.0017	0.0167
	(0.0213)	(0.0166)	(0.0069)	(0.0140)	(0.0123)
Observations	12,282	12,279	12,282	12,282	12,282
Panel B: 2SLS					
Happiness	-0.0958	-0.0270	-0.0342	-0.0055	-0.0539
	(0.0714)	(0.0546)	(0.0220)	(0.0454)	(0.0419)
Observations	12,282	12,279	12,282	12,282	12,282
1st Stage F-Stat	21.90	21.87	21.90	21.90	21.90

Notes: Second-stage 2SLS models reported, using visual exposure to gloomy 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 dummies for day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

measure of productivity. Besides, none of these other two channels seem to come at a quality cost in terms of affecting workers' ability to convert more calls to sales.

Though this chapter's main focus is on labor productivity, we also investigate the impact of positive mood on labor supply decisions. 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 3.4, we find no robust evidence of any happiness effects on the amount of time spent selling. We also observe additional high-frequency data related to workers' overall labor supply. 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.[32] Here too, we find no robust evidence of any significant mood effects on attendance. This rules out the possibility that happier workers would choose to attend work more often, for instance due to higher levels of motivation, or that good weather may somehow affect the opportunity cost of leisure. Equally, we find negative but non-significant coefficients for over-time working and paid vacation. We also observe the number and length of breaks taken by workers. Here we code the overall number of hours taken during the week. The

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

coefficient is negative (though imprecisely estimated), effectively ruling out any large effect whereby happier workers take longer breaks to chat with their co-workers (what we might think of as "distraction" effects).

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 interpret this evidence as strongly suggesting that happiness does not affect labor supply decisions in general. As we noted above in relation to the validity of our weather instrument, the very limited labor supply flexibility in our call center field site effectively provides us with an ideal setting in which to test for the pure productivity effects.

3.5 Discussion

A long-running literature has sought to explain heterogeneity in productivity across individuals as well as firms. We contribute to this growing body of work by showing the effect of a typically-overlooked aspect of workers' lives: how happy they feel while at work. We follow the existing literature on 'mood effects in the field' by presenting reduced-form evidence of the effects of gloomy 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 access to windows. Second, we demonstrate using survey data that visual exposure to gloomy 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 suggesting a strong causal effect of affect on productivity.

3.5.1 Psychological Mechanisms

While the main focus of the paper has been to assess the extent to which there is a main effect of happiness on productivity, our data and setting nevertheless allow us to discuss some suggestive evidence on mechanisms. One may think of two broad psychological pathways through which happiness may feed through into higher sales: cognitive processing and emotional or social skills.

First, better mood states may improve cognitive functioning. Indeed, positive mood states have been shown to lead to better cognitively processing (Estrada et al., 1997) as well as more creative problem solving (Amabile et al., 2005; Isen et al., 1987).[33] Relatedly, it has been shown that the thoughts of unhappier people are more likely to "wander" (Killingsworth and Gilbert, 2010), a mechanism that has been formalized into an economic model in which happiness reduces the amount of time spent worrying about negative aspects of people's lives, and thus drives productivity (Oswald et al., 2015). Though we cannot measure cognition directly, we already saw that a worker's adherence, a proxy arguably capturing the capacity to stay focused on a given task, increases with happiness (see Table 3.3). In our setting, however, adherence happens to have little influence on sales; however, in other settings with different types of work, this mechanism could well be more consequential.

Positive mood could also augment (non-cognitive) socio-emotional skills. For example, happy workers have been shown to demonstrate higher powers of self-control, empathy (or sociability), and abilities to manage their emotions, which is particularly useful in customer-facing industries and occupations (Tice et al., 2001). On the other hand, it has been argued that sociability may reduce work speed by raising workers' vulnerability to distractions (Coviello et al., 2018). In our particular setting, where team work is absent and workers are

[33]However, for a natural experiment, using deaths of friends and relatives of artists, suggesting null effects, see Graddy and Lieberman (2018).

heavily monitored, we find no evidence suggesting sociability may be detrimental to productivity. On the contrary, happier workers take calls slightly faster and we find no effect of happiness on break-taking. 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 in Table C.19, we find that, although all of our estimates are less precise, the magnitude is close to zero for regular order-taking. This is consistent with the main mechanism being call-to-sales conversion rather than working faster, since line sales are largely mechanical order-taking. A more strongly positive effect is evident for TV and cell-phone contracts, where there is more leeway for persuasion and negotiation. A even larger positive effect is evident for the re-contracting sales, which usually involve the largest amount of social and negotiating skills (and also typically involve disgruntled customers).

Lastly, the sociological literature on emotional labor (see, e.g. Hochschild, 1983) has long argued that in tasks involving interactions with customers, it becomes particularly costly for unhappy employees to leverage their social skills and manage their emotions, as they need to “fake" happiness. This is particularly the case when dealing with unsatisfied customers, as employees' ability to empathize and “tame" customers' negative emotions then comes at a larger psychological cost. Consistent with that hypothesis, we find that our effect is stronger during weeks where average customers are the most unsatisfied (versus high). Table C.20 replicates our main analysis by terciles of weekly customer satisfaction. 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).

3.5.2 Magnitude of Effects

Depending on specification, we find an IV point estimate for the effect of a one unit increase in happiness on sales in our IV analysis of somewhere between .08 to .15 log points. There are a number of different ways to think about the size of this effect. First, we can think in terms of the absolute number of sales our estimates correspond to. As noted above, with our main specification, 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), with a relatively large bootstrapped confidence interval of around half a sale to a little over 6 additional sales.

Second, 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 C-5, we show that workers who have at least 1 year of experience on the job (compared with those with less than 1 year of tenure) generally make around 24% more sales per week. Equally, those who are paid in a more strongly incentivized way (a point we return to below) also make around 23% more sales per week. Male workers and workers under 40 in general make around 10% more sales per week on average.

Third, we can compare our instrumented estimate with the non-instrumented one. The non-IV within-person partial correlation estimate is much smaller. We show below that the IV estimate is roughly in line with what one would expect from the experimental (i.e. causal) literature on affect and productivity. The non-IV equation is likely to be biased, which is what motivates the need for natural experimental evidence. In particular, the lower within-person estimate is in line with the direct evidence presented above on the strong negative relationship between work volume and happiness, as well as with the presence of measurement error in our main right-hand side variable. That is, the within-person estimate is likely to be strongly downward biased by both of these factors. In addition, while our non-instrumented estimate is relatively precisely estimated, our IV estimates are much less so, which implies there is still the possibility that it could be smaller.

3.5.3 Comparisons with Existing Literature

One place to begin any comparison with previous studies is prior controlled laboratory experiments. Here we focus on the seminal paper of Oswald et al. (2015), who manipulate happiness in the laboratory and subsequently have subjects do a piece-rate productivity task. While this is clearly a very different context, call center work is not entirely dissimilar, in that it involves a series of individual tasks with very little teamwork. In this experimental set-up, a short-run one standard deviation increase in happiness (induced by viewing a comedy as opposed to a placebo video clip) causes participants to correctly do around 29 to 35 percent more incentivized additions (See Appendix C.8 for further detail). In our setting, a one standard deviation increase in happiness (equivalent to a 2.4 point increase on our scale), leads to around a 30% increase in sales.

We can also 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. Though they are primarily interested in the impact of working-from-home on the performance outcomes of non-sales workers (typically the total number of calls), they also assess how the new policy impacted workers' affective states. 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. The absence of a speed-quality trade-off when it comes to the effect of mood on performance is consistent with prior work by Miner and Glomb (2010). Though they cannot claim causality, they find suggestive evidence supportive of the same pattern. However, the closest paper in nature to ours is Coviello et al. (2018), who study a sample of largely non-sales call center workers in the USA, and instrument for worker engagement—on a 1-to-5 scale corresponding to feeling 'unstoppable', 'good', 'so so', 'exhausted', and 'frustrated'—using weather patterns. While we focus on the productivity of sales workers, the authors focus more clearly on short-run labor supply decisions (i.e. percentage of unproductive time at work) and work speed (i.e. number of calls per hour). The study focuses on customer service representatives (CSR), for whom the task is not to sell but rather to advise and inform customers. 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,[34] 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, further research is required in order to understand the issue of when and where productivity effects of happiness are likely to be stronger or weaker – for example, in different industries and occupations.

3.5.4 Managerial Implications

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

[34] Besides their focus on the labor supply decisions of non-sales workers, the authors rely on a work engagement survey conducted by the firm rather than externally. Further, the geographical dispersion of call centers does not allow for the inclusion of time fixed effects, which we argue is critical to identify mood effects in a two-sided market. There is also the possibility that the performance gains from positive mood may play a stronger role at lower levels of happiness: while a significant fraction of workers in our sample report feeling unhappy, 70% of workers in Coviello et al. (2018) report feeling "good" or "unstoppable".

"100 Best Companies To Work For In America" outperform industry benchmarks in terms of long-run stock returns (Edmans, 2011), as well as a positive link between aggregate firm-level employee well-being and financial performance (Böckerman and Ilmakunnas, 2012; Bryson et al., 2017).[35]

A natural question is the extent to which the local average treatment effect of weather-induced unhappiness on productivity is a useful (i.e. policy-relevant) parameter to estimate. 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. Oswald et al., 2015) is an important step forwards in the literature. In other words, while it may be true that aspects of our results are managerially relevant in terms of natural lighting within workplaces, it is important to note that we employ variation in weather and proximity to windows here on a more fundamental level as an econometric device in order to isolate the causal effect of happiness on productivity. Our study is thus best thought of as a form of basic research, but in an applied setting. We rely on a mood shock that occurs while employees are at work in order to isolate and estimate the causal effect of happiness on performance.

This leads to the inevitable question of whether this weather-induced causal effect is reflective of a more generalizable causal effect. Here we note our use of an overall measure of affective well-being, the week-to-week variation of which we refer to as mood. It is worth recalling that within the broad category of affective measures of SWB, one can distinguish between moods and emotions (see, e.g. Frijda, 1986). In the psychological literature, 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, one is often not aware of the source of a good or bad mood.

[35] In a related literature, work has shown performance impacts of aggregate eudaimonic well-being (see, e.g., Chadi et al., 2017; Gartenberg et al., 2019).

Given this, there seems 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 the sport events[36] to line manager behavior. A great deal of research – increasingly using field experiments – has shown workplace mood and happiness can be influenced by a range of management practices (Bloom et al., 2014; Breza et al., 2018; **?**). 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 (see HBR Analytical Services, 2020), 95% believed that they have some or a high degree of control when it comes to influencing the happiness of their employees (and 19% of firms say they have a workforce happiness strategy in place).[37]

Recent work by Bryson and MacKerron (2016), using high-frequency data on people's mood states in the UK, suggests that doing paid work is ranked almost the lowest in terms of happiness out of 40 activities individuals can report engaging in (only being sick in bed is associated with more unhappiness). In the USA, detailed time use surveys suggest that the most unhappy periods of people's days are when they are with their boss (Krueger et al., 2009). There thus seems to be considerable room for improvement in the happiness of employees while they are at work. While this clearly is in the interest of workers themselves, the analysis presented here suggests it may also be in the interests of their employers. Moreover, it is now well-accepted that the number of jobs requiring workers to interact socially is increasing rapidly (Deming, 2017). Our analysis, which points to a strong role for happiness when workers have to deal with unsatisfied customers, suggests that the importance of employee happiness in driving productivity growth is likely to rise in the coming years.

[36]We considered using sport events as an alternative to visual exposure to gloomy weather. However, it is not clear which team any particular worker would support as most workers in our setting are located in small towns between major cities.

[37]As discussed earlier, the LATE coefficient may be the relevant "targeted" population policy estimate, as it captures the effect of happiness on productivity conditional on workers' mood being reactive to shocks.

3.5.5 Generalizability

A final issue relates to how generalizable our findings are.[38] First, it is worth noting that in terms of selection, our data covers the universe of sales workers at British Telecom, one of the largest private employers in the UK. About 80% participated in the survey, and we have administrative firm records for all workers, including non-survey participants. Neither productivity nor happiness or gloomy weather predicted participation to the study. In terms of attrition, we show that the absence of response in a given week is unrelated to gloomy weather or workers' productivity or happiness, but to variation in work schedules, which are set by the firm in advance. Considering naturalness of the choice task, setting, and time frame, we rely on a natural field experiment in a firm context where participants have to perform a clearly defined real-life productivity task. We also study daily and weekly sales performance for a period of 6-months, which covers three seasons from summer to winter. Lastly, in terms of scaling and insights, our results suggest that being in a positive mood state increases both cognitive and socio-emotional skills in the workplace, which is likely to apply to other firm contexts where (a) workers have clearly defined and autonomous targets, (b) interactions with other co-workers are somewhat limited, but (c) interactions with customers are important (e.g. the whole of the service industry in general).

3.6 Conclusion

A growing number of firms claim to care about employee well-being – with some high-profile examples, such as Google and Zappos, going so far as to appoint a "Chief Happiness Officer" (Knowles, 2015). At least one motivation for this is the expectation that happier employees will be more productive. But while this belief in the happy-productive worker may be held widely by many outside of academia (Fisher, 2003), a long history of research across various disciplines has found only very modest evidence, if any, of any such empirical relationship.

[38]Here we follow List (2020) in discussed the so-called SANS (selection-attrition-naturalness-scale) conditions.

As a result, the link between happiness and productivity has been consigned by many in the research community to the category of "management mythology." Although the long-running stream of research on job satisfaction has produced mixed findings of a relationship with job performance, there has been a more recent turn away from satisfaction measures and onto workers' affective states. This research has produced more robust causal findings – at least in laboratory settings where happiness can be manipulated in a controlled manner (e.g. Oswald et al., 2015). In this chapter, we leverage variation in exposure to local weather conditions, together with a novel survey and detailed quantitative data on workplace behaviors and performance, and find a strong positive impact of employee happiness on productivity in a field setting.[39]

[39]We are very 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.

Appendix A

Supplementary Materials for Chapter 1

A.1 Survey Question Wordings

Users on the site are invited to answer questions on the workplace well-being of their workplace. The wording of these questions is as follows. See `https://www.indeed.com/about/happiness` for more details. All questions ask jobseekers the extent to which they agree with the statement from strongly disagree to strongly agree.

- *I feel happy at work most of the time.*
- *My work has a clear sense of purpose.*
- *I am paid fairly for my work.*
- *There are people at work who give me support and encouragement.*
- *There are people at work who appreciate me as a person.*
- *I can trust people in my company.*
- *I feel a sense of belonging in my company.*
- *My manager helps me succeed.*
- *My work environment feels inclusive and respectful of all people.*

- *My work has the time and location flexibility I need.*
- *In most of my work tasks, I feel energized.*
- *I am achieving most of my goals at work.*
- *I often learn something at work.*

A.2 Balance and Summary Statistics

Table A.1: Balance Between Control and Treatment job seekers

	(1) Control		(2) Treated		(3) Total		T-test P-value
Variable	N	Mean/SD	N	Mean/SD	N	Mean/SD	(1)-(2)
Cookie Age	1167000	125.229 (278.498)	22204762	125.512 (278.638)	23371762	125.498 (278.631)	0.285
Registered User	1167000	0.472 (0.497)	22204762	0.473 (0.497)	23371762	0.473 (0.497)	0.113
Desktop User	1167000	0.626 (0.484)	22204762	0.625 (0.484)	23371762	0.625 (0.484)	0.057
CZ: Unempoyment Rate	1129172	0.083 (0.033)	21487432	0.083 (0.033)	22616604	0.083 (0.033)	0.344
User Has Resume	1167000	0.255 (0.436)	22204762	0.256 (0.436)	23371762	0.256 (0.436)	0.239
Number of Jobs on Resume	298067	3.534 (3.041)	5682210	3.533 (3.037)	5980277	3.533 (3.037)	0.810
Total Experience (months)	298067	95.405 (101.438)	5682210	95.123 (101.302)	5980277	95.137 (101.309)	0.137
Year Entered Labor Force	224743	2009.480 (8.668)	4282258	2009.508 (8.643)	4507001	2009.507 (8.645)	0.122
User is Employed	298067	0.473 (0.499)	5682210	0.473 (0.499)	5980277	0.473 (0.499)	0.967
User Wants Full-Time Job	301891	0.518 (0.500)	5755899	0.518 (0.500)	6057790	0.518 (0.500)	0.448
User has BA or higher	469519	0.402 (0.490)	8951705	0.401 (0.490)	9421224	0.401 (0.490)	0.412

Notes: Each observation is a job seeker.

Table A.2: Balance Between Control and Treatment Jobseeker–Company Observations

	(1) Control		(2) Treated		(3) Total		T-test P-value
Variable	N	Mean/SD	N	Mean/SD	N	Mean/SD	(1)-(2)
Applied = 100	1865835	20.049 (40.037)	35507316	19.775 (39.830)	37373151	19.788 (39.840)	0.000
Happiness Score	1865835	62.769 (8.836)	35507316	62.764 (8.833)	37373151	62.764 (8.833)	0.491
Number of Happiness Surveys	1865835	1278.794 (3723.817)	35507316	1281.649 (3724.995)	37373151	1281.506 (3724.936)	0.308
Company is Fortune 500	1865104	0.157 (0.364)	35493558	0.157 (0.364)	37358662	0.157 (0.364)	0.776
Company is Staffing Agency	1865104	0.014 (0.115)	35493558	0.013 (0.115)	37358662	0.013 (0.115)	0.925
Number of Jobs Listed	1865098	6068.453 (32154.709)	35493464	6084.784 (32099.241)	37358562	6083.969 (32102.012)	0.498
Company Star Rating	1865026	3.521 (0.492)	35492273	3.520 (0.492)	37357299	3.520 (0.492)	0.177
Number of Reviews	1865026	6300.275 (19105.253)	35492273	6317.131 (19134.965)	37357299	6316.289 (19133.483)	0.241
Number of Employees	1865835	5847.607 (9663.622)	35507316	5845.556 (9659.707)	37373151	5845.659 (9659.902)	0.777

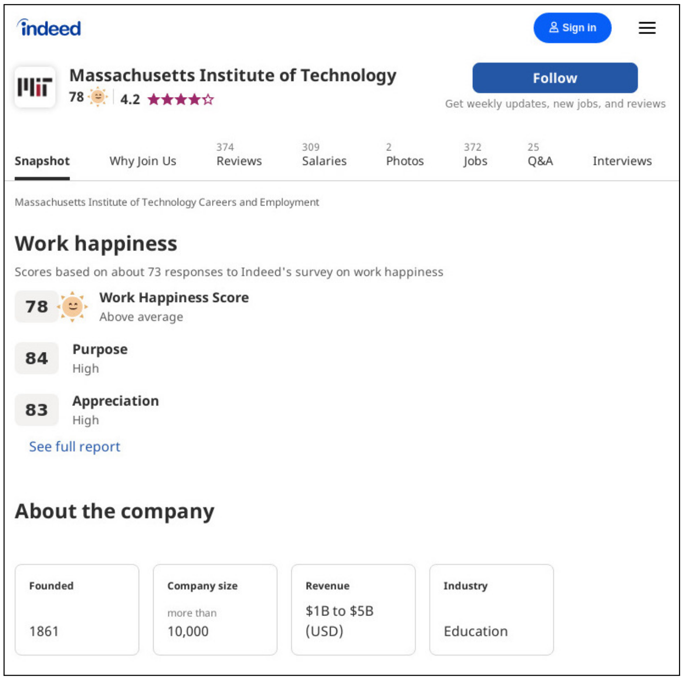
Notes: Each observation is a jobseeker–company pair on the first day that the job seeker visits that company's page.

Figure A-1: Company Happiness Score Variation Within Industries

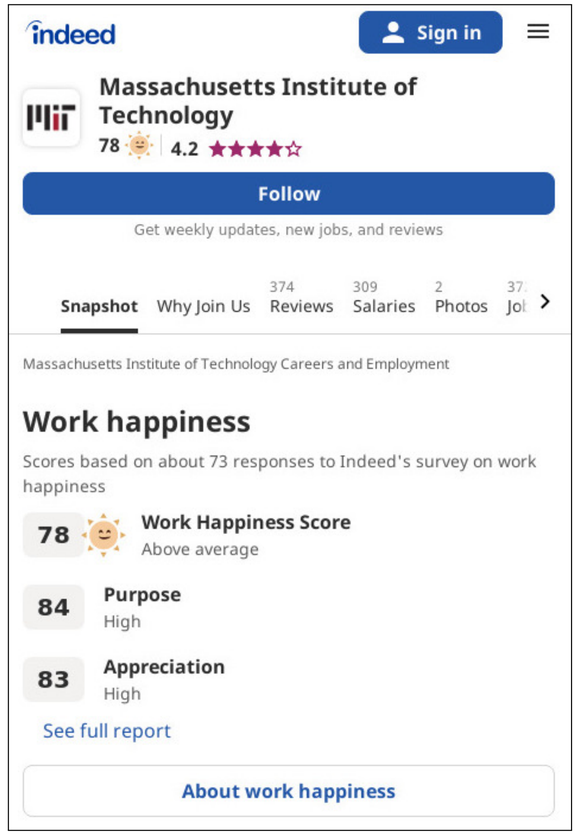
Note: Plotted are the residuals from a regression of company-level happiness score on the final day of the experiment on a set of industry fixed effects.

A.3 Treatment Conditions on Different Device Types

Figure A-2: Treatment Screenshots on Other Devices



(a) Tablet

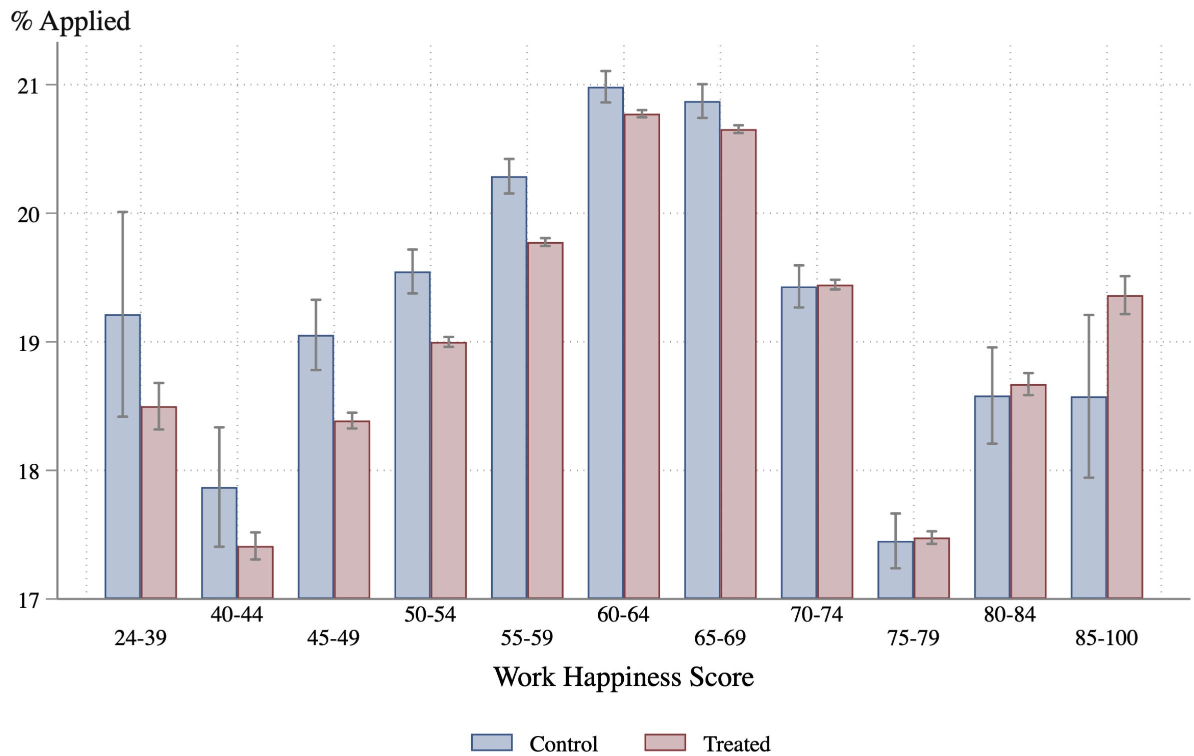


(b) Phone

Note: Screenshots in main text are for desktop computers. Shown here is what the treatment looks like on a mobile phone and on a tablet.

A.4 Sensitivity to Alternative Specifications

Figure A-3: Non-Parametric Estimates



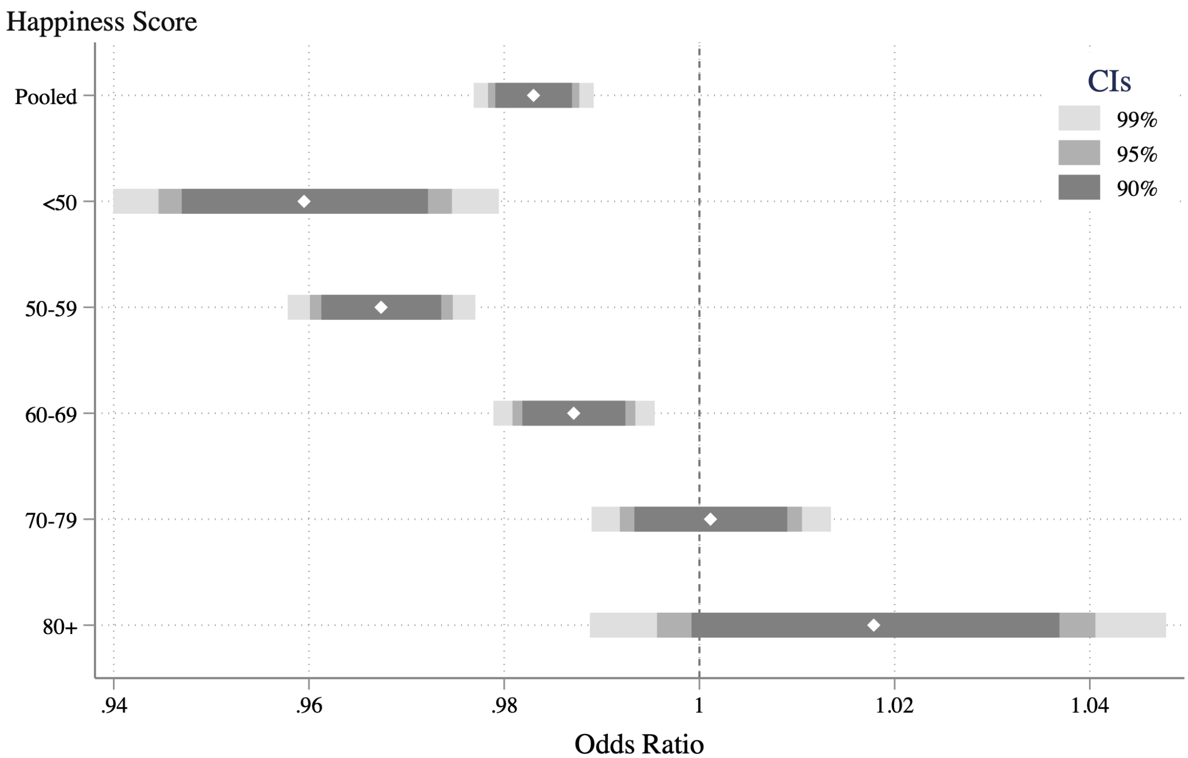
Note: Plotted are the simple means of applied $\times$ 100 *across treatment and control, at different levels of work happiness score displayed.*

Table A.3: Logistic Regression Estimates

	(1) All	(2) 20-49	(3) 50-59	(4) 60-69	(5) 70-79	(6) 80-100
Treated	-0.017***	-0.041***	-0.033***	-0.013***	0.001	0.018
	(0.002)	(0.008)	(0.004)	(0.003)	(0.005)	(0.011)
Observations	37363079	2306602	11059770	15951318	6924676	1120713
Log-Likelihood	-18586738.3	-1094469.1	-5458399.7	-8140686.9	-3341673.1	-542246.8

Notes: Robust standard errors are in parentheses, adjusted for clustering on job seekers. Outcome in all models is equal to 1 if the job seeker applied to a job at that company, 0 otherwise. Logistic regression coefficients are reported. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure A-4: Logistic Regression Estimates: Odds Ratios



Note: Plotted are the exponentiated coefficients reported in Table A.3.

Table A.4: Main Effect of Treatment: Sensitivity to Different Controls and Samples

(a) Sample restricted to first day per jobseeker–company pair

	Applied = 100							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	-0.273***	-0.273***	-0.277***	-0.260***	-0.278***	-0.280***	-0.283***	-0.269***
	(0.039)	(0.039)	(0.038)	(0.038)	(0.038)	(0.038)	(0.037)	(0.037)
N	37,309,899	37,309,899	37,309,899	37,309,899	37,309,899	37,309,899	37,309,899	37,309,899
Date FEs		✓	✓			✓	✓	
Company FEs			✓				✓	
Company-Date FEs				✓				✓
User Controls					✓	✓	✓	✓

(b) Whole Sample: all jobseeker-company-days observed

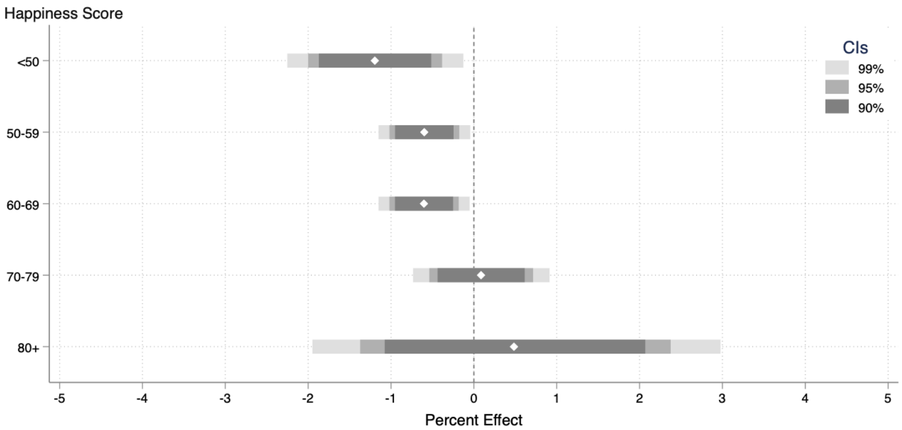
	Applied = 100							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	-0.280***	-0.281***	-0.285***	-0.267***	-0.285***	-0.287***	-0.292***	-0.276***
	(0.041)	(0.041)	(0.040)	(0.040)	(0.040)	(0.039)	(0.039)	(0.039)
Observations	39,712,785	39,712,785	39,712,785	39,712,785	39,712,785	39,712,785	39,712,785	39,712,785
Date FEs		✓	✓			✓	✓	
Company FEs			✓				✓	
Company-Date FEs				✓				✓
User Controls					✓	✓	✓	✓

(c) Sample restricted to first company-day pair per job seeker

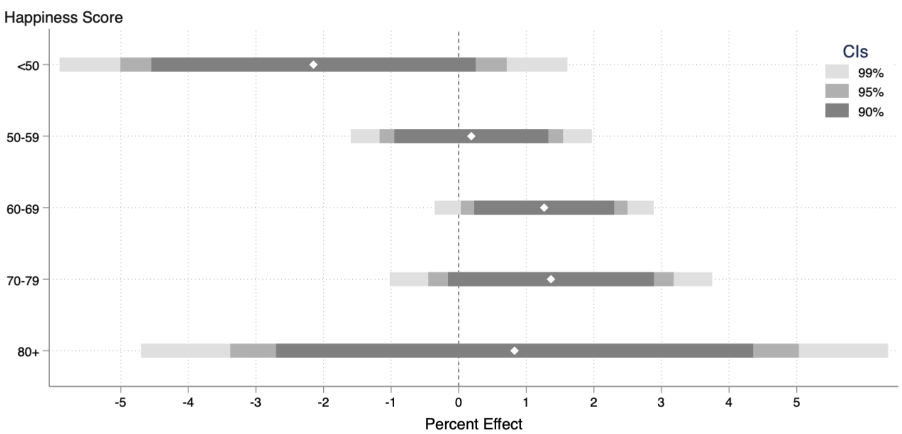
	Applied = 100							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	-0.266***	-0.269***	-0.261***	-0.259***	-0.270***	-0.272***	-0.265***	-0.266***
	(0.038)	(0.038)	(0.038)	(0.040)	(0.038)	(0.038)	(0.037)	(0.040)
N	22,997,526	22,997,526	22,997,526	22,997,526	22,997,525	22,997,525	22,997,525	22,997,525
Date FEs		✓	✓			✓	✓	
Company FEs			✓				✓	
Company-Date FEs				✓				✓
User Controls					✓	✓	✓	✓

Notes: Robust standard errors are in parentheses in Panel (c); robust standard errors adjusted for clustering on jobseekers are in parentheses in Panels (a) and (b). Outcome in all models is equal to 100 if the job seeker applied to a job at that company, 0 otherwise. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

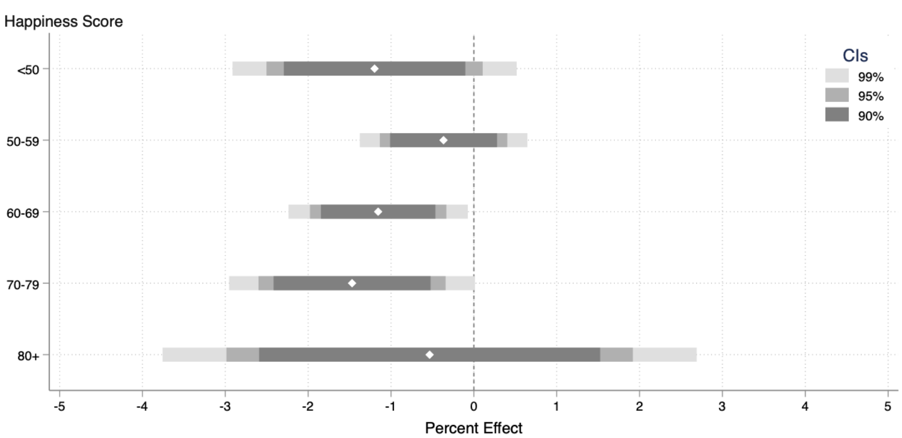
Figure A-5: Alternative Outcomes



(a) Outcome: Job Clicks



(b) Outcome: Followed Company = 1



(c) Outcome: Looked at Reviews Tab = 1

Note: panel (a) is based on a series of Poisson regressions, with the sample split according to the happiness score of the company in question; panels (b) and (c) are based on a series of LPMs. Percent effects are calculated using the control group mean of Pr(Apply) in panels (b) and (c), and in panel (a) are based on (the exponent of) the coefficients from the Poisson models. Company-by-day fixed effects included in all models. Standard errors are adjusted for clustering on job seekers.

A.5 Replication Studies in Canada and the UK

Table A.5: Effect of Showing Happiness Score on Application Behavior in the United Kingdom

	Applied = 100			
	(1)	(2)	(3)	(4)
Main Effect				
Treated	-0.603***	-0.570***	-0.570***	-0.599***
	(0.069)	(0.068)	(0.068)	(0.101)
Interactions: Treated				
× Happiness (z-score)			0.160***	
			(0.059)	
× score is 50-59				-0.206
				(0.137)
× score is 60-69				0.406**
				(0.168)
× score is 70-79				0.197
				(0.238)
× score is 80-100				1.197**
				(0.498)
Observations	2,496,398	2,493,551	2,493,551	2,493,551
User Controls		✓	✓	✓
Company-by-Date FEs		✓	✓	✓

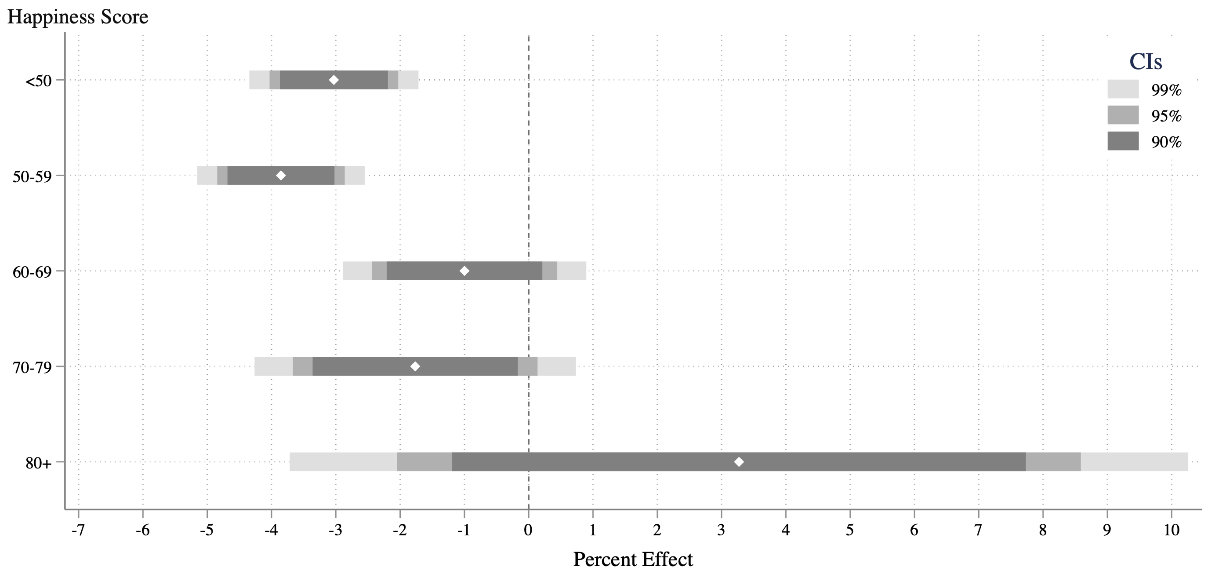
Notes: Robust standard errors are in parentheses, adjusted for clustering on individuals. A linear probability model is estimated in which the outcome is an indicator variable equal to 100 if the job seeker applies to that company. jobseeker controls: logged in job seeker, desktop job seeker, cookie age. In column (4), the omitted happiness score category is 20-49. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.6: Effect of Showing Happiness Score on Application Behavior in Canada

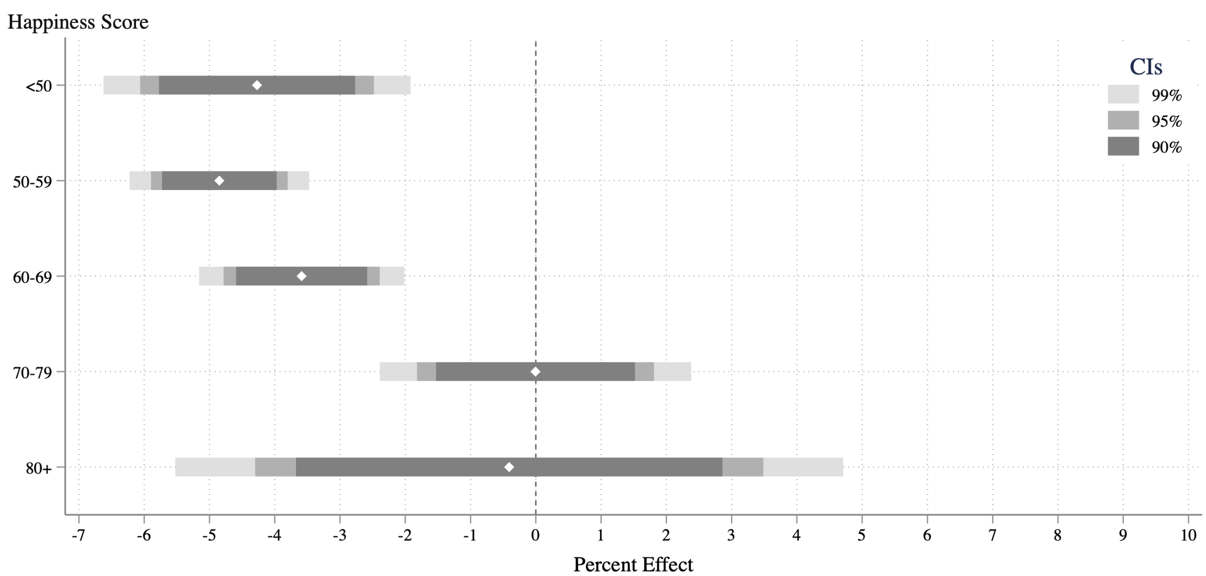
	Applied = 100			
	(1)	(2)	(3)	(4)
Main Effect				
Treated	-0.851***	-0.701***	-0.701***	-0.741***
	(0.072)	(0.070)	(0.070)	(0.158)
Interactions: Treated				
× Happiness (z-score)			0.204***	
			(0.061)	
× score is 50-59				-0.286
				(0.189)
× score is 60-69				0.050
				(0.192)
× score is 70-79				0.746***
				(0.222)
× score is 80-100				0.677*
				(0.376)
Observations	2,037,505	2,036,455	2,036,455	2,036,455
User Controls		✓	✓	✓
Company-by-Date FEs		✓	✓	✓

Notes: Robust standard errors are in parentheses, adjusted for clustering on individuals. A linear probability model is estimated in which the outcome is an indicator variable equal to 100 if the job seeker applies to that company. jobseeker controls: logged in job seeker, desktop job seeker, cookie age. In column (4), the omitted happiness score category is 20-49. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure A-6: Replication Studies: Split-Sample Percent Effects of Showing Happiness Score



(a) United Kingdom

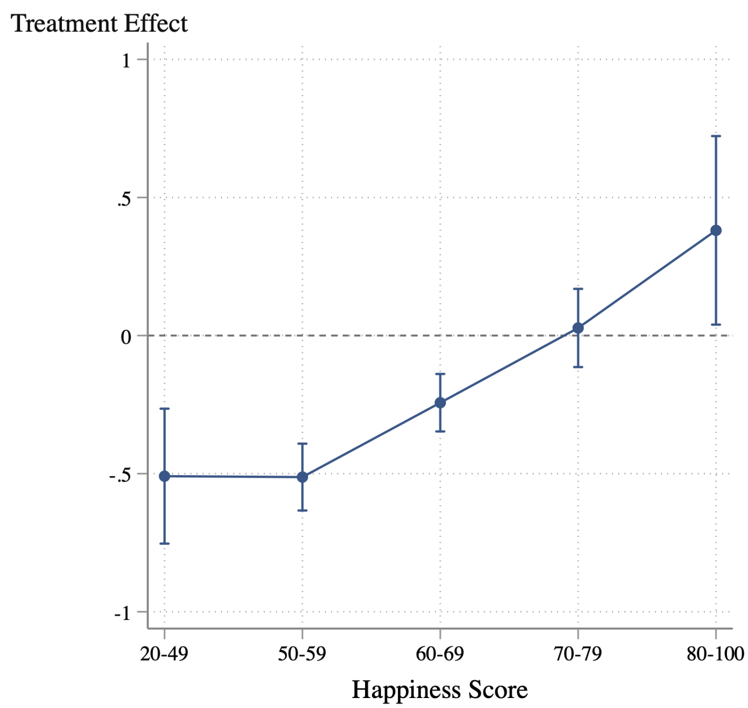


(b) Canada

Note: Percent effects are calculated using the control group mean of Pr(Apply) in each model. Results are derived from separate LPMs, with the sample split according to the happiness score of the company in question. Controls are included in all models for jobseeker characteristics as well as fixed effects for commuting zone and company-by-day. Standard errors are adjusted for clustering on job seekers.

A.6 Treatment Effect Heterogeneity

Figure A-7: Implied Treatment Effects from Interaction Model with Bins of Happiness Score Displayed



Note: Figure plots the implied treatment effects from the interaction models reported in column (4) of Table 1.1. In that model, the outcome is equal to 100 if the job seeker applies, and zero otherwise. The model includes controls for jobseeker characteristics, and a set of company-by-date fixed effects. The model includes a dummy for treatment, which is interacted with a series of dummies of the binned happiness score shown to job seekers. This figure plots the linear combination of treated and the interaction effect for each level of happiness, along with 95% confidence intervals.

Table A.7: Treatment Effect By Happiness Score of the Company Visited

(a) Sample restricted to first day per jobseeker–company pair

	(1) 20-49	(2) 50-59	(3) 60-69	(4) 70-79	(5) 80-100
Treated	-0.506***	-0.515***	-0.242***	0.026	0.385**
	(0.127)	(0.062)	(0.053)	(0.072)	(0.176)
Observations	2,299,865	11,042,034	15,933,238	6,916,461	1,118,283
Control Mean	18.81	20.01	20.94	18.74	18.59

(b) Whole Sample: all jobseeker-company-days observed

	(1) 20-49	(2) 50-59	(3) 60-69	(4) 70-79	(5) 80-100
Treated	-0.536***	-0.489***	-0.256***	0.004	0.337*
	(0.126)	(0.063)	(0.054)	(0.072)	(0.174)
Observations	2,456,781	11,776,805	16,962,276	7,336,604	1,180,299
Control Mean	18.60	19.91	20.92	18.83	18.56

(c) Sample restricted to first company-day pair per job seeker

	(1) 20-49	(2) 50-59	(3) 60-69	(4) 70-79	(5) 80-100
Treated	-0.598***	-0.472***	-0.243***	-0.015	0.413*
	(0.163)	(0.075)	(0.061)	(0.088)	(0.218)
Observations	1,396,195	6,700,754	9,848,111	4,321,047	731,390
Control Mean	18.81	20.11	20.97	18.51	18.35

Notes: Robust standard errors are in parentheses in Panel (c); robust standard errors adjusted for clustering on jobseekers are in parentheses in Panels (a) and (b). Outcome in all models is equal to 100 if the job seeker applied to a job at that company, 0 otherwise. Linear probability models reported. All models include company-day and commuting zone fixed effects, as well as controls for account age, whether logged in or not, and whether on desktop computer or not. Each column represents a separate model, with the sample determined by the Work Happiness Score of the company the job seeker was looking at. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.8: Treatment Effect by Well-Being Scores

	Applied = 100	
	(1)	(2)
Main Effect		
Treated	-0.262***	-0.260***
	(0.038)	(0.038)
× Happiness	0.022***	0.048***
	(0.004)	(0.017)
× Achievement		-0.017
		(0.013)
× Appreciation		-0.020
		(0.020)
× Belonging		-0.006
		(0.021)
× Energized		0.006
		(0.016)
× Flexibility		0.013
		(0.008)
× Inclusvity		-0.009
		(0.016)
× Learning		0.022*
		(0.012)
× Management		-0.003
		(0.014)
× Fair Pay		-0.006
		(0.006)
× Purpose		-0.005
		(0.012)
× Support		-0.000
		(0.021)
× Trust		-0.002
		(0.020)
Observations	37,309,899	37,219,457

Notes: Robust standard errors are in parentheses, adjusted for clustering on job seekers. Outcome in all models is equal to 100 if the job seeker applied to a job at that company, 0 otherwise. Linear probability models reported. All models include company-by-day fixed effects. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.9: Treatment Effect For Registered and Non-Registered job seekers

	(1) < 50	(2) 50-59	(3) 60-69	(4) 70-79	(5) 80+
Treated	-0.219	-0.375***	-0.288***	0.125	0.374*
	(0.168)	(0.081)	(0.069)	(0.091)	(0.221)
Registered User	11.620***	11.332***	10.715***	10.208***	9.581***
	(0.241)	(0.118)	(0.101)	(0.140)	(0.347)
× Registered	-0.485**	-0.286**	0.081	-0.193	0.031
	(0.247)	(0.121)	(0.104)	(0.144)	(0.355)
Observations	2,283,640	10,961,636	15,807,291	6,870,993	1,110,181
Control Mean (Registered)	11.53	13.24	14.63	13.06	13.20
Control Mean (Non-Registered)	23.90	25.22	26.34	24.18	24.15

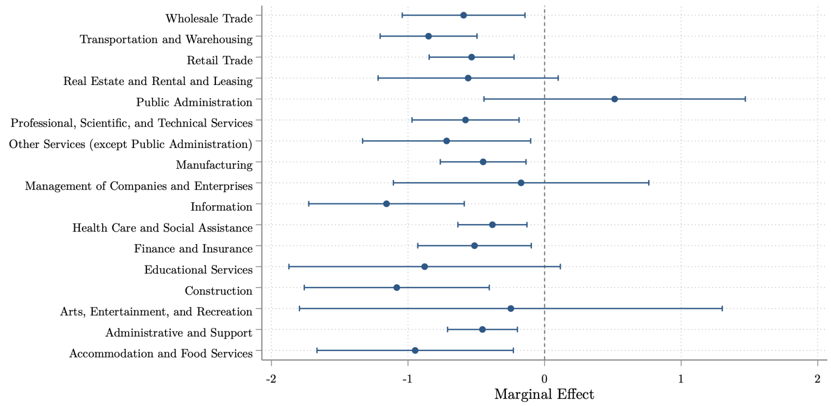
Notes: Robust standard errors are in parentheses, adjusted for clustering on job seekers. Outcome in all models is equal to 100 if the job seeker applied to a job at that company, 0 otherwise. Linear probability models reported. All models include company-by-day fixed effects. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.10: Treatment Effect Heterogeneity: jobseeker and Company Characteristics

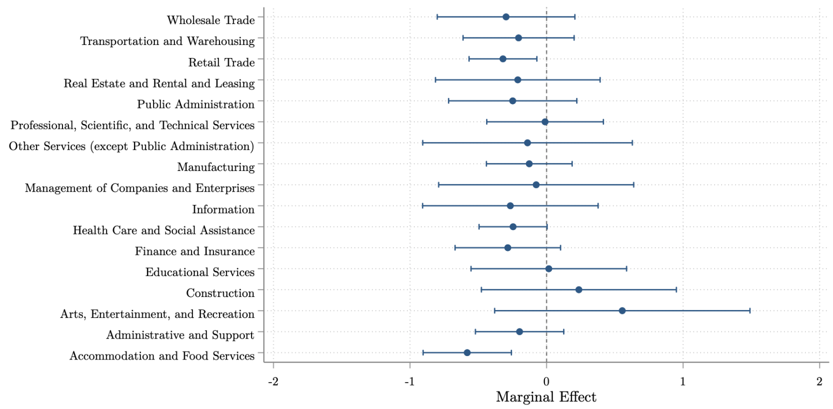
	< 50		50-59		60-69		70-79		80+	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Main Treatment Effect										
Treated	-0.880***	-2.382***	-0.658***	-0.500*	-0.245**	-0.232	-0.084	-0.440	0.512	0.857
	(0.251)	(0.915)	(0.128)	(0.261)	(0.115)	(0.215)	(0.159)	(0.355)	(0.386)	(1.629)
Covariate Effects										
Cookie Age (z-score)	-0.507***	-0.452*	-0.490***	-0.330***	-0.595***	-0.536***	-0.495***	-0.733***	-0.484***	-0.644*
	(0.051)	(0.237)	(0.026)	(0.120)	(0.024)	(0.106)	(0.033)	(0.143)	(0.081)	(0.351)
Employed User	0.346***	-0.006	0.220***	0.307	0.080	-0.243	0.529***	0.111	0.627***	0.749
	(0.112)	(0.512)	(0.057)	(0.266)	(0.052)	(0.236)	(0.073)	(0.326)	(0.179)	(0.793)
User high Education (BA or more)	-2.933***	-2.305***	-2.754***	-2.531***	-2.209***	-2.164***	-0.019	-0.086	0.022	0.843
	(0.114)	(0.497)	(0.058)	(0.255)	(0.054)	(0.228)	(0.075)	(0.315)	(0.184)	(0.767)
User Total Work Experience (z-score)	-0.942***	-0.712***	-1.063***	-1.074***	-1.346***	-1.239***	-0.908***	-0.861***	-0.954***	-1.052**
	(0.055)	(0.244)	(0.028)	(0.127)	(0.026)	(0.116)	(0.037)	(0.165)	(0.095)	(0.409)
Local Unemployment Rate (z-score)	0.246***	-0.010	0.431***	0.439***	0.550***	0.629***	0.653***	0.521***	0.553***	0.432
	(0.084)	(0.275)	(0.040)	(0.132)	(0.035)	(0.116)	(0.049)	(0.156)	(0.118)	(0.370)
Num Happiness Surveys (z-score)	19.727***	22.835***	-1.029**	-1.257***	0.200	0.193	4.865***	4.931***	0.523	-0.104
	(3.751)	(4.370)	(0.438)	(0.481)	(0.144)	(0.178)	(0.800)	(0.916)	(6.332)	(7.306)
Large Company (10,000+ employees)	2.297**	1.361	0.770	1.108**	2.309***	2.593***	1.587**	1.338*	8.865***	8.774***
	(0.942)	(1.206)	(0.477)	(0.540)	(0.537)	(0.576)	(0.716)	(0.781)	(1.783)	(2.091)
Company Star Rating (z-score)	0.133	0.712**	0.091	0.145	-0.154***	0.046	0.073	0.122	0.088	-0.175
	(0.194)	(0.322)	(0.079)	(0.162)	(0.053)	(0.133)	(0.159)	(0.289)	(0.300)	(0.683)
Company Num of Jobs Listed (z-score)	1.080**	-0.656	1.093**	1.757**	3.395***	3.370***	0.311**	0.222	10.312***	9.981***
	(0.535)	(0.882)	(0.476)	(0.821)	(0.174)	(0.487)	(0.121)	(0.226)	(0.425)	(1.198)
Interactions: Treated										
× Cookie Age		-0.058		-0.168		-0.063		0.251*		0.169
		(0.242)		(0.123)		(0.109)		(0.147)		(0.361)
× Employed		0.370		-0.091		0.340		0.440		-0.129
		(0.524)		(0.272)		(0.242)		(0.334)		(0.814)
× Education		-0.660		-0.234		-0.047		0.071		-0.865
		(0.508)		(0.260)		(0.233)		(0.322)		(0.784)
× User Work Experience		-0.243		0.012		-0.112		-0.049		0.103
		(0.250)		(0.130)		(0.119)		(0.169)		(0.420)
× Unemployment Rate		0.270		-0.008		-0.084		0.139		0.128
		(0.276)		(0.133)		(0.117)		(0.156)		(0.369)
× Num Happiness Surveys		-3.267		0.242		0.007		-0.066		0.640
		(2.354)		(0.211)		(0.111)		(0.473)		(3.758)
× Large Company		0.990		-0.354		-0.300		0.261		0.100
		(0.798)		(0.266)		(0.219)		(0.329)		(1.141)
× Star Rating		-0.612**		-0.056		-0.210		-0.052		0.275
		(0.274)		(0.149)		(0.129)		(0.254)		(0.631)
× Num Jobs Listed		1.809**		-0.696		0.026		0.093		0.352
		(0.755)		(0.701)		(0.478)		(0.200)		(1.179)
Observations	638,671	638,671	2,971,360	2,971,360	3,900,874	3,900,874	1,661,821	1,661,821	260,162	260,162

Notes: Robust standard errors are in parentheses, adjusted for clustering on job seekers. Outcome in all models is equal to 100 if the job seeker applied to a job at that company, 0 otherwise. Linear probability models reported. All models include company fixed effects and day fixed effects. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

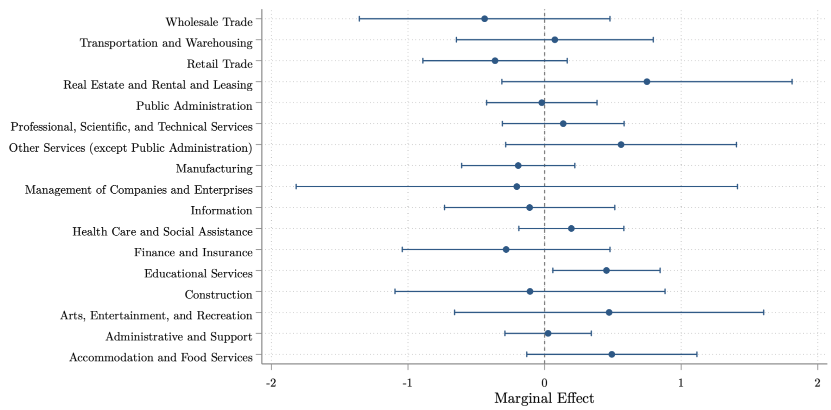
Figure A-8: Treatment Effect by Industry



(a) Low: Happiness Score < 60



(b) Average: Happiness Score $60-69$

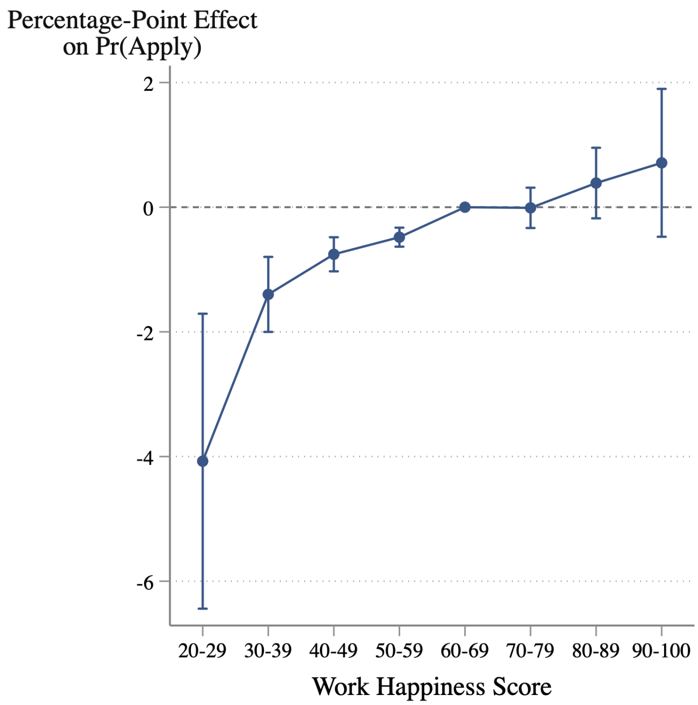


(c) High: Happiness Score 70+

Note: Each panel reports a regression with the sample split according to the happiness score shown. In each of the 3 regressions, the treatment indicator is interacted with industry fixed effects. Marginal effects are then calculated and reported, along with 95% confidence intervals.

A.7 Effects of Happiness Score on Applications: Functional Form

Figure A-9: Happiness Split into Categorical Variables



Note: The figure plots the coefficients and 95% confidence intervals from a regression of the probability of applying to a job on a series of indicators for the work happiness score displayed. 60–70 is the omitted category. The sample includes only treated job seekers. A linear probability model is estimated in which the outcome is an indicator variable equal to 100 if the job seeker applies to that company. The regression includes a full set of jobseeker and date fixed effects, as well as controls for company industry, company size, and indicators for whether the firm is a staffing agency and in the Fortune 500. Confidence intervals are derived from standard errors that are adjusted for two-way clustering on companies and job seekers.

A.8 IV Estimates of Score Effects

In this section, I build on the fixed effects and RD estimates by making use of the fact that i can observe the data generating process of these scores. In doing so I am able to pursue an alternative identification strategy that leverages plausibly exogenous variation in the happiness score.

A.8.1 Identification Strategy

Happiness scores are calculated and shown when a company has 20 or more individual-level surveys. These surveys are filled in by job seekers, who arrive at the survey page from a variety of different sources. Some arrive directly, by going to their company's review page and clicking through to the survey. Others arrive after being directed to the survey from an email or mobile alert. Still others arrive there from the resume pages of the website. jobseekers may upload their resume details to the website, and in doing so fill in their employment history. From here, the website directs these jobseekers to the happiness survey for each of the companies they have been employed by or are currently working for. As can be seen in Figure A-12, those arriving from the resume pages report on average much higher happiness, on the 1-to-5 agreement scale. While those arriving to the survey from company pages answer on average just over 2.5, for those arriving from resume pages this is over 1 point higher.

Across companies, there is variation in the "make-up" of their happiness scores – in terms of where its individual-level surveys originate from. Among their respondents, some have a larger proportion of resume respondents than others, as can be seen in Figure A-12. On average, around a quarter of responses come from the resume section, but this varies from zero up to around 60%. In order to identify the causal effect of the happiness score on applications, I make the (untestable) assumption that the percentage of a firm's responses coming from the resume section is as-good-as-random, controlling for observables like a company's industry and size. If this is the case, then the proportion of a score's responses arising from clicks from the resume section of the site will be a valid instrument for the score itself.

I estimate the equation

$$A_{ijt} = \beta H_{jt} + U_i + T_t + X'_j + \varepsilon_{ijt} \tag{A.1}$$

where U_i and T_t are jobseeker and date fixed effects, and $X`_j$ includes 26 fixed effects for the company's industry, 8 fixed effects for the company's number of employees, and indicator variables for if the company is a staffing agency and is in the Fortune 500. I instrument for H_{jt} using the proportion of individual-level surveys that make up H_{jt} that come from click-throughs from the resume section of the site, Z_{jt}.

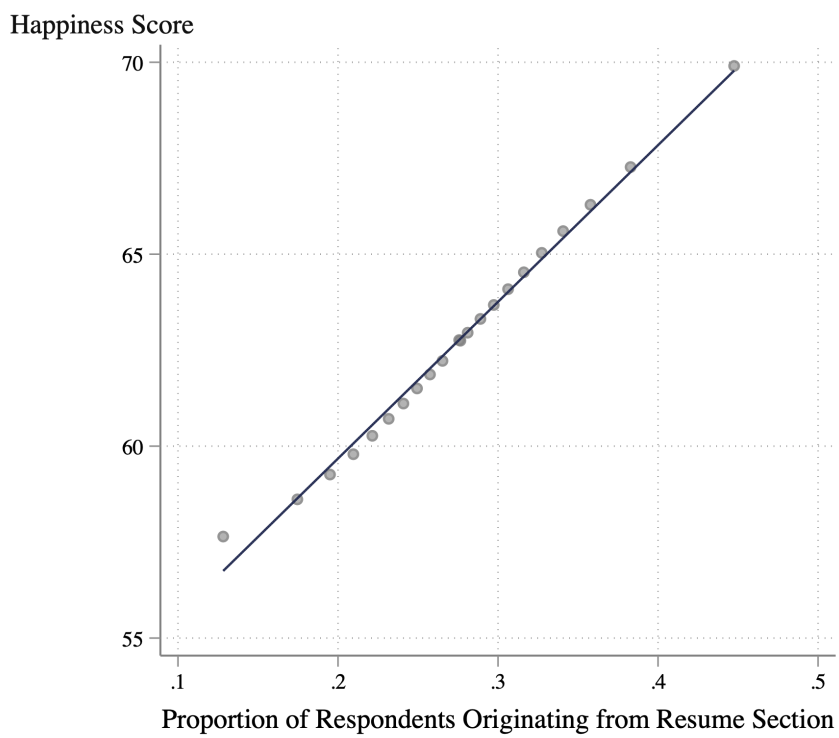
One key assumption in this set-up is that the share of resume respondents a company has is sufficiently correlated with the happiness score. Figure A-10 shows a binned scatterplot of the "first-stage" relationship between survey source and happiness score. Included in the regression are the full set of jobseeker and date fixed effects, along with a set of company observables. The proportion of surveys coming from the resume section has a strong positive impact on the aggregate score. The instrument has a first stage cluster-robust F-statistic of over 990, suggesting that the instrument is sufficiently strong to be valid.

A further key (untestable) assumption is that the effect that the share of resume responses a company has on application behavior runs through its influence on the happiness score only, condition on the covariates and fixed effects in the model.

A.8.2 Results

Column (1) of Table A.11 reports the key coefficient and standard error from this first-stage regression. Column (2) of Table A.11 reports the reduced form relationship between survey source and the probability of applying. The proportion of surveys coming from the resume

Figure A-10: IV First Stage Effect



Note: Figure shows the relationship between the composition of a company-day's aggregate happiness score in terms of the source of individual-level surveys, a using binned scatter plot. Both the happiness score and the proportion of surveys coming from resumes are first regressed on the full set of controls and fixed effects. The residuals from these regressions are binned across 40 quantiles and plotted as grey dots. The blue line shows the linear fit from a regression using all of the data.

Table A.11: IV Estimates of the Effect of Happiness Score on Application Behavior

	DV: Happiness	DV: Applied = 100			
	(1)	(2)	(3)	(4)	(5)
	1st Stage	Red.-Form	2SLS	Red.-Form	2SLS
Proportion Surveys from Resume	4.064***	0.434***		0.611***	
	(0.128)	(0.088)		(0.085)	
Happiness Score			0.107***		0.104***
			(0.022)		(0.023)
Resume-Source Surveys^2				-0.256***	
				(0.037)	
Happiness Score^2					-0.013***
					(0.002)
Observations	18,639,593	18,639,593	18,639,593	18,639,593	18,639,593
Kleibergen-Paap rk Wald F-Stat			1007.4		159.7

Notes: Robust standard errors are in parentheses, adjusted for clustering on companies. Outcome in all models is equal to 100 if the job seeker applied to a job at that company, 0 otherwise. All models include jobseeker and date fixed effects, as well as controls for company industry, company size and dummies for being a staffing agency and in the Fortune 500. Happiness score is re-centered around 0. Proportion surveys from resume is the IV in model (3); proportion surveys from resume and it's square are the instruments in model (5). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

section has a strong positive effect on application decisions. A coefficient of 4 suggests that displaying a score made up entirely of resume clickers increases the probability of applying by 4 percentage points, compared with a score made up with no jobseekers from the resume source. A one standard deviation increase in this proportion (which is around 0.11) increases the probability of applying by around half a percentage point.

Turning to the 2SLS estimate reported in column (3), a statistically significant coefficient of 0.104 [95% CI: 0.064, 0.149] suggests that a one-point increase in the score, on the 100-point scale, increases the application probability by around 0.1 percentage points.[1] In the remaining columns of Table A.11, I again estimate the 2SLS model, but this time introduce a quadratic term for the happiness score, as well as a quadratic term of the instrument. The use of two instrumental variables weakens the strength of the first stage, though it remains relatively strong with a cluster-robust F-statistic of over 160.[2] In column (5) of Table A.11, both linear and quadratic happiness terms enter significantly into the equation, with the

[1] Although the point estimate is slightly larger than the fixed-effect estimate reported above, the fixed-effect estimate falls well within the confidence interval of the 2SLS estimate.

[2] As above, when using polynomials, I first re-centre the happiness score to have a mean of zero.

expected signs. The happiness score has a positive causal effect on applications, which then flattens out at higher levels of the score.

Figure A-11: Source of Happiness Surveys Over Time

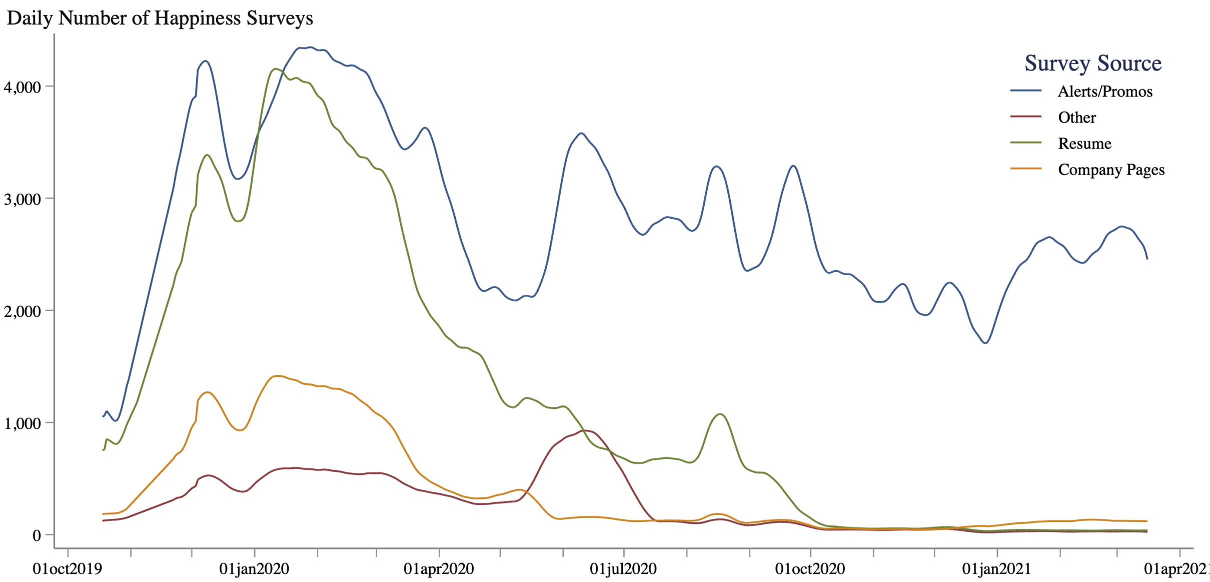
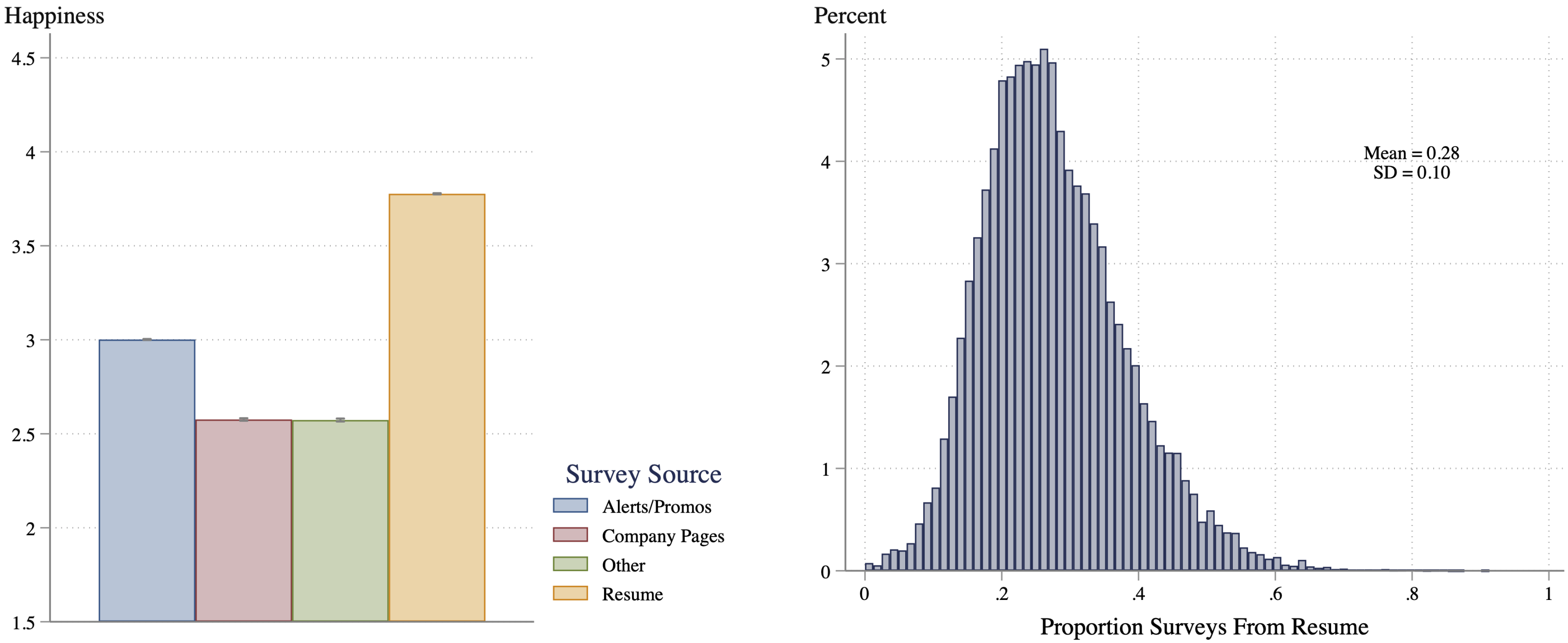


Figure A-12: Happiness by Survey Source



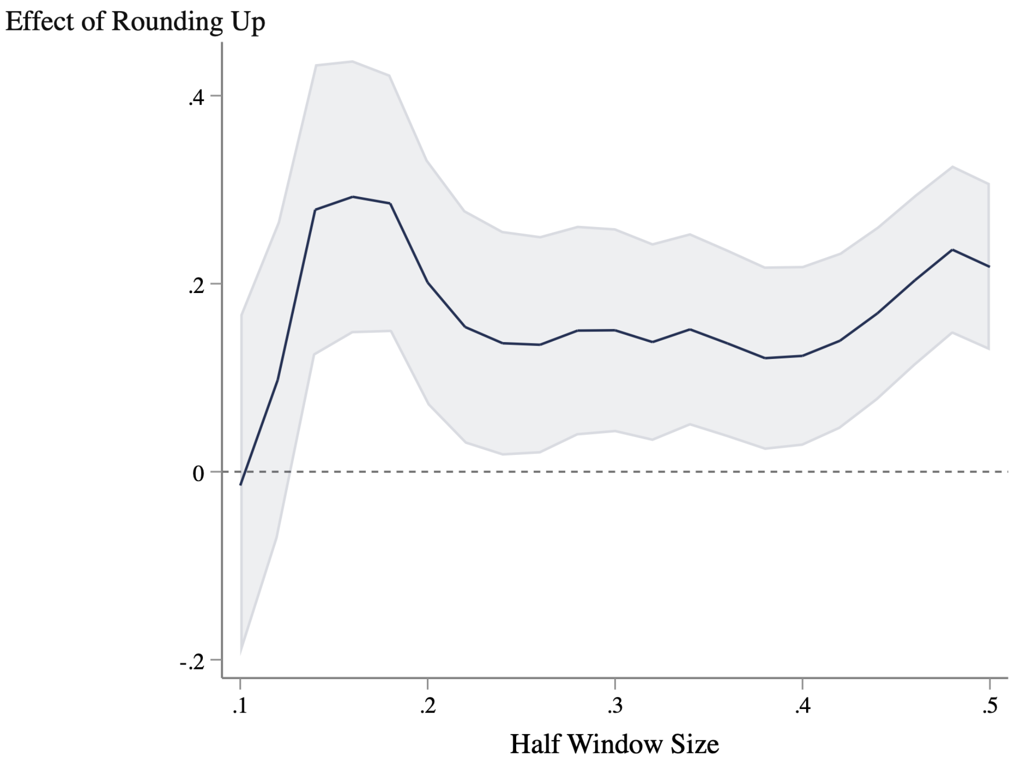
(a) Happiness by Survey Source

(b) Survey Source Across Companies

Note: Panel (a) plots the raw differences in happiness responses by survey source. Panel (b) shows the source of happiness surveys across company-days.

A.9 Local Randomization RD Analysis: Robustness

Figure A-13: Sensitivity of Window Size Selection



Note: Point estimates and 95% confidence intervals reported, from a series of regressions with differing window sizes around the cutoff points for inclusion (in increments of 0.02 from 0.1 to 0.5). All regressions are linear probability models, with the outcome equal to 100 if the job seeker applies and zero otherwise. All models include a dummy equal to 1 if the score is rounded up to the nearest integer, as well as cutoff fixed effects.

Table A.12: Local Randomization RD Analysis: Extra Results

RD Effect	Fisher p-value	Large-sample p-value	Window	N_{below}	N_{above}	Kernel
0.313	< 0.0001	< 0.0001	[-0.2, 0.2]	728,047	725,941	Uniform
0.290	< 0.0001	< 0.0001	[-0.2, 0.2]	728,047	725,941	Triangular

Notes: Table reports local randomization RD estimates of the effect of rounding up the score, using data pooled over the 5 cutoffs. P-values based on 1,000 replications. See Cattaneo et al. (2015, 2017) for more details of the approach.

Appendix B

Supplementary Materials for Chapter 2

B.1 Data Collection

Table B.1: Survey Quotas: Age

Age	Quota	US Sample	Quota	CA Sample	Quota	UK Sample
18-24	13%	7%	11%	6%	13%	9%
25-34	26%	26%	23%	24%	26%	23%
35-44	23%	25%	25%	27%	23%	26%
45-54	22%	23%	27%	27%	22%	26%
55-64	12%	14%	12%	13%	12%	14%
65 and over	3%	4%	2%	3%	3%	3%

Table B.2: Survey Quotas: Gender

	Quota	USA	Quota	CA	Quota	UK
Male	52%	53%	52%	49%	54%	52%
Female	48%	47%	48%	50%	46%	48%

Table B.3: Survey Quotas: Income

US Income	Quota	USA	CA Income	Quota	CA	UK Income	Quota	UK
< $20k	8%	5%	Under $20k	4%	2%	< £20k	15%	14%
$20k- $50k	27%	23%	$20k- $50k	16%	17%	£20k - £40k	33%	35%
$50k- $100k	36%	38%	$50k- $100k	40%	43%	£40k - £60k	25%	24%
$100k+	30%	34%	$100k or higher	39%	38%	£60k+	27%	26%

Table B.4: Survey Quotas: Education

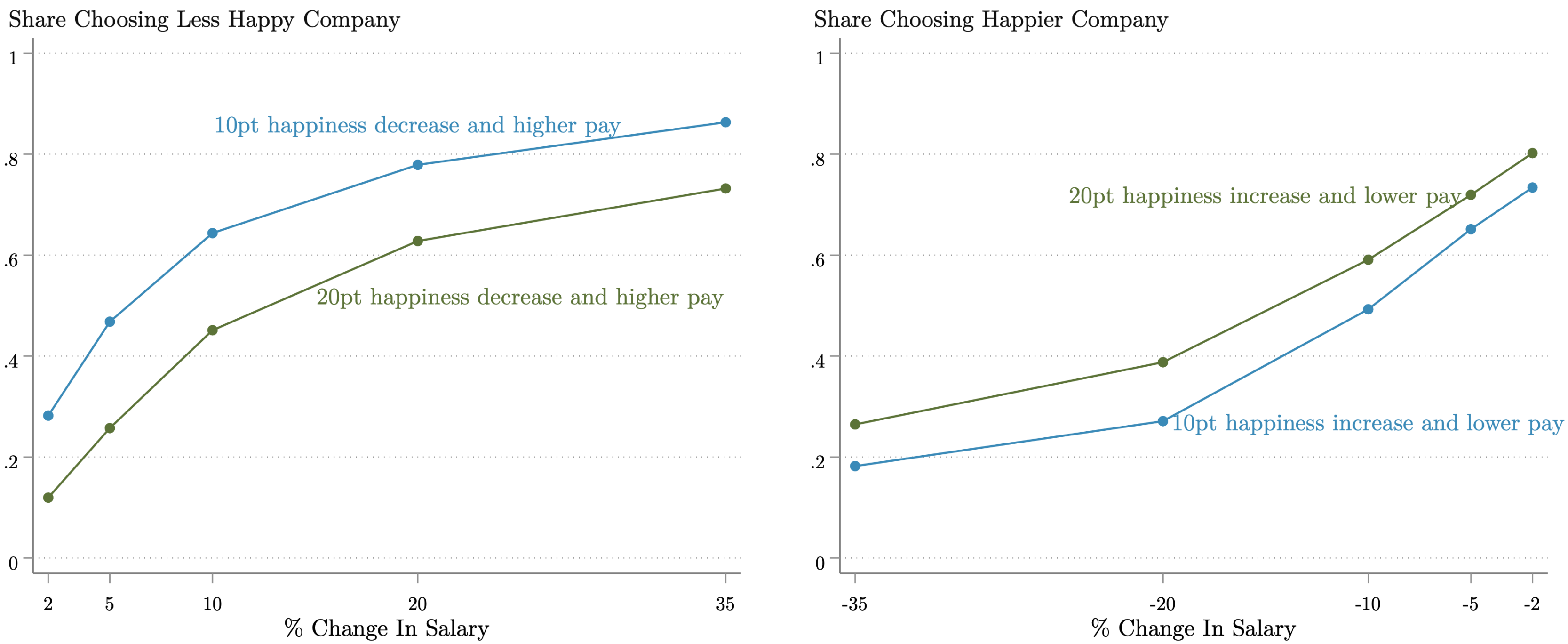
US Education Level	Quota	US	CA Education Level	Quota	CA	UK Education Level	Quota	UK
High School or Less	12%	10%	High School or Less	13%	10%	Secondary School	39%	41%
Some College; Associate's	26%	26%	Some College; Technical; CEGEP	35%	36%	Bachelor's Degree	35%	36%
BA	38%	39%	BA	33%	36%	Beyond Degree	25%	23%
Graduate Degree	24%	25%	Graduate Degree	18%	18%			

Table B.5: Survey Quotas: Region

UK Region	Quota	UK	CA Region	Quota	CA	US Region	Quota	US
North East	4%	4%	Alberta	11%	12%	Northeast	18%	19%
North West	11%	11%	British Columbia	13%	13%	Midwest	21%	22%
Yorkshire and Humber	8%	8%	Manitoba	3%	3%	South	36%	38%
East Midlands	7%	6%	New Brunswick	1%	1%	West	23%	21%
West Midlands	9%	9%	Newfoundland and Labrador	1%	1%			
East of England	10%	10%	Nova Scotia	3%	2%			
London	12%	14%	Ontario	43%	47%			
South East	14%	15%	Prince Edward Island	0%	0%			
South West	9%	9%	Quebec	23%	20%			
Wales	5%	4%	Saskatchewan	3%	2%			
Scotland	8%	9%						
Northern Ireland	3%	1%						

B.2 Replications in Canada and the UK

Figure B-1: Distribution of Survey Responses in Canada

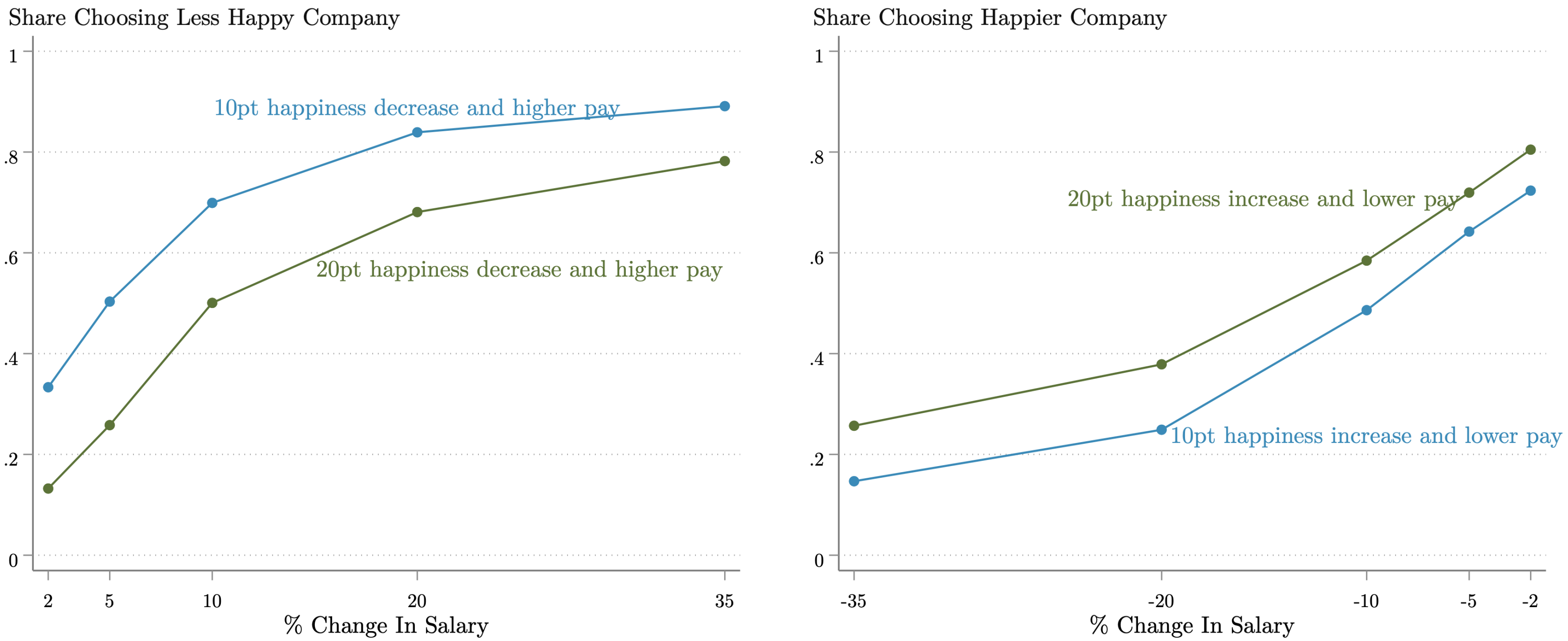


(a) Higher Pay, Lower Happiness

(b) Lower Pay, Higher Happiness

Note: In all cases, the base job is at a company with a happiness score of 65. Panel (a) reports the share of respondents preferring the company with lower happiness, at differing levels of pay increase. Panel (b) reports the share of respondents preferring the company with higher happiness, at differing levels of pay decrease.

Figure B-2: Distribution of Survey Responses in the United Kingdom



(a) Higher Pay, Lower Happiness

(b) Lower Pay, Higher Happiness

Note: In all cases, the base job is at a company with a happiness score of 65. Panel (a) reports the share of respondents preferring the company with lower happiness, at differing levels of pay increase. Panel (b) reports the share of respondents preferring the company with higher happiness, at differing levels of pay decrease.

B.3 Ordered Logit Models for WTP/WTA

Table B.6: Heterogeneity in WTA (Happiness 45 vs. 65)

		Marginal Effects, for Outcome:					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	O-Logit	2% (1)	5% (2)	10% (3)	20% (4)	35% (5)	None (6)
Demographics							
Men (d)	-0.096	0.009	0.008	0.006	-0.001	-0.003	-0.019
	(0.068)	(0.007)	(0.006)	(0.005)	(0.001)	(0.002)	(0.013)
Age	0.056***	-0.005***	-0.005***	-0.004***	0.001**	0.002***	0.011***
	(0.018)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)	(0.004)
Age^2	-0.001***	0.000***	0.000***	0.000***	-0.000**	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married (d)	-0.047	0.005	0.004	0.003	-0.001	-0.002	-0.009
	(0.083)	(0.008)	(0.007)	(0.006)	(0.001)	(0.003)	(0.016)
Has Children (d)	0.107	-0.010	-0.009	-0.007	0.002	0.004	0.021
	(0.079)	(0.008)	(0.007)	(0.005)	(0.001)	(0.003)	(0.015)
Has BA (d)	0.139*	-0.014*	-0.012*	-0.009*	0.002	0.005*	0.027*
	(0.076)	(0.008)	(0.006)	(0.005)	(0.001)	(0.003)	(0.015)
Black (v. white non-Hispanic) (d)	-0.151	0.015	0.013	0.009	-0.003	-0.006	-0.029
	(0.144)	(0.015)	(0.012)	(0.008)	(0.004)	(0.006)	(0.027)
Hispanic (d)	-0.326*	0.036	0.028*	0.018**	-0.009	-0.013	-0.060*
	(0.186)	(0.023)	(0.016)	(0.008)	(0.007)	(0.008)	(0.031)
Asian (d)	-0.069	0.007	0.006	0.004	-0.001	-0.003	-0.013
	(0.113)	(0.011)	(0.010)	(0.007)	(0.002)	(0.004)	(0.022)
Other race (d)	-0.039	0.004	0.003	0.003	-0.001	-0.001	-0.008
	(0.165)	(0.016)	(0.014)	(0.011)	(0.003)	(0.006)	(0.032)
Current Job							
Work Happiness	0.210***	-0.020***	-0.017***	-0.014***	0.003***	0.008***	0.041***
	(0.035)	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)	(0.007)
Income (log)	0.169***	-0.016***	-0.014***	-0.011***	0.002**	0.006***	0.033***
	(0.060)	(0.006)	(0.005)	(0.004)	(0.001)	(0.002)	(0.012)
Tenure (company)	-0.012*	0.001*	0.001*	0.001*	-0.000*	-0.000*	-0.002*
	(0.006)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Current Weekly Work Hours	-0.000	0.000	0.000	0.000	-0.000	-0.000	-0.000
	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
County (v. USA)							
UK (d)	-0.266***	0.027***	0.022***	0.016***	-0.005**	-0.010***	-0.050***
	(0.087)	(0.010)	(0.008)	(0.005)	(0.002)	(0.004)	(0.016)
Canada (d)	-0.112	0.011	0.009	0.007	-0.002	-0.004	-0.022
	(0.085)	(0.009)	(0.007)	(0.005)	(0.002)	(0.003)	(0.016)
cut1	1.729**						
	(0.726)						
cut2	2.657***						
	(0.726)						
cut3	3.607***						
	(0.728)						
cut4	4.360***						
	(0.729)						
cut5	4.829***						
	(0.730)						
Observations	2926	2926	2926	2926	2926	2926	2926

Notes: Robust standard errors are in parentheses. Ordered Logit regression reported in column (1). Marginal effects are reported in columns (2) to (7), separately for each of the six possible outcomes. (d) for discrete change of dummy variable from 0 to 1. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.7: Heterogeneity in WTA (Happiness 55 vs. 65)

		Marginal Effects, for Outcome:					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	O-Logit	2% (1)	5% (2)	10% (3)	20% (4)	35% (5)	None (6)
Demographics							
Men (d)	0.094	-0.019	-0.004	0.002	0.006	0.005	0.010
	(0.068)	(0.014)	(0.003)	(0.001)	(0.005)	(0.003)	(0.008)
Age	0.059***	-0.012***	-0.003***	0.001***	0.004***	0.003***	0.007***
	(0.019)	(0.004)	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)
Age^2	-0.001***	0.000***	0.000**	-0.000**	-0.000**	-0.000**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married (d)	-0.114	0.022	0.006	-0.002	-0.008	-0.006	-0.013
	(0.083)	(0.016)	(0.004)	(0.001)	(0.006)	(0.004)	(0.010)
Has Children (d)	-0.003	0.001	0.000	-0.000	-0.000	-0.000	-0.000
	(0.079)	(0.016)	(0.004)	(0.001)	(0.005)	(0.004)	(0.009)
Has BA (d)	0.062	-0.012	-0.003	0.001	0.004	0.003	0.007
	(0.076)	(0.015)	(0.004)	(0.002)	(0.005)	(0.004)	(0.008)
Black (v. white non-Hispanic) (d)	-0.206	0.043	0.009*	-0.005	-0.014	-0.010	-0.021
	(0.147)	(0.031)	(0.005)	(0.005)	(0.011)	(0.007)	(0.014)
Hispanic (d)	-0.201	0.042	0.008	-0.005	-0.014	-0.010	-0.021
	(0.189)	(0.041)	(0.007)	(0.007)	(0.014)	(0.009)	(0.018)
Asian (d)	-0.014	0.003	0.001	-0.000	-0.001	-0.001	-0.002
	(0.112)	(0.022)	(0.005)	(0.002)	(0.008)	(0.006)	(0.012)
Other race (d)	-0.179	0.037	0.008	-0.005	-0.012	-0.009	-0.019
	(0.167)	(0.036)	(0.006)	(0.006)	(0.012)	(0.008)	(0.016)
Current Job							
Work Happiness	0.148***	-0.029***	-0.007***	0.003***	0.010***	0.007***	0.016***
	(0.035)	(0.007)	(0.002)	(0.001)	(0.002)	(0.002)	(0.004)
Income (log)	0.184***	-0.037***	-0.009***	0.003***	0.012***	0.009***	0.020***
	(0.061)	(0.012)	(0.003)	(0.001)	(0.004)	(0.003)	(0.007)
Tenure (company)	-0.013**	0.003**	0.001*	-0.000*	-0.001*	-0.001*	-0.001**
	(0.006)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Current Weekly Work Hours	-0.002	0.000	0.000	-0.000	-0.000	-0.000	-0.000
	(0.004)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
County (v. USA)							
UK (d)	-0.309***	0.064***	0.013***	-0.008**	-0.022***	-0.015***	-0.032***
	(0.089)	(0.019)	(0.003)	(0.003)	(0.006)	(0.004)	(0.009)
Canada (d)	-0.089	0.018	0.004	-0.002	-0.006	-0.004	-0.010
	(0.085)	(0.017)	(0.004)	(0.002)	(0.006)	(0.004)	(0.009)
cut1	2.818***						
	(0.735)						
cut2	3.570***						
	(0.737)						
cut3	4.393***						
	(0.738)						
cut4	5.170***						
	(0.740)						
cut5	5.715***						
	(0.740)						
Observations	2926	2926	2926	2926	2926	2926	2926

Notes: Robust standard errors are in parentheses. Ordered Logit regression reported in column (1). Marginal effects are reported in columns (2) to (7), separately for each of the six possible outcomes. (d) for discrete change of dummy variable from 0 to 1. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.8: Heterogeneity in WTP (Happiness 75 vs. 65)

		Marginal Effects, for Outcome:					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	O-Logit	None (1)	2% (2)	5% (3)	10% (4)	20% (5)	35% (6)
Demographics							
Men (d)	0.143**	-0.028**	-0.004**	-0.004**	0.006**	0.008**	0.022**
	(0.068)	(0.013)	(0.002)	(0.002)	(0.003)	(0.004)	(0.010)
Age	0.009	-0.002	-0.000	-0.000	0.000	0.000	0.001
	(0.020)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Age^2	-0.000	0.000	0.000	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married (d)	0.080	-0.016	-0.002	-0.002	0.003	0.005	0.012
	(0.083)	(0.016)	(0.002)	(0.002)	(0.003)	(0.005)	(0.013)
Has Children (d)	0.247***	-0.049***	-0.007***	-0.006***	0.010***	0.014***	0.037***
	(0.078)	(0.015)	(0.002)	(0.002)	(0.003)	(0.005)	(0.012)
Has BA (d)	0.070	-0.014	-0.002	-0.002	0.003	0.004	0.011
	(0.076)	(0.015)	(0.002)	(0.002)	(0.003)	(0.004)	(0.011)
Black (v. white non-Hispanic) (d)	0.317**	-0.058**	-0.010**	-0.011*	0.008***	0.017**	0.053**
	(0.144)	(0.024)	(0.005)	(0.006)	(0.002)	(0.007)	(0.026)
Hispanic (d)	0.374**	-0.067**	-0.012*	-0.014	0.008***	0.020**	0.064*
	(0.187)	(0.030)	(0.006)	(0.009)	(0.002)	(0.009)	(0.035)
Asian (d)	-0.001	0.000	0.000	0.000	-0.000	-0.000	-0.000
	(0.113)	(0.022)	(0.003)	(0.003)	(0.004)	(0.006)	(0.017)
Other race (d)	0.149	-0.028	-0.005	-0.005	0.005	0.008	0.024
	(0.160)	(0.029)	(0.005)	(0.006)	(0.004)	(0.009)	(0.027)
Current Job							
Work Happiness	0.048	-0.009	-0.001	-0.001	0.002	0.003	0.007
	(0.034)	(0.007)	(0.001)	(0.001)	(0.001)	(0.002)	(0.005)
Income (log)	-0.003	0.001	0.000	0.000	-0.000	-0.000	-0.001
	(0.059)	(0.012)	(0.002)	(0.002)	(0.002)	(0.003)	(0.009)
Tenure (company)	-0.009	0.002	0.000	0.000	-0.000	-0.001	-0.001
	(0.006)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Current Weekly Work Hours	0.006	-0.001	-0.000	-0.000	0.000	0.000	0.001
	(0.004)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
County (v. USA)							
UK (d)	-0.222**	0.044**	0.006***	0.005***	-0.010**	-0.013**	-0.033***
	(0.088)	(0.018)	(0.002)	(0.002)	(0.004)	(0.005)	(0.012)
Canada (d)	-0.069	0.014	0.002	0.002	-0.003	-0.004	-0.010
	(0.085)	(0.017)	(0.002)	(0.002)	(0.004)	(0.005)	(0.013)
cut1	-0.656						
	(0.740)						
cut2	-0.307						
	(0.740)						
cut3	0.311						
	(0.740)						
cut4	1.202						
	(0.740)						
cut5	1.816**						
	(0.740)						
Observations	2928	2928	2928	2928	2928	2928	2928

Notes: Robust standard errors are in parentheses. Ordered Logit regression reported in column (1). Marginal effects are reported in columns (2) to (7), separately for each of the six possible outcomes. (d) for discrete change of dummy variable from 0 to 1. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

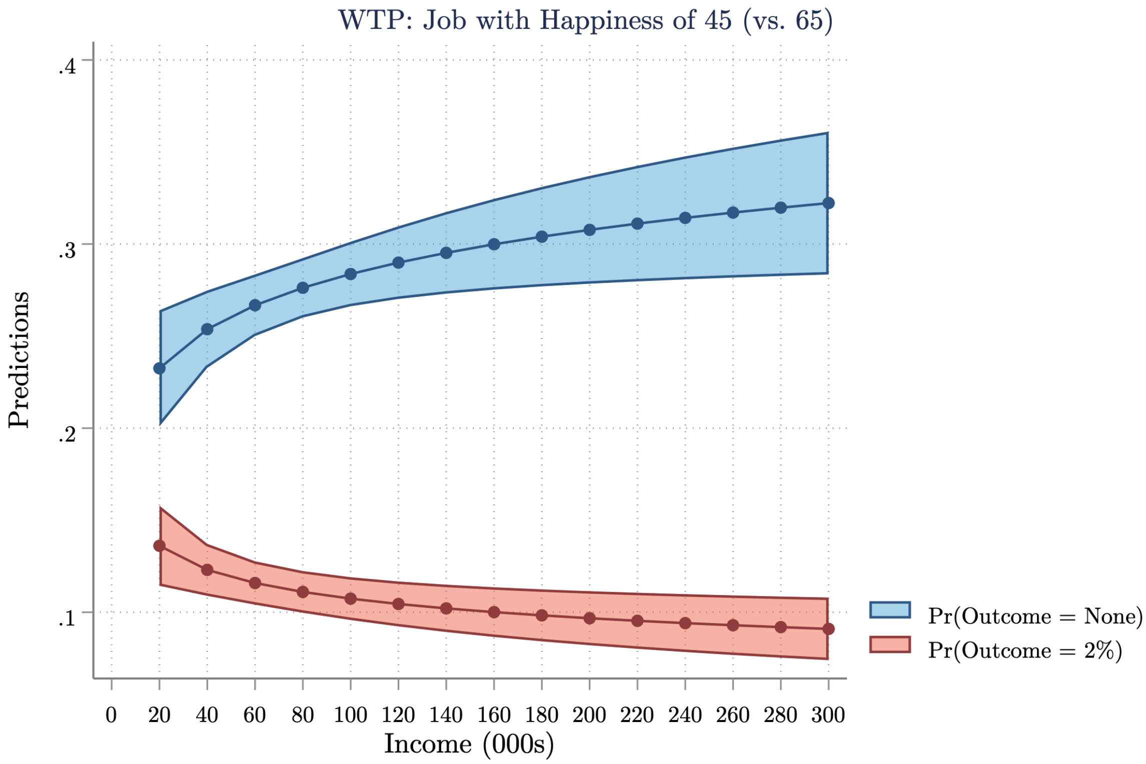
Table B.9: Heterogeneity in WTP (Happiness 85 vs. 65)

		Marginal Effects, for Outcome:					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	O-Logit	None (1)	2% (2)	5% (3)	10% (4)	20% (5)	35% (6)
Demographics							
Men (d)	0.098	-0.014	-0.004	-0.005	-0.001	0.004	0.020
	(0.068)	(0.010)	(0.003)	(0.003)	(0.001)	(0.003)	(0.014)
Age	0.002	-0.000	-0.000	-0.000	-0.000	0.000	0.001
	(0.020)	(0.003)	(0.001)	(0.001)	(0.000)	(0.001)	(0.004)
Age^2	-0.000	0.000	0.000	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married (d)	0.047	-0.007	-0.002	-0.002	-0.000	0.002	0.010
	(0.083)	(0.012)	(0.004)	(0.004)	(0.001)	(0.003)	(0.017)
Has Children (d)	0.338***	-0.050***	-0.015***	-0.015***	-0.002*	0.013***	0.069***
	(0.078)	(0.012)	(0.003)	(0.004)	(0.001)	(0.003)	(0.016)
Has BA (d)	0.111	-0.016	-0.005	-0.005	-0.001	0.004	0.023
	(0.076)	(0.011)	(0.003)	(0.003)	(0.001)	(0.003)	(0.016)
Black (v. white non-Hispanic) (d)	0.364**	-0.047***	-0.016**	-0.019**	-0.009	0.010***	0.080**
	(0.150)	(0.017)	(0.006)	(0.008)	(0.006)	(0.003)	(0.035)
Hispanic (d)	0.492**	-0.061***	-0.021***	-0.026**	-0.014	0.011***	0.111**
	(0.193)	(0.020)	(0.008)	(0.011)	(0.009)	(0.002)	(0.046)
Asian (d)	0.113	-0.016	-0.005	-0.005	-0.002	0.004	0.024
	(0.114)	(0.016)	(0.005)	(0.006)	(0.002)	(0.004)	(0.025)
Other race (d)	0.244	-0.033	-0.011	-0.012	-0.005	0.007*	0.053
	(0.164)	(0.020)	(0.007)	(0.009)	(0.005)	(0.004)	(0.037)
Current Job							
Work Happiness	0.067**	-0.010**	-0.003*	-0.003*	-0.001	0.002*	0.014**
	(0.034)	(0.005)	(0.001)	(0.002)	(0.000)	(0.001)	(0.007)
Income (log)	0.008	-0.001	-0.000	-0.000	-0.000	0.000	0.002
	(0.060)	(0.009)	(0.003)	(0.003)	(0.001)	(0.002)	(0.012)
Tenure (company)	-0.011*	0.002*	0.000*	0.001*	0.000	-0.000*	-0.002*
	(0.006)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Current Weekly Work Hours	0.010**	-0.001**	-0.000**	-0.000**	-0.000*	0.000**	0.002**
	(0.004)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
County (v. USA)							
UK (d)	-0.243***	0.037***	0.011***	0.011***	0.001	-0.010**	-0.049***
	(0.088)	(0.014)	(0.004)	(0.004)	(0.001)	(0.004)	(0.017)
Canada (d)	-0.141*	0.021	0.006*	0.006*	0.001	-0.005	-0.029*
	(0.085)	(0.013)	(0.004)	(0.004)	(0.000)	(0.003)	(0.017)
cut1	-1.111						
	(0.752)						
cut2	-0.655						
	(0.752)						
cut3	-0.061						
	(0.752)						
cut4	0.735						
	(0.752)						
cut5	1.302*						
	(0.752)						
Observations	2928	2928	2928	2928	2928	2928	2928

Notes: Robust standard errors are in parentheses. Ordered Logit regression reported in column (1). Marginal effects are reported in columns (2) to (7), separately for each of the six possible outcomes. (d) for discrete change of dummy variable from 0 to 1. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

B.4 Income and Heterogeneity in WTP/WTA

Figure B-3: Marginal Effects for WTA (45 vs. 65)



Note: Margins are plotted from an ordered logit regression (reported in column (1) of Table B.6). The two lines and 95% confidence intervals are based on columns (2) and (7) of the same table.

Appendix C

Supplementary Materials for Chapter 3

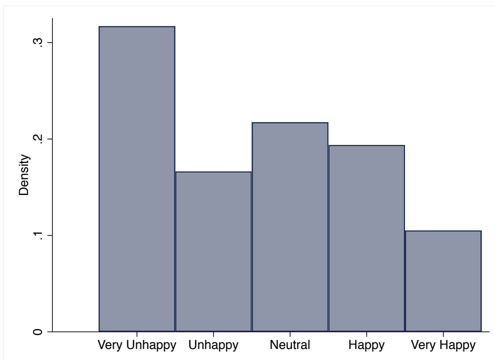
By *Clément Bellet, Jan-Emmanuel De Neve, and George Ward*

C.1 Summary Statistics

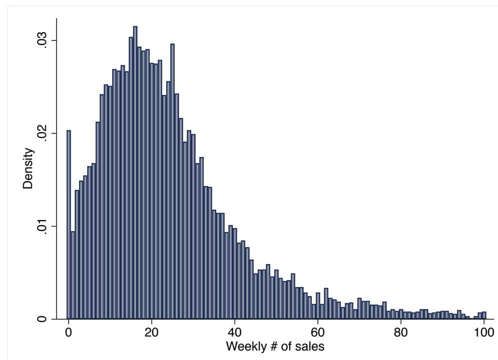
Table C.1: Summary Statistics

	N	Mean	Standard Deviation			Min	Max
			Overall	Between	Within		
Happiness	12,282	4.01	3.44	2.55	2.37	0	10
Sales	12,282	25.25	19.16	14.89	12.41	0	101
Selling Time	12,282	19.99	8.18	5.69	6.38	0	52.43
Internal Shrinkage	12,282	10.52	13.30	10.11	10.74	0	100
Customer Satisfaction	10,120	7.90	2.26	1.20	2.06	0	10
Adherence	12,174	91.99	5.48	3.58	4.34	0	100
Calls per Hour	12,282	4.96	1.62	1.16	1.14	0	67.29
Conversion Rate	11,850	26.63	17.98	15.77	10.59	0	100
Sick Leave	12,279	1.58	7.73	3.37	7.26	0	100
Attendance	12,279	92.58	14.14	6.55	13.15	0	100
Breaks (hrs)	12,282	3.77	1.49	1.15	1.09	0	9.25
Overtime (hrs)	12,282	0.15	0.98	0.48	0.89	0	15.58
Paid Time Off (hrs)	12,282	1.07	2.99	1.36	2.78	0	38
Gloomy Weather	12,282	4.06	1.78	1.18	1.36	0	10
Visual Exposure to Gloomy Weather	12,282	0.77	0.62	0.53	0.32	0	3.61
Windows (% of wall surface)	1,157	0.20		0.14		0.03	0.59
Age	1,157	33.81		10.41		17.5	67.5
Female	1,157	0.41		0.50		0	1
Tenure	1,157	4.99		7.17		0.09	43.79
Left Firm During Study	1,157	0.04		0.26		0	1

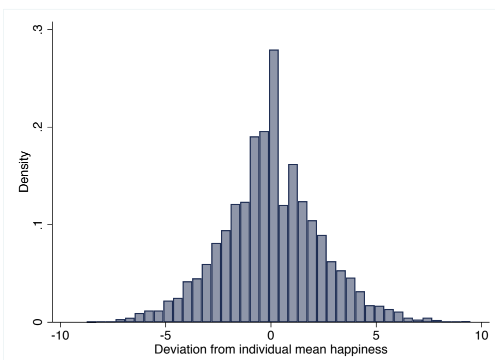
Figure C-1: Distribution of Happiness and Sales



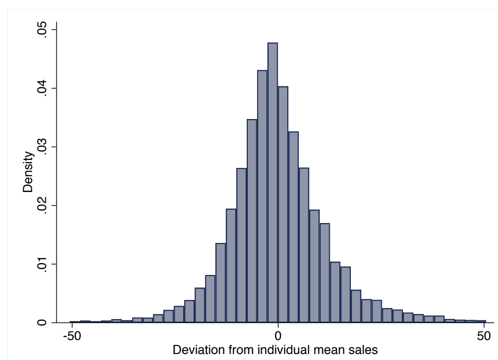
(a) Happiness



(b) Sales



(c) Happiness (within-worker)



(d) Sales (within-worker)

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

C.2 Enrollment/Attrition/Non-Response

Table C.2: 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	-0.004	-0.005	-0.002	-0.001
	(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)	(0.036)	(0.035)	(0.034)
Tenure (months)	-0.019*	-0.018	-0.018	-0.020	-0.020
	(0.010)	(0.011)	(0.011)	(0.012)	(0.012)
Mean selling hours during study	0.009***	0.009***	0.010***	0.010***	0.007*
	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Mean sales during study		-0.000	-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.002)	(0.002)
Mean team happiness (excl. worker)		-0.007	-0.007	-0.005	-0.006
		(0.006)	(0.006)	(0.007)	(0.007)
Mean gloomy weather (call center)			-0.006		
			(0.005)		
Building window coverage			-0.066		
			(0.059)		
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	0.060	0.069

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

Table C.3: 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**	-0.083**	-0.071**	-0.071**
	(0.037)	(0.036)	(0.037)	(0.035)	(0.031)
Left firm during study	-0.862***	-0.850***	-0.847***	-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)	(0.023)	(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 gloomy weather (call center)			0.019		
			(0.012)		
Building window coverage			0.029		
			(0.093)		
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.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C.4: Predictors of attrition/non-response

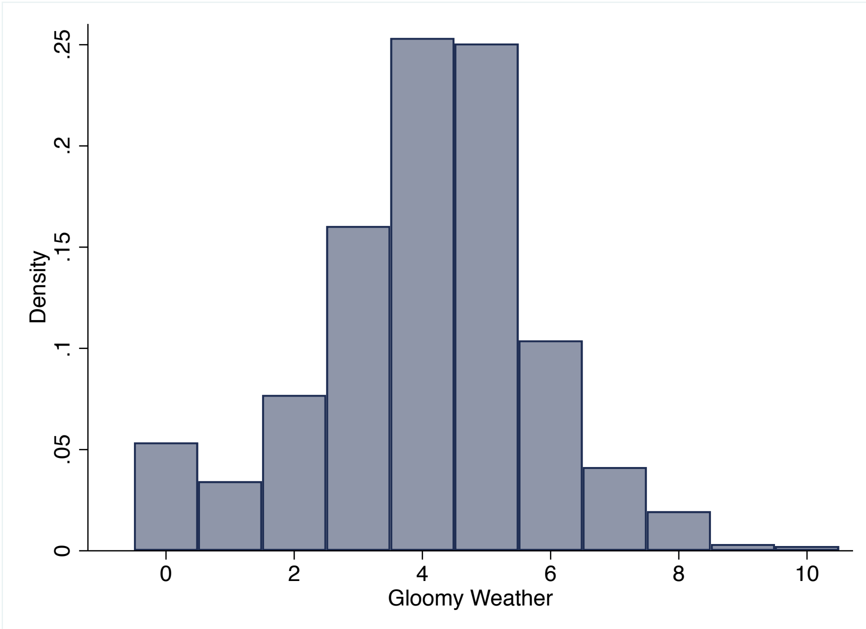
	Responded to Survey = 1					
	(1)	(2)	(3)	(4)	(5)	(6)
Sales	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Selling time	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Not working on thursday/friday	-0.153***	-0.153***	-0.153***	-0.153***	-0.153***	-0.153***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Gloomy weather index		0.001			0.001	
		(0.002)			(0.002)	
Gloomy 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.

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

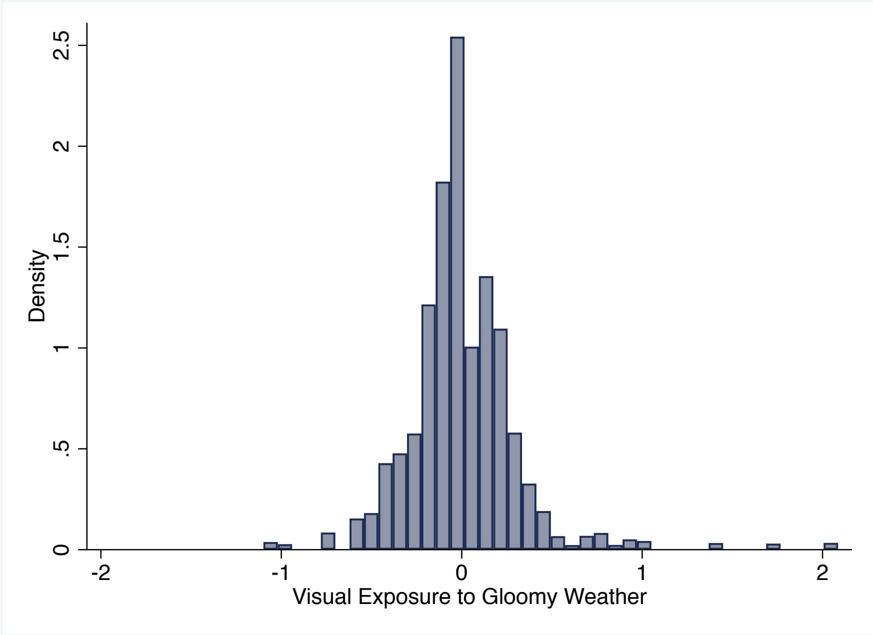
C.3 IV 1st Stage

Figure C-2: Distribution of (Raw) Gloomy Weather Index



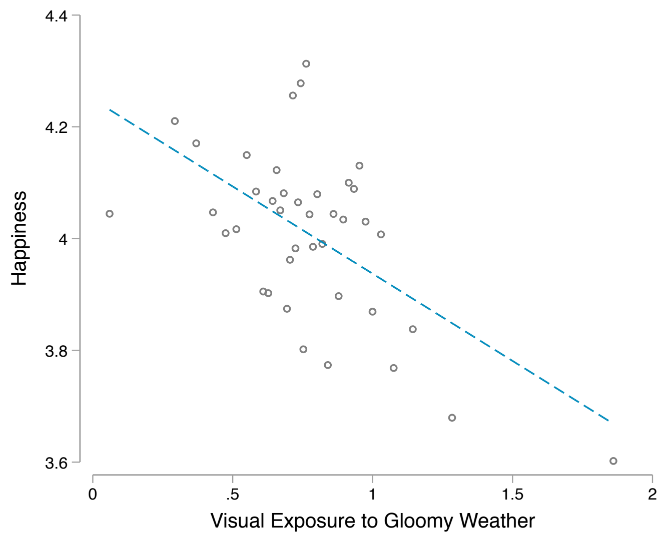
Note: Figure shows the distribution of gloomy weather.

Figure C-3: Distribution of Instrument: Visual exposure to gloomy weather



*Note: Figure shows the distribution of visual exposure to gloomy weather (adverse weather index * window coverage). We plot the residuals following a regression of visual exposure to gloomy weather on location and week fixed effects.*

Figure C-4: Weather IV First Stage: Graphical Representation



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

Table C.5: Alternative IV Definitions

	IV: Weather^2		IV: asinh(Weather)		IV: Gloom Dummy	
	(1)	(2)	(3)	(4)	(5)	(6)
	1st Stage IV	2nd Stage IV	1st Stage IV	2nd Stage IV	1st Stage IV	2nd Stage IV
Happiness		0.0976*		0.1557**		0.0772
		(0.0556)		(0.0709)		(0.0690)
(Gloomy Weather $\times$ % Windows)^2	-0.1063***					
	(0.0185)					
asinh(Gloomy Weather $\times$ % Windows)			-0.4369***			
			(0.1069)			
Gloomy Weather Dummy $\times$ % Windows					-1.7197***	
					(0.4156)	
Observations	12,282	12,282	12,282	12,282	12,282	12,282
1st Stage F-Stat	32.96		16.69		17.12	

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 dummies for day of week of response to survey. Bad weather dummy is equal to 1 if the index is 7 or above, 0 otherwise. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C.4 Additional Robustness

Table C.6: Sensitivity to Gloomy Weather: Heterogeneous Effects

	Happiness					
	(1)	(2)	(3)	(4)	(5)	(6)
Gloomy Weather	-0.0617***	-0.0438*	-0.0324	-0.0658***	-0.0850***	-0.0711***
	(0.0204)	(0.0264)	(0.0585)	(0.0246)	(0.0271)	(0.0272)
Weather Interaction:						
x Female Worker		-0.0449				
		(0.0354)				
x Worker's Age			-0.0008			
			(0.0015)			
x Worker's Tenure (Years)				0.0007		
				(0.0021)		
x Worker Avg Sickness > Median					0.0477	
					(0.0348)	
x # Sales						0.0004
						(0.0008)
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. Column (6) also controls for weekly sales as it varies within workers.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C.7: Gloomy Weather vs. Local Share of Sick Employees

	Local Sickness	
	(1)	(2)
Panel A: Daily Call-Center-Level Regressions		
Gloomy Weather	-0.1070	
	(0.1052)	
Gloomy Weather (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		
Gloomy Weather	-0.3487	
	(0.3711)	
Gloomy Weather (1-week lag)		-0.3541
		(0.2290)
Observations	289	278
R^2	0.333	0.329

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

Table C.8: Daily Call-Center-Level Regressions

	Local Demand (Poisson-FE)	Local Speed (OLS-FE)
Gloomy Weather	-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.

Table C.9: Results When Using Weather Index Dis-Aggregated

	Sales	
	(1)	(2)
	1st Stage IV	2nd Stage IV
Happiness		0.1127**
		(0.0564)
Exposure to Fog	-0.4198***	
	(0.1296)	
Exposure to Rain	-0.2578**	
	(0.1003)	
Exposure to Snow	-0.3084	
	(0.2450)	
Observations	12,282	12,282
1st Stage F-Stat	8.23	

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 dummies for day of week of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C.10: Controlling for Temperature

	Poisson-IV	
	(1)	(2)
	1st Stage	2nd Stage
Happiness		0.1478*
		(0.0871)
Exposure to Gloomy Weather	-0.2952***	
	(0.0670)	
Weekly mean temperature	0.0322	-0.0078
	(0.0221)	(0073)
Observations	12,121	12,121
1st Stage F-Stat	19.44	

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 dummies for day of week of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C.11: Null Effects of Visual Exposure to Temperature on Sales and Happiness

	(1)	(2)	(3)
Panel A: Reduced Form (Dep. Variable: Sales)			
Temperature	-0.0015	-0.0016	
	(0.0037)	(0.0037)	
Temperature × (z-scored) Window Coverage		-0.0012	
		(0.0009)	
Temperature (weighted by window coverage)			-0.0086
			(0.0058)
Observations	12,131	12,131	12,131
Panel B: First Stage (Dep. Variable: Happiness)			
Temperature	0.0454*	0.0438*	
	(0.0237)	(0.0238)	
Temperature × (z-scored) Window Coverage		-0.0023	
		(0.0043)	
Temperature (weighted by window coverage)			-0.0102
			(0.0305)
Observations	12,131	12,131	12,131
F-Stat of IV(s)	3.66	1.95	0.11

Notes: Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. Panel A: Poisson-FE models reported with weekly sales as dependent variable. Panel B: OLS-FE models reported with weekly happiness as dependent variable. 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 C.12: Happiness and Customer Satisfaction

	Customer Satisfaction (z-score)			
	(1)	(2)	(3)	(4)
	OLS-FE	OLS-FE	2SLS	2SLS
Exposure to Gloomy Weather	0.0270	0.0282		
	(0.0258)	(0.0253)		
Happiness			-0.0824	-0.0886
			(0.0795)	(0.0795)
# Sales		0.0032***		0.0052**
		(0.0011)		(0.0022)
Observations	10,059	10,059	10,059	10,059
R^2	0.173	0.174	-0.055	-0.060
1st Stage F-Stat			21.30	19.46

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

Table C.13: 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*
	(0.1004)	(0.0475)
Observations	12,282	12,282
Employees	1,157	1,157
R^2	0.535	
Pseudo-R^2		0.620
Week FEs	✓	✓
Individual FEs	✓	✓

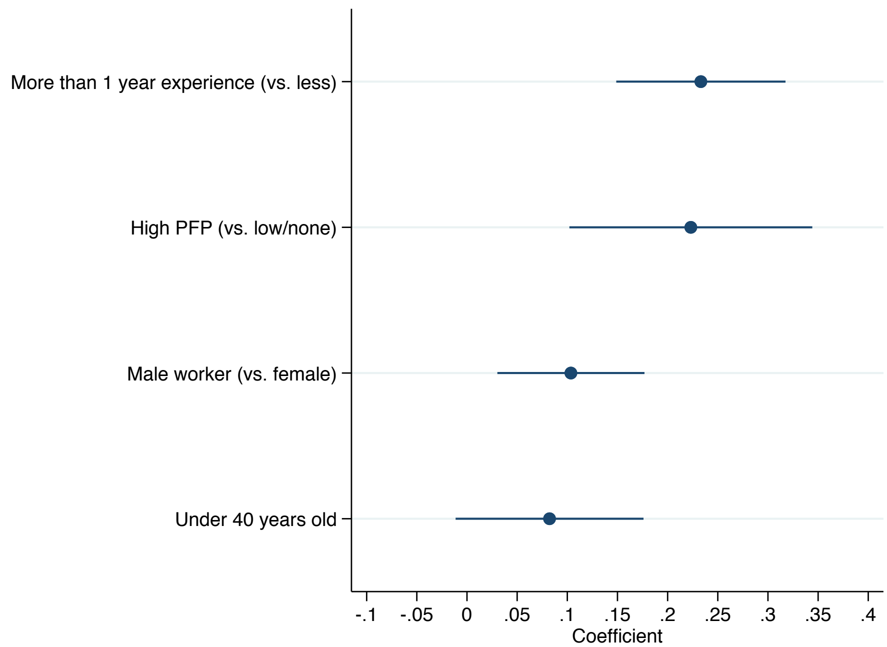
Notes: Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. All models include controls for the working hours, internal shrinkage, and day of week dummies for response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C.14: Visual Exposure to Gloomy Weather: 1st Stage Treatment Effects (LATE)

	Happiness					
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure to Gloomy Weather	-0.3122***	-0.2823***	-0.1346	-0.2803***	-0.3732***	-0.3328***
	(0.0654)	(0.0907)	(0.2276)	(0.0779)	(0.0963)	(0.0827)
Weather Interaction:						
x Female Worker		-0.0785				
		(0.1306)				
x Worker's Age			-0.0051			
			(0.0058)			
x Worker's Tenure (Years)				-0.0072		
				(0.0084)		
x Worker Avg Sickness > Median					0.1243	
					(0.1308)	
x # Sales						0.0012
						(0.0029)
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. Column (6) also controls for weekly sales as it varies within workers.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure C-5: Effect Size Benchmarking



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

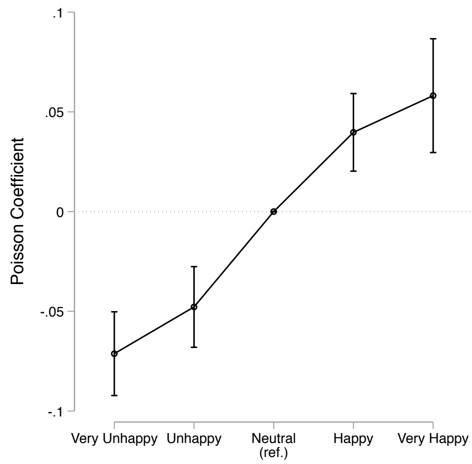
Table C.15: Impact of Happiness on Sales: Controlling for Channels Other than Conversion

	Sales (Poisson-IV)			
	(1)	(2)	(3)	(4)
Happiness	0.1331**	0.1356**	0.1368**	0.1365**
	(0.0594)	(0.0615)	(0.0586)	(0.0597)
Adherence (Met Target=1)		-0.0257*		-0.0267*
		(0.0150)		(0.0150)
Calls per hour (ln)			0.0531	0.0545
			(0.0608)	(0.0631)
Observations	12,282	12,169	12,100	12,033
1st Stage F-Stat	22.82	19.17	22.65	19.82

Notes: Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. Poisson-FE models reported where happiness is instrumented using the control function approach. All models include individual and week fixed effects, work schedule controls, and dummies for day of week of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C.5 Functional Form of Happiness

Figure C-6: 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.

C.6 Estimates using Daily Data

Table C.16: Happiness and Sales Performance using Daily Data

	Sales			
	(1)	(2)	(3)	(4)
	Poisson	Poisson	Poisson-IV	Poisson-IV
Very Unhappy	-0.0732***			
	(0.0107)			
Unhappy	-0.0498***			
	(0.0104)			
Happy	0.0355***			
	(0.0096)			
Very Happy	0.0621***			
	(0.0139)			
Happiness		0.0143***	0.1585	0.1542*
		(0.0014)	(0.1150)	(0.0793)
Observations	43,927	43,927	40,929	40,929
1st Stage F-Stat			11.04	15.02
IV	—	—	Daily Weather	Weekly Weather

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

Table C.17: Reduced Form Effects of Daily Weather on Sales

	Whole Sample		Non-Missing Happiness Data Only	
	(1)	(2)	(3)	(4)
Gloomy Weather Exposure (Daily)	-0.0499		-0.0539	
	(0.0330)		(0.0360)	
Gloomy Weather Exposure (Weekly)		-0.0370***		-0.0388***
		(0.0123)		(0.0143)
Observations	79,296	80,152	42,111	42,613
Pseudo-R^2	0.372	0.372	0.363	0.364

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

Table C.18: Impact of Adherence and Speed on Sales Conditional on Labor Supply

	Sales		
	(1)	(2)	(3)
Adherence (Met Target=1)	0.0034		0.0034
	(0.0076)		(0.0076)
Total number of calls per hour (ln)		-0.0579	-0.0616*
		(0.0374)	(0.0373)
Labor Supply:			
Total number of selling hours (ln)	0.9831***	0.9926***	0.9967***
	(0.0184)	(0.0134)	(0.0134)
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, and controls working hours, internal shrinkage, and day of response to survey. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C.7 Suggestive Evidence on Psychological Mechanisms

Table C.19: Effects by Sales Type

	# Sales		
	(1) Phone + Internet Lines	(2) TV + Cell Phone Contracts	(3) Upgrades + Re-Contracting
Happiness	-0.0324	0.0987	0.1860
	(0.1647)	(0.0840)	(0.1497)
Observations	12,241	12,268	12,264
1st Stage F-Stat	22.55	23.04	23.21

Notes: Poisson-FE-IV 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 C.20: Effects by Average Weekly Customer Satisfaction

	Sales (Poisson-FE)	
	Red.-Form	2nd Stage IV
Panel A: Unsatisfied Customers Weeks		
Happiness		0.2753***
		(0.0866)
Exposure to Gloomy Weather	-0.0997***	
	(0.0318)	
Observations	4,161	4,161
Panel B: Satisfied Customers Weeks		
Happiness		0.1321**
		(0.0551)
Exposure to Gloomy Weather	-0.0612**	
	(0.0253)	
Observations	4,081	4,081
Panel C: Highly Satisfied Customers Weeks		
Happiness		-0.0873
		(0.1564)
Exposure to Gloomy Weather	0.0243	
	(0.0432)	
Observations	3,462	3,462

Notes: Poisson-FE models reported. Robust standard errors in parentheses, adjusted for two-way clustering on individuals and location-week. We divide the sample in terms of average weekly customer satisfaction by terciles. All models include individual and week fixed effects, work schedule controls, and day of response to survey.

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C.8 Extra Detail on Effect Size Comparisons

Oswald et al. (2015): The figure we quote 1 SD increase of happiness leading to 29 to 35 percent more 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.

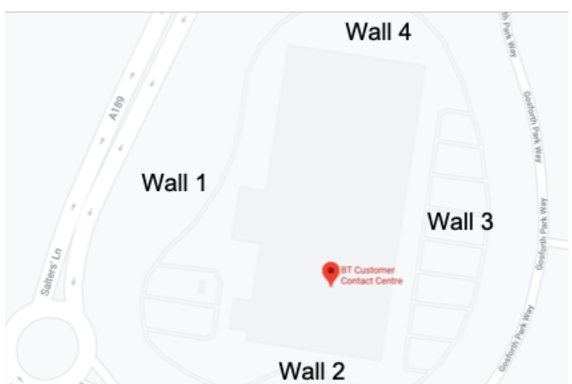
Bloom et al. (2014): We use the replication data provided by the authors to re-run regressions of Table 7, replacing the logged dependent variables with standardized dependent variables. The experiment reduced negative emotions by .44 SDs and increased positive emotions by .55 SDs. Similar effect sizes are found for evaluative satisfaction measures. Keeping the logged dependent variables and inferring standardized effect sizes using the control group SDs provides similar estimates.

C.9 Calculation of Share 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. Figure C-7 illustrates this first step with the map of Newcastle call center, which has four clearly defined walls (two long walls that capture 35% of total building surface each, and two short walls that capture 15% of total building surface each). See column 3 of Table C.21 for the full sample.

Figure C-7: Newcastle BT Call Center (Google Map View)



Note: Screenshot from Newcastle BT Call Center Map. Source: Google Maps.

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. As illustrated in Figure C-8, 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.

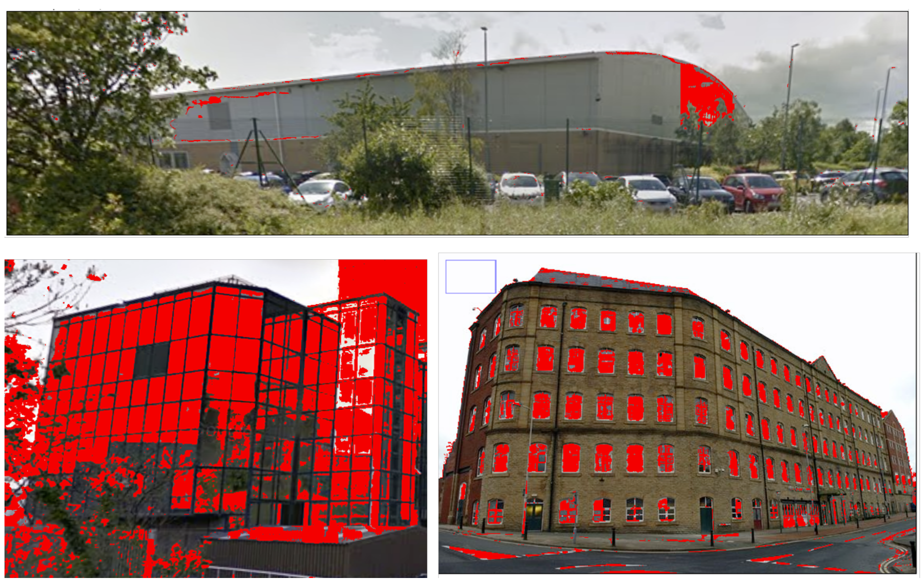
To compute the share of wall surface with windows, we use the open source image processing software ImageJ (see `https://imagej.nih.gov/ij/`). ImageJ can calculate area and pixel value statistics of user-defined selections and intensity-thresholded objects, like the color of wall surface or windows. 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. Figure C-9 illustrates ImageJ window processing for the three wall pictures described earlier. 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). See column 4 of Table C.21 for the full sample.

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%.

Figure C-8: Selected Wall Types (Google Street View)

Note: Screenshots from various BT call center walls. Top panel: Walls 2 and 3 from Newcastle call center. Bottom left panel: wall 3 from Swansea call center. Bottom right panel: wall 3 from Accrington call center.

Figure C-9: Selected Wall Types (Google Street View) After ImageJ Processing



Note: Screenshots from various BT call center walls after window processing with ImageJ. Top panel: Walls 2 and 3 from Newcastle call center. Bottom left panel: wall 3 from Swansea call center. Bottom right panel: wall 3 from Accrington call center.

Table C.21: Call Center Window Share

Call Center	*Wall #*	*Wall Size (%)*	*Wall Windows (%)*	*Final Windows (%)*
Accrington	1	0,48	0,38	0,32
Accrington	2	0,13	0,30	0,32
Accrington	3	0,39	0,26	0,32
Canterbury	1	0,50	0,12	0,11
Canterbury	2	0,10	0,08	0,11
Canterbury	3	0,10	0,08	0,11
Canterbury	4	0,30	0,12	0,11
Doncaster*	1+2	0,50	0,03	0,03
Doncaster*	3+4	0,50	0,03	0,03
Dundee	1	0,18	0,26	0,20
Dundee	2	0,21	0,33	0,20
Dundee	3	0,25	0,33	0,20
Dundee	4	0,36	0,00	0,20
Glasgow	1	0,50	0,34	0,31
Glasgow	2	0,50	0,29	0,31
Lancaster	1	0,39	0,40	0,36
Lancaster	2	0,39	0,40	0,36
Lancaster	3	0,11	0,22	0,36
Lancaster	4	0,11	0,22	0,36
Newcastle	1	0,35	0,04	0,06
Newcastle	2	0,15	0,10	0,06
Newcastle	3	0,35	0,00	0,06
Newcastle	4	0,15	0,18	0,06
South Shields	1	0,24	0,18	0,29
South Shields	2	0,29	0,18	0,29
South Shields	3	0,12	0,53	0,29
South Shields	4	0,12	0,53	0,29
South Shields	5	0,07	0,30	0,29
South Shields	6	0,07	0,30	0,29
South Shields	7	0,05	0,23	0,29
South Shields	8	0,05	0,23	0,29
Swansea	1	0,33	0,77	0,59
Swansea	2	0,20	0,52	0,59
Swansea	3	0,07	0,89	0,59
Swansea	4	0,17	0,45	0,59
Swansea	5	0,23	0,40	0,59
Truro	1	0,29	0,13	0,17
Truro	2	0,29	0,13	0,17
Truro	3	0,21	0,24	0,17
Truro	4	0,21	0,24	0,17
Warrington	1	0,17	0,38	0,14
Warrington	2	0,37	0,17	0,14
Warrington	3	0,12	0,02	0,14
Warrington	4	0,34	0,01	0,14

*Notes: Wall size measurement based on building map as shown on Google Maps. Wall pictures collected from Google Street View. Share wall surface with windows computed from the share of wall pixels with windows using ImageJ software. *Only interior angle pictures available for Doncaster from Google Images, with one long and short wall shown for each angle picture.*

Bibliography

Abowd, John M, John Haltiwanger, Ron Jarmin, Julia Lane, Paul Lengermann, Kristin McCue, Kevin McKinney, and Kristin Sandusky, "The relation among human capital, productivity, and market value: Building up from micro evidence," in "Measuring capital in the new economy," University of Chicago Press, 2005, pp. 153–204.

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

Achor, Shawn, "Is happiness the secret of success?," *CNN*, 2012. [Online; accessed 2022-03-19].

Adda, Jérôme, "Economic activity and the spread of viral diseases: Evidence from high frequency data," *The Quarterly Journal of Economics*, 2016, *131* (2), 891–941.

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

Allas, Tera and Bill Schaninger, "The boss factor: Making the world a better place through workplace relationships," *The McKinsey Quarterly*, 2020.

Altmann, Steffen, Armin Falk, Simon Jäger, and Florian Zimmermann, "Learning about job search: A field experiment with job seekers in Germany," *Journal of Public Economics*, 2018, *164*, 33–49.

Amabile, Teresa M, Sigal G Barsade, Jennifer S Mueller, and Barry M Staw, "Affect and creativity at work," *Administrative Science Quarterly*, 2005, *50* (3), 367–403.

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

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

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

Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott S Lee, "Losing prosociality in the quest for talent? Sorting, selection, and productivity in the delivery of public services," *American Economic Review*, 2020, *110* (5), 1355–94.

Baker, Scott R and Andrey Fradkin, "The impact of unemployment insurance on job search: Evidence from Google search data," *Review of Economics and Statistics*, 2017, *99* (5), 756–768.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social connections and incentives in the workplace: Evidence from personnel data," *Econometrica*, 2009, *77* (4), 1047–1094.

_ , _ , and _ , "Social incentives in the workplace," *Review of Economic Studies*, 2010, *77* (2), 417–458.

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

Baumeister, Roy F, Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D Vohs, "Bad is stronger than good," *Review of general psychology*, 2001, *5* (4), 323.

Belot, Michele, Philipp Kircher, and Paul Muller, "Providing advice to jobseekers at low cost: An experimental study on online advice," *Review of Economic Studies*, 2019, *86* (4), 1411–1447.

Benson, Alan, Aaron Sojourner, and Akhmed Umyarov, "Can Reputation Discipline the Gig Economy? Experimental Evidence from an Online Labor Market," *Management Science*, 2020, *66* (5), 1802–1825.

Bhattacharya, Utpal, Craig W Holden, and Stacey Jacobsen, "Penny wise, dollar foolish: Buy–sell imbalances on and around round numbers," *Management Science*, 2012, *58* (2), 413–431.

Bhole, Monica, Andrey Fradkin, and John Horton, "Information About Vacancy Competition Redirects Job Search," *SocArXiv*, 2021, *p82fk.*

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

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

Bond, Timothy N and Kevin Lang, "The sad truth about happiness scales," *Journal of Political Economy*, 2019, *127* (4), 1629–1640.

Borusyak, Kirill and Peter Hull, "Non-Random Exposure to Exogenous Shocks: Theory and Applications," *NBER Working Paper 27845*, 2020.

Bradler, Christiane, Robert Dur, Susanne Neckermann, and Arjan Non, "Employee recognition and performance: A field experiment," *Management Science*, 2016, *62* (11), 3085–3099.

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

_ **and W.H. Crockett**, "Employee attitudes and employee performance," *Psychological Bulletin*, 1955, *52*, 396–424.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani, "The Morale Effects of Pay Inequality," *Quarterly Journal of Economics*, 2018, *133* (2), 611–663.

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

Brown, Jennifer and David A Matsa, "Boarding a sinking ship? An investigation of job applications to distressed firms," *The Journal of Finance*, 2016, *71* (2), 507–550.

Brown, Thomas C, "Loss aversion without the endowment effect, and other explanations for the WTA–WTP disparity," *Journal of Economic Behavior & Organization*, 2005, *57* (3), 367–379.

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

_ **, John Forth, and Lucy Stokes**, "Does employees' subjective well-being affect workplace performance?," *Human Relations*, 2017, *70* (8), 1017–1037.

Burbano, Vanessa C, "Social responsibility messages and worker wage requirements: Field experimental evidence from online labor marketplaces," *Organization Science*, 2016, *27* (4), 1010–1028.

Cahlíková, Jana, Lubomír Cingl, and Ian Levely, "How stress affects performance and competitiveness across gender," *Management Science*, 2020, *66* (8), 3295–3310.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Robust Inference With Multiway Clustering," *Journal of Business & Economic Statistics*, 2011, *29* (2), 238–249.

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

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez, "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction," *American Economic Review*, 2012, *102* (6), 2981–3003.

Carmichael, H Lorne, "Reputations in the labor market," *The American Economic Review*, 1984, *74* (4), 713–725.

Carnahan, Seth, David Kryscynski, and Daniel Olson, "When does corporate social responsibility reduce employee turnover? Evidence from attorneys before and after 9/11," *Academy of Management Journal*, 2017, *60* (5), 1932–1962.

Cassar, Lea and Stephan Meier, "Nonmonetary Incentives and the Implications of Work as a Source of Meaning," *Journal of Economic Perspectives*, 2018, *32* (3), 215–38.

Cattaneo, Matias D, Brigham R Frandsen, and Rocio Titiunik, "Randomization inference in the regression discontinuity design: An application to party advantages in the US Senate," *Journal of Causal Inference*, 2015, *3* (1), 1–24.

_ , Rocio Titiunik, and Gonzalo Vazquez-Bare, "Comparing inference approaches for RD designs: A reexamination of the effect of head start on child mortality," *Journal of Policy Analysis and Management*, 2017, *36* (3), 643–681.

Chadi, Adrian, Sabrina Jeworrek, and Vanessa Mertins, "When the meaning of work has disappeared: experimental evidence on employees' performance and emotions," *Management Science*, 2017, *63* (6), 1696–1707.

Chamberlain, Andrew, "Does company culture pay off? Analyzing stock performance of "Best Places to Work" companies," *Glassdoor Research Report*, 2015.

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

Chang, Tom Y, Joshua Graff Zivin, Tal Gross, and Matthew Neidell, "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China," *American Economic Journal: Applied Economics*, 2019, *11* (1), 151–72.

Chen, Te-Ping and Ray A. Smith, "American Workers Are Burned Out, and Bosses Are Struggling to Respond," *Wall Street Journal*, 2021. [Online; accessed 2022-03-19].

Chen, Yan, F Maxwell Harper, Joseph Konstan, and Sherry Xin Li, "Social comparisons and contributions to online communities: A field experiment on MovieLens," *American Economic Review*, 2010, *100* (4), 1358–98.

Chevalier, Judith A and Dina Mayzlin, "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research*, 2006, *43* (3), 345–354.

Clark, Andrew E, "What really matters in a job? Hedonic measurement using quit data," *Labour Economics*, 2001, *8* (2), 223–242.

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

_ , "Four decades of the economics of happiness: Where next?," *Review of Income and Wealth*, 2018, *64* (2), 245–269.

— , **Nicolai Kristensen, and Niels Westergård-Nielsen**, "Job satisfaction and co-worker wages: Status or signal?," *The Economic Journal*, 2009, *119* (536), 430–447.

Coffman, Lucas C, Clayton R Featherstone, and Judd B Kessler, "Can social information affect what job you choose and keep?," *American Economic Journal: Applied Economics*, 2017, *9* (1), 96–117.

Commons, John R, "American Shoemakers, 1648–1895 A Sketch of Industrial Evolution," *Quarterly Journal of Economics*, 1909, *24* (1), 39–84.

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

Coviello, Decio, Erika Deserranno, Nicola Persico, and Paola Sapienza, "Effect of Mood on Workplace Productivity," *Mimeo*, 2018.

Cullen, Zoë and Ricardo Perez-Truglia, "How much does your boss make? the effects of salary comparisons," *Journal of Political Economy*, 2020, *Forthcoming.*

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

Delfgaauw, Josse, Robert Dur, OA Onemu, and Joeri Sol, "Team Incentives, Social Cohesion, and Performance: A Natural Field Experiment," *Management Science*, 2022, *68* (1), 230–256.

Dellarocas, Chrysanthos, "The digitization of word of mouth: Promise and challenges of online feedback mechanisms," *Management science*, 2003, *49* (10), 1407–1424.

DellaVigna, Stefano and Devin Pope, "What Motivates Effort? Evidence and Expert Forecasts," *Review of Economic Studies*, 2017, *85* (2), 1029–1069.

— **and M Daniele Paserman**, "Job search and impatience," *Journal of Labor Economics*, 2005, *23* (3), 527–588.

— , **Attila Lindner, Balázs Reizer, and Johannes F. Schmieder**, "Reference-Dependent Job Search: Evidence from Hungary," *Quarterly Journal of Economics*, 2017, *132* (4), 1969–2018.

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

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

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

Diener, E., S.J. Heintzelman, K. Kushlev, L. Tay, D. Wirtz, L.D. Lutes, and S. Oishi, "Findings All Psychologists Should Know From the New Science on Subjective Well-Being„" *Canadian Psychology/Psychologie Canadienne*, 2017, *58*, 87–104.

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

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker, "Referral-based job search networks," *The Review of Economic Studies*, 2016, *83* (2), 514–546.

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

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

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

Erez, Amir and Alice M Isen, "The influence of positive affect on the components of expectancy motivation.," *Journal of Applied Psychology*, 2002, *87* (6), 1055–1067.

Erickson, Rebecca J and Amy S Wharton, "Inauthenticity and depression: Assessing the consequences of interactive service work," *Work and occupations*, 1997, *24* (2), 188–213.

Estrada, Carlos A, Alice M Isen, and Mark J Young, "Positive affect facilitates integration of information and decreases anchoring in reasoning among physicians," *Organizational behavior and human decision processes*, 1997, *72* (1), 117–135.

Faberman, R Jason and Marianna Kudlyak, "The intensity of job search and search duration," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 327–57.

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

Fernandez, Roberto M, Emilio J Castilla, and Paul Moore, "Social capital at work: Networks and employment at a phone center," *American journal of sociology*, 2000, *105* (5), 1288–1356.

Filippas, Apostolos, John J. Horton, and Joseph M. Golden, "Reputation Inflation," *NBER Working Paper 25857*, 2019.

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

_ , "Why do lay people believe that satisfaction and performance are correlated? Possible sources of a commonsense theory," *Journal of Organizational behavior*, 2003, *24* (6), 753–777.

Fisher, V.E. and J.V. Hanna, *The dissatisfied worker*, New York: Macmillan, 1931.

Flory, Jeffrey A, Andreas Leibbrandt, and John A List, "Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions," *Review of Economic Studies*, 2015, *82* (1), 122–155.

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

Freeman, Richard, "Job Satisfaction as an Economic Variable," *American Economic Review*, 1978, *68* (2), 135–141.

Frey, Bruno S and Stephan Meier, "Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment," *American Economic Review*, 2004, *94* (5), 1717–1722.

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

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

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

Gee, Laura K, "The more you know: information effects on job application rates in a large field experiment," *Management Science*, 2019, *65* (5), 2077–2094.

_ **and Michael J Schreck**, "Do beliefs about peers matter for donation matching? Experiments in the field and laboratory," *Games and Economic Behavior*, 2018, *107*, 282–297.

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

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

Graddy, Kathryn and Carl Lieberman, "Death, bereavement, and creativity," *Management Science*, 2018, *64* (10), 4505–4514.

Graham, Carol, Kate Laffan, and Sergio Pinto, "Well-being in metrics and policy," *Science*, 2018, *362* (6412), 287–288.

Green, Francis, "Well-being, job satisfaction and labour mobility," *Labour Economics*, 2010, *17* (6), 897–903.

Green, T. Clifton, Ruoyan Huang, Quan Wen, and Dexin Zhou, "Crowdsourced employer reviews and stock returns," *Journal of Financial Economics*, 2019, *134* (1), 236–251. ISBN: 0304-405X publisher: Elsevier.

Groth, Markus and Alicia Grandey, "From bad to worse: Negative exchange spirals in employee–customer service interactions," *Organizational Psychology Review*, 2012, *2* (3), 208–233.

Hall, P. A. and D. Soskice, "An Introduction to Varieties of Capitalism," in P. A. Hall and D. Soskice, eds., *Varieties of Capitalism. The Institutional Foundations of Comparative Advantage*, Oxford: Oxford University Press, 2001, pp. 1–68.

Hampson, E. and A. Jacob, "Mental health and employers: Refreshing the case for investment," *Deloitte*, 2020, *https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/consultancy/deloitte-uk-mental-health-and-employers.pdf.*

Harrell, Frank E, *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*, Springer, 2015.

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

Hedblom, Daniel, Brent R Hickman, and John A List, "Toward an Understanding of Corporate Social Responsibility: Theory and Field Experimental Evidence," *NBER Working Paper No. 26222*, 2019.

Helmers, Christian, Pramila Krishnan, and Manasa Patnam, "Attention and saliency on the internet: Evidence from an online recommendation system," *Journal of Economic Behavior & Organization*, 2019, *161*, 216–242.

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

_ , *Workers' emotions in shop and home: A study of industrial workers from the psychological and physiological standpoint*, Philadelphia, PA: University of Pennsylvania Press, 1932.

Hersey, R.B., *Workers' emotions in shop and home: A study of industrial workers from the psychological and physiological standpoint*, Philadelphia, PA: University of Pennsylvania Press, 1932.

Hill, Amelia, "'Commercial no-brainer': why the role of happiness officer is taking off," *The Guardian*, 2022, *https://www.theguardian.com/lifeandstyle/2022/feb/25/commercial-no-brainer-why-role-of-happiness-officer-is-taking-off.*

Hill, Andrew, "Six ways to improve the wellbeing of workers," *Financial Times*, 2019. [Online; accessed 2022-03-19].

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

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

Hoppock, R., "Job satisfaction of psychologists," *Journal of Applied Psychology*, 1937, *21*, 300–303.

Horton, John J and Prasanna Tambe, "Labor economists get their microscope: Big data and labor market analysis," *Big data*, 2015, *3* (3), 130–137.

Houser, J.D., *What the employer thinks: Executives' attitudes toward employees*, Cambridge, MA: Harvard University Press, 1927.

Hsieh, Chih-Wei, "Burnout among public service workers: The role of emotional labor requirements and job resources," *Review of Public Personnel Administration*, 2014, *34* (4), 379–402.

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

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

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

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

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

_ , Kimberly A Daubman, and Gary P Nowicki, "Positive affect facilitates creative problem solving," *Journal of Personality and Social Psychology*, 1987, *52* (6), 1122.

Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer, "Worker Beliefs About Rents and Outside Options," *NBER Working Paper 29623*, 2021.

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

Johnston, Robert J, Kevin J Boyle, Wiktor Adamowicz, Jeff Bennett, Roy Brouwer, Trudy Ann Cameron, W Michael Hanemann, Nick Hanley, Mandy Ryan, Riccardo Scarpa et al., "Contemporary guidance for stated preference studies," *Journal of the Association of Environmental and Resource Economists*, 2017, *4* (2), 319–405.

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

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

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

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

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

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

Kim, Younjun, Catherine L Kling, and Jinhua Zhao, "Understanding behavioral explanations of the WTP-WTA divergence through a neoclassical lens: Implications for environmental policy," *Annual Review of Resource Economics*, 2015, *7* (1), 169–187.

Kirchsteiger, Georg, Luca Rigotti, and Aldo Rustichini, "Your morals might be your moods," *Journal of Economic Behavior and Organization*, 2006, *59* (2), 155–172.

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

Knowles, Vikki, "If you're happy and you know it ... become a chief happiness officer," *The Guardian,*, 2015, (https://www.theguardian.com/careers/2015/jul/13/if-youre-happy-and-you-know-it-become-a-chief-happiness-officer).

Kornhauser, Arthur W and Agnes A Sharp, "Employee attitudes: Suggestions from a study in a factory," *Personnel Journal*, 1932, *10*, 393–404.

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

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

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

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

Krueger, Alan B and Andreas I Mueller, "A contribution to the empirics of reservation wages," *American Economic Journal: Economic Policy*, 2016, *8* (1), 142–79.

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

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

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

Lacetera, Nicola, Devin G Pope, and Justin R Sydnor, "Heuristic thinking and limited attention in the car market," *American Economic Review*, 2012, *102* (5), 2206–36.

Landy, F.W., *The psychology of work behavior*, 3rd ed., Homewood, IL: Dorsey Press, 1985.

Lawler, Edward E, *Motivation in work organizations.*, Monteray, CA: Brooks/Cole Publishing, 1973.

_ **and Lyman W Porter**, "The effect of performance on job satisfaction," *Industrial relations*, 1967, *7* (1), 20–28.

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

_ **and Kathryn L Shaw**, "Personnel economics: The economist's view of human resources," *Journal of Economic Perspectives*, 2007, *21* (4), 91–114.

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

Levy-Garboua, L., C. Montmarquette, and V. Simonnet, "Job satisfaction and quits," *Labour Economics*, 2007, *14*, 2–251.

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

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

Lippman, Steven A and John J McCall, "The economics of job search: A survey," *Economic inquiry*, 1976, *14* (2), 155–189.

List, John A, "Field experiments: a bridge between lab and naturally occurring data," *The BE Journal of Economic Analysis & Policy*, 2007, *5* (2).

_ , "Non est Disputandum de Generalizability? A Glimpse into The External Validity Trial," *NBER Working Paper 27535*, 2020.

_ **and Fatemeh Momeni**, "When corporate social responsibility backfires: Evidence from a natural field experiment," *Management Science*, 2021, *67* (1), 8–21.

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

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

Luca, Michael, "Reviews, reputation, and revenue: The case of Yelp.com," *Harvard Business School Working Paper*, 2016.

MacGregor, Douglas, *The human side of enterprise*, Vol. 21, New York McGraw-Hill, 1960.

Maestas, Nicole, Kathleen J Mullen, David Powell, Till Von Wachter, and Jeffrey B Wenger, "The value of working conditions in the united states and implications for the structure of wages," *NBER Working Paper 25204*, 2018.

Malcolm, Sarah, "If You Want To Be More Productive At Work, Get Happy," *Forbes*, 2021.

Marinescu, Ioana and Daphné Skandalis, "Unemployment insurance and job search behavior," *The Quarterly Journal of Economics*, 2021, *136* (2), 887–931.

Marinescu, Ioana E, "Job search monitoring and assistance for the unemployed," *IZA World of Labor*, 2017.

Mas, Alexandre and Amanda Pallais, "Valuing Alternative Work Arrangements," *American Economic Review*, 2017, *107* (12), 3722–3759.

Maslow, Abraham Harold, "A theory of human motivation.," *Psychological review*, 1943, *50* (4), 370.

McCall, John Joseph, "Economics of information and job search," *The Quarterly Journal of Economics*, 1970, pp. 113–126.

McGee, Andrew D, "How the perception of control influences unemployed job search," *Industrial and Labor Relations Review*, 2015, *68* (1), 184–211.

McMurry, R.N., "Efficiency, work satisfaction and neurotic tendency," *Personnel Journal*, 1932, *11*, 201–210.

Melián-González, Santiago, Jacques Bulchand-Gidumal, and Beatriz González López-Valcárcel, "New evidence of the relationship between employee satisfaction and firm economic performance," *Personnel Review*, 2015, *44* (6), 906–929.

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

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

Moen, Phyllis, Erin L Kelly, Wen Fan, Shi-Rong Lee, David Almeida, Ellen Ernst Kossek, and Orfeu M Buxton, "Does a flexibility/support organizational initiative improve high-tech employees' well-being? Evidence from the work, family, and health network," *American Sociological Review*, 2016, *81* (1), 134–164.

Mortensen, Dale T, "Job search, the duration of unemployment, and the Phillips curve," *The American Economic Review*, 1970, *60* (5), 847–862.

_ , "Job search and labor market analysis," *Handbook of labor economics*, 1986, *2* (15), 849–919.

_ **and Christopher A Pissarides**, "Job creation and job destruction in the theory of unemployment," *Review of Economic Studies*, 1994, *61* (3), 397–415.

Myers, D.G., D.G. Myers, and E. Diener, "The pursuit of happiness," *Psychological Science*, 1993, *6*, 1 19.

Nagaraj, Abhishek and Henning Piezunka, "How Competition Affects Contributions to Open Source Platforms: Evidence from OpenStreetMap and Google Maps," *Working Paper*, 2020.

Neve, Jan-Emmanuel De and George Ward, "Happiness at work," in Jeffrey Sachs, John F. Helliwell, and Richard Layard, eds., *World Happiness Report*, Columbia Earth Institute, 2017.

OECD, *OECD Guidelines on Measuring Subjective Wellbeing*, OECD Publishing, 2013.

Osterman, Paul, Thomas A Kochan, Richard M Locke, and Michael J Piore, *Working in America: A blueprint for the new labor market*, MIT Press, 2001.

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

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

Piore, M, "Revitalizing industrial relations," in Charles J Whalen, ed., *New Directions in the Study of Work and Employment: Revitalizing Industrial Relations as an Academic Enterprise*, 2008, pp. 163–172.

Purcell, John, "Disengaging from engagement," *Human Resource Management Journal*, 2014, *24* (3), 241–254.

Reimers, Imke and Joel Waldfogel, "Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings," *American Economic Review*, 2021, *111* (6), 1944–71.

Reynolds, Lloyd George, *The structure of labor markets: Wages and labor mobility in theory and practice*, Westport, Conn: Greenwood Press, 1951.

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

Rosen, Sherwin, "The theory of equalizing differences," *Handbook of labor economics*, 1986, *1*, 641–692.

Rothbard, Nancy P. and Steffanie L. Wilk, "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*, 2011, *54* (5), 959–980.

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

Schwab, D.P. and L.L. Cummings, "Theories of performance and satisfaction: A review," *Industrial Relations*, 1970, *9*, 408–430.

Segel, Liz Hilton, "The priority for workplaces in the new normal? Wellbeing," *McKinsey & Company*, 2021. [Online; accessed 2022-02-19].

Seppälä, Emma and Kim Cameron, "Proof that positive work cultures are more productive," *Harvard Business Review*, 2015, *12* (1), 44–50.

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

Sockin, Jason, Aaron Sojourner, and Evan Starr, "Non-Disclosure Agreements and Externalities from Silence," *SSRN Working Paper 3900285*, 2021.

_ **and** _ , "What's the Inside Scoop? Addressing Workers' Information Problems with Firm Reputation," *Addressing Workers' Information Problems with Firm Reputation (May 4, 2020)*, 2020.

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

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

Stern, Scott, "Do scientists pay to be scientists?," *Management Science*, 2004, *50* (6), 835–853.

Stinchcombe, Arthur L., "Social structure and organizations," in J.G. March, ed., *Handbook of Organizations*, 1965, pp. 142–193.

Super, D.E., "Occupational level and job satisfaction," *Journul of Applied Psychology*, 1939, *23*, 547–564.

Symitsi, Efthymia, Panagiotis Stamolampros, and George Daskalakis, "Employees? online reviews and equity prices," *Economics Letters*, 2018, *162*, 53–55.

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

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

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

Tucker, Catherine and Juanjuan Zhang, "How does popularity information affect choices? A field experiment," *Management Science*, 2011, *57* (5), 828–842.

Vroom, Victor H, *Work and motivation*, New York, NY: John Wiley and Sons, 1964.

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

Webb, Sidney and Beatrice Webb, *Industrial democracy*, Longmans, Green and Company, 1897.

Wiswall, Matthew and Basit Zafar, "Preference for the workplace, investment in human capital, and gender," *The Quarterly Journal of Economics*, 2018, *133* (1), 457–507.

Wright, T.A. and D.G. Bonett, "Job satisfaction and psychological well-being as non-additive predictors of workplace turnover," *Journal of Management*, 2007, *33*, 141–160.

Wright, Thomas A, "The emergence of job satisfaction in organizational behavior: A historical overview of the dawn of job attitude research," *Journal of Management history*, 2006.

_ **and Russell Cropanzano**, "The role of psychological well-being in job performance:: a fresh look at an age-old quest," *Organizational Dynamics*, 2004, *33* (4), 338–351.

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

Zivin, Joshua Graff and Matthew Neidell, "The Impact of Pollution on Worker Productivity," *American Economic Review*, 2012, *102* (7), 3652–3673.