The Balancing Act:
Corporate Norms and Practices that Affect Work-Life Balance

Heather A. Haveman and **Jasmine M. Sanders**[1]

University of California, Berkeley
haveman@berkeley.edu and jmsanders@berkeley.edu

17 November, 2022

Abstract

Balancing work and life outside of work is a challenge for many employees. Lack of work-life balance hampers employee performance and health, and prompts exit. It has especially negative effects on female employees because women do more domestic labor than men. We focus on firms in the tech sector and capture norms and practices concerning work-life balance by analyzing employees' descriptions of their firms using natural-language-processing techniques. We develop and test arguments about who discusses work-life balance. One-quarter of employees discuss work-life balance. Supporting our arguments, these are mostly women in their 30s, and women in privately owned and smaller firms. Second, we develop and test arguments about opinions about work-life balance. Most tech employees express positive opinions about work-life balance. But, contrary to expectations, there are no gender differences. Analyzing a random sample of reviews for other themes revealed a common concern that firms did not have uniform work-life balance policies; instead, managers had considerable discretion to approve or deny employee requests to deal with life outside work. Because this is an important issue for many employees, male and female alike, all firms would gain from standard policies that offer cafeteria-style work-life-balance benefits that fit employees' personal circumstances.

---

[1] We thank Glassdoor.com for access to the data, and Sameer Srivastava for facilitating an introduction to the firm. For funding, we acknowledge the Center for Equity, Gender, and Leadership at the Haas School of Business, and the Institute for Research on Labor and Employment. We thank Will Rathje for his research assistance and advice, and Minori Jaggia and Emily Guo for their programming efforts, as well as earlier work by Lily Bhattarjee, Grace Kull, and Sandeep Sainath. For helpful comments, we thank seminar participants at the University of Michigan, the American Sociological Association annual meetings, the Stockholm School of Economics, and the Center for Equity, Gender, and Leadership.

Work-life balance, the extent to which employees can thrive at work *and* enjoy their lives outside work, is difficult for many employees because of managerial expectations that employees be available at all hours, and the practices that reflect and reinforce these expectations (Kelly and Moen, 2020).[2] Insisting on around-the-clock availability stresses employees because it erodes boundaries between work and life outside work, creating conflict between the two spheres of life (Nippert-Eng, 1996; Moen, 2003). Employees are less likely to seek jobs in firms where they perceive that work-life balance is difficult to achieve, less likely to perform well, more likely to be dissatisfied with their jobs, and more likely to exit (Kelly and Moen, 2020). Such employee reactions impose considerable costs on firms as well as employees (e.g., James, 2014), especially when labor markets are tight.

Corporate norms and practices that hinder work-life balance have especially pernicious effects on female employees because women do more domestic labor than men, so working long hours is more difficult for women (Hochschild, 1989; Jacobs and Gerson, 2004; Moen, 2003; Kelly and Moen, 2020). Thus, norms and practices that hinder work-life balance have gendered effects on hiring, performance, satisfaction, and exit (Pedulla and Thébaud, 2015). But not all firms are alike. Instead, some make it easier than others for workers to balance work and family responsibilities (Kelly and Moen, 2020). Therefore, we investigate for whom work-life balance is most salient and how positive or negative their opinions about work-life balance are: *which workers, in which firms*.

To gain empirical leverage, we analyze the language employees use when they describe their firms. Language has shadings of beliefs and values baked into it; it is a medium for expressing fundamental assumptions and values (Barley, 1983). For example, studying the language of performance evaluations reveals how gender stereotypes infuse managers' perceptions of workers (Correll, Weisshaar, Wynn, and Wehner, 2020). Language also reveals

[2] We generally use the term "work-life balance" to describe this goal. It includes not just work-family balance, but also spending time with friends; political and social causes; religious life; exercise, sports, entertainment, and hobbies; and furthering one's education.

formal corporate policies and everyday practices. Using natural-language-processing (NLP) techniques, we analyze employee descriptions of U.S. tech firms from 2014 to 2020 taken from Glassdoor.com, an online job-search platform. We create NLP-derived measures of the salience of and opinions about work-life balance.

We begin by asking *who talks about work-life balance*? Our baseline assumption is that workers talk about things that matter to them. Both female workers and workers with young children who need considerable parental effort (those in their 30s and early 40s) are particularly sensitive to work-life balance, so we expect the salience of work-life balance to be greatest among female workers in this age range. We also expect that context matters. Specifically, we expect gender differences in discussing work-life balance to be less in publicly traded and larger firms because they tend to be more formalized, more visible, and better resourced than privately owned and smaller firms. Next, we probe *what do employees talk about when they talk about work-life balance*? We develop and test hypotheses about how employee characteristics and workplace context affect opinions about work-life balance, based on arguments similar to those about the salience of work-life balance. We then analyze a random sample of the data in depth, to uncover other themes in employee discourse about work-life balance and assess gender differences.

We study employing organizations in the tech sector: industries with computing technology at their core, including hardware, software, network systems, internet services, and video games. Tech firms are often touted as offering good work-life balance relative to firms in other sectors (Thomas, 2021), but they have long had a gender problem, with women underrepresented in the technical and managerial ranks (Rangarajan, 2018), and loud complaints about discrimination and harassment (Fowler, 2017; Kolhatkar, 2017; Chang, 2018). Given the importance of work-life balance for women's career success, it is not clear what we will find.

We find considerable variation in tech workers' discourse about work-life balance. Supporting our arguments, we find that women in the middle of the age range are most likely to discuss work-life balance, as are those in smaller and privately owned firms. Although most tech workers expressed positive sentiment about work-life balance, there is considerable variation. Contrary to expectations, we find that women and men express similar opinions about work-life balance, although we find some of the expected differences by age and between publicly traded and privately owned firms. Most workers offer purely factual statements about work-life balance, but some reveal concerns about lack of standard work-life-balance policies that apply to all employees equally, and hence their reliance upon managerial discretion. These findings fit with small-N, interview-based research that revealed flexible work arrangements to be perks that workers must negotiate with their managers (Kelly and Kalev, 2014).

Our analysis reveals considerable variation in the salience of this topic across workers (its salience was greater for female employees and employees who are (based on age) most likely to have young children) but it was relatively common among men and employees of all ages. Work-life balance was salient among workers in all firms, but especially women in small and privately owned firms. These findings suggest that *all* employers (large and small, public and private) might gain from designing cafeteria-style work-life balance policies, with workers able to choose the options that fit their particular situations. Moreover, our analysis indicates that these policies should be formalized and applied to all employees, without leeway for managerial discretion.

Theory

Gender is the biggest divide in many societies, including the U.S. For over 40 years, sociologists and management scholars have demonstrated that employing organizations can both generate and reduce gender inequality (e.g., Kanter, 1977; Bielby and Baron, 1986; Gorman, 2005; Castilla, 2015; Kelly and Moen, 2020). Such work has focused on organizational structures (e.g., Baron, Davis-Blake, and Bielby, 1986; Reskin, 1993) and policies (e.g., Kelly and Dobbin, 2009; Dobbin, Schrage, and Kalev, 2015; Castilla, 2015). Less studied, especially at scale, are the norms that shape how work is done and evaluated, and the on-the-ground practices that reflect and reinforce these norms. In this section, we focus on workplace norms and practices that affect how well employees can balance work and life outside of work, and the gendered implications of those norms and practices. We then discuss the way language – how employees describe their firms in their own words – can give us insight into workplace norms and practices. Finally, we develop hypotheses concerning how workers' personal characteristics and work contexts affect their attention to and views on work-life balance.

Work-life balance: Workplace norms and practices

For three decades, journalists and academics have documented the rise of workplace norms and practices that hamper employees' ability to balance work and life outside work (Venkatesh and Vitalari, 1992; King, 1997; Jacobs and Gerson, 2004; Correll, Kelly, O'Connor, and Williams, 2014; Pedulla and Thébaud, 2015; Kelly and Moen, 2020). These are especially prominent in some industries such as finance, where many employees are expected to work very long hours in exchange for very high compensation. But long hours are also prevalent in academia, law, management consulting, medicine, information technology, and startups in many industries, and they hit managerial and professional occupations especially hard (Schor, 1992; Cha and Weeden, 2014; Kelly and Moen, 2020).

The rise of work-around-the-clock norms and practices has been made possible by the development of new electronic technologies that facilitate communication, starting with workplace email in the late 1980s, text messaging on smartphones in the late 1990s, and shared files on cloud servers like Google docs and sheets in the 2000s. The most recent developments include web meeting platforms like Zoom and Microsoft Teams, and collaboration applications like Slack and Trello. Together, these communication technologies have changed the nature of work, increasingly blurring the boundary between work and life outside work (Kelly and Moen, 2020).

These technologies have become widespread because they fit with expectations of what work is, how it should be done, and who should do it; in turn, these expectations drive the development and maintenance of formal policies and everyday workplace practices, including those related to job design, recruiting, training, compensation, and scheduling. Most germane is the belief that the "ideal worker" is a man who can devote all of his time to work because his spouse manages their domestic life (Acker, 1990; Britton, 2000). The male ideal-worker belief creates the expectation that employees are available at all hours and on all days, and the practices that reflect and reinforce this expectation, such as valuing "face time," scheduling meetings late in the day, assigning tasks to be done overnight or on the weekend, and sending messages around the clock (Perlow, 2012; Boushey, 2016; Padavic, Ely, and Reid, 2020). Such norms and practices are also driven by masculinity competition cultures, which demand that employees put work first and let nothing outside their workplace interfere with work (Berdahl, Cooper, Glick, Livingston, and Williams, 2018).

How to balance work and life outside of work is deeply cognitive and cultural. To bring order to their lives, people draw mental boundaries around activities, between themselves and other people, and around physical objects and spaces (Zerubavel, 1991). Although people vary in the degree to which they prefer to integrate or segment work and family, most recognize them as distinct spheres of life (Nippert-Eng, 1996). People generally strive to separate their roles in the two spheres so they can devote their full attention to the sphere they occupy at a particular time.[3] In general, spillovers between work and home life obliterate those boundaries, undermining the distinction between the cultural schemas associated with each sphere of life. Integrating work and home life makes it more difficult to focus on either and so causes stress (Moen, 2003). Employees are less likely to seek jobs in firms where they perceive

[3] The two spheres of life became separated during the industrial revolution, when work moved out of homes and into mines and factories. In the process, work life came to be conceptualized as public and masculine, while home life came to be conceptualized as private and feminine (Davies and Frink, 2014) – even though some women always worked for pay outside the home and women have over the past 50 years become almost as likely to do so as men.

that work-life balance is difficult to achieve, less likely to perform well, more likely to be dissatisfied with their jobs, more likely to be absent from work, and more likely to exit (Dalton and Mesch, 1990; Anderson, Coffey, and Byerly, 2002; McNall, Masuda, and Nicklin, 2009; Kelly and Moen, 2020; Choper, Schneider, and Harknett, 2022). Such employee dynamics impose considerable costs on firms (e.g., James, 2014), especially when labor markets are tight.

Lack of work-life balance has especially deleterious effects on female employees because women do more domestic labor than men (Hochschild, 1989; Jacobs and Gerson, 2004; Moen, 2003; Kelly and Moen, 2020), so female employees find it more difficult than their male counterparts to work around the clock. If women with children do put in long hours at work, they are seen as bad mothers (Cuddy, Fiske, and Glick, 2004). Moreover, women find it more difficult to reconcile their roles as employees with their roles in family life (Mogenroth, Ryan, Rink, and Begeny, 2021).

We recognize, however, that firms vary greatly: some are more willing to help workers balance work and family responsibilities (Kelly and Moen, 2020), which can reduce gender inequality. For example, firms that accommodate work-life balance reduce the salience of gender at work and so reduce gender discrimination (Stainback, Ratliff, and Roscigno, 2011). To investigate heterogeneity in norms and practices concerning work-life balance, we investigate who is most affected: not just women more than men, but which women, in which firms.

Language provides insights into workplace norms and practices

To probe the workplace norms and the related practices that hamper work-life balance, we focus on the language employees use when they describe where they work. Such language reveals employees' understandings of their firms and their experiences at work – their mental models (Carley and Palmquist, 1992) or cognitive schemas (DiMaggio, 1997; Hunzaker and Valentino, 2019). Language allows people to categorize the world around them in symbolic terms that "have meaning, are cues to behavior, and organize behavior" (Stryker, 1980: 56). For instance, studying the language used in job postings reveals to prospective applicants whether women are expected to fit into those jobs (Gorman, 2005), while the language used in performance evaluations reveals how gender stereotypes infuse managers' perceptions of workers (Correll et al., 2020). Most germane to this analysis is that language is a medium for expressing how firms actually operate, as well as expectations about how they should operate.

The *content* of language indicates what matters and how: words and phrases that are common are associated with central cultural elements, while words and phrases that are rare (or missing) are associated with peripheral (or foreign) cultural elements (Whorf, 1956; Sapir, 1958). That is not to say that all central cultural elements are valued; indeed, some that are highly central to an organization's operations may be contested or disdained by employees. If so, they will be prevalent *and* take the form of complaints about the current workplace norms and practices. For example, if employees frequently complain about overwork and their struggles to balance job requirements with their lives outside work, then the expectation that work is more important than all else is *both* central to employees' mental models of their firms *and* contested by them. Values, regardless of their centrality, can also be contested if employees disagree about them, some expressing positive views, others negative views.

The meaning of language, including its emotional valence, derives from its *context* – nearby grammatical elements such as words and punctuation marks (Harris, 1954; Firth, 1957). For example, the word "bank" means the edge of a river when it is surrounded by words such as fish, rapids, boat, and swim. In contrast, "bank" means a financial institution when it is surrounded by words such as save, borrow, collateral, and assets. In particular, we can analyze the context of the language employees use to measure its emotional valence, as we explain in the research design section.

Who talks about work-life balance?

We expect people to vary in their attention to and attitudes toward norms and practices related to work-life balance depending on their personal circumstances and the contexts in which they work. Some of our arguments are based on generally applicable theory and broadbased empirical evidence, but some incorporate knowledge of our research site, tech workplaces. To begin, we expect that employees are more likely to discuss work-life balance when it is salient to them – when it matters to them. So below, we focus on two employee characteristics – gender and age – that have strong effects on the salience of work-life balance.

Gender. This is the most important personal characteristic affecting work-life balance. Women shoulder most of the burden for domestic labor, not just for childcare and housework (Hochschild, 1989; Jacobs and Gerson, 2004), but also for planning and arranging the details of family life: who must be where and when, what must be done to meet familial deadlines (Daminger, 2019). Women are also overrepresented among frontline workers, who must be physically present to do their jobs and cannot leave work to handle personal matters. While women constitute just under half of the workforce, they represent nearly two-thirds of workers in frontline jobs, including retail sales, food service, health care, cleaning, and child care (Rho, Brown, and Fremstad, 2020).[4]

Our analysis focuses on the tech sector, where the core tasks are technical, so core jobs are technical. These jobs have tight labor markets, with the number of open positions far greater than the number of qualified candidates (O'Conner, 2011; Loten, 2019; Statistica, 2022). For these reasons, technical jobs drive the design of work, including norms about work-life balance. Women are vastly outnumbered by men in technical jobs (McKinsey and Company, 2018; Rangarajan, 2018), so men tend to be the ideal workers for technical jobs and those jobs tend to be perceived as male-typed. Women in technical jobs are often viewed with skepticism, even suspicion (Wynn and Correll, 2018; Luhr, 2020), and so have to work harder than men to demonstrate that they can succeed. Extra effort is most clearly manifested by sheer presence ("face time"): working long hours and responding (rapidly) to work-related communications, no matter when they arrive. Finally, in tech as in other sectors, women tend

---

[4] Ideally, we would test hypotheses about job type (e.g., technical vs. nontechnical, managerial vs. nonmanagerial), but we do not have clean data about job type.

to be overrepresented in the lowest-level frontline positions, such as receptionists and administrative assistants, which usually require workers' physical presence to get tasks done.

Overall, work-life balance should be more salient to women than men. Thus, women should be more likely to mention work-life balance when they talk about their firms. This is especially true for frontline workers, who have the most difficulty getting time off to deal with personal chores and emergencies, and for female technical workers, who have to strive to demonstrate that they can perform as well as men. Therefore we predict:

> **Hypothesis 1**: Women are more likely than men to mention work-life balance when they talk about their employing organizations.

Age. The relationship between employee age and concern about work-life balance is complicated. Older employees are more likely to have children – the biggest source of work-life tension (Kelly and Moen, 2020) – than their younger counterparts. It is workers in the middle of the age range – those in their 30s and early 40s – who are likely to face the most demands on their time for domestic labor. Younger workers are likely to be childless, while older workers are likely to have older children (U.S. Bureau of Labor Statistics, 2019). Older children spend considerable time in school and after-school activities; while at home, older children are more self-sufficient than younger children. Given age-related differences in family composition, we expect that workers in the middle of the age range suffer the most conflict between their responsibilities at work and to their families, while those who are younger and older experience less work-family conflict. If so, work-life balance will be most salient to workers in the middle of the age range. Therefore, we predict:

> **Hypothesis 2**: Workers in the middle of the age range are more likely than younger or older ones to mention work-life balance when they talk about their employing organizations.

Age and gender are interrelated. Employee gender and age are likely to accentuate each other's relationships with employee discourse about work-life balance. Female workers bear more of the burden of domestic labor than their male counterparts, and workers in the middle of the age range are the most likely to have children who need the most care and supervision. As a result, female workers in the middle of the age range are more likely to bear the burden of domestic labor, compared with their younger and older female colleagues and with their male colleagues in the middle of the age range. Therefore, we predict:

> **Hypothesis 3**: Any observed gender gap in mentioning work-life balance is larger for workers in the middle of the age range than for younger or older workers.

Context matters: firm ownership. There are fundamental differences between privately owned and publicly traded firms that can affect relationships between gender and employee discourse about work-life balance. First, going public gives firms access to new sources of capital, dramatically increasing their financial resources. As a result, compared to privately owned firms, publicly traded firms tend to have more slack resources that they can choose to devote to programs and policies that could help employees balance work and life outside work, such as gym memberships and childcare subsidies.

Second, publicly traded firms are much more visible than privately owned ones because they are legally required to report more on their operations and finances. Such reports are intended for current shareholders and government regulators (primarily the Securities and Exchange Commission), as well as the public, meaning investment advisors, potential shareholders, and potential employees. Heightened visibility drives publicly traded firms to adopt more formalized rules and procedures, including those concerning the scope (timing and location) of work, which may help employees achieve work-life balance. Such rules and procedures help by constraining potentially harmful managerial discretion (Reskin and McBrier, 2000); for example, by preventing managers from scheduling meetings outside standard work hours, denying employees leave to deal with family responsibilities, or refusing to allow employees to take accumulated vacation time. Such rules and procedures can also have affirmative impacts (rather than preventing negative ones); for example, by offering childcare subsidies.

Although publicly traded firms could commit more resources and implement more policies and programs to help employees balance work and life outside work, they may not do so because of a focus on meeting short-term profit and growth goals. But our expectation of a positive association between public ownership and resources, policies, and procedures to help employees achieve work-life balance is bolstered by research showing that publicly traded firms donate more to charities (e.g., Brammer and Millington, 2006). Corporate philanthropy has a positive social impact, so this study suggests that it is reasonable to propose a positive association between ownership type (public vs. private) and corporate actions that facilitate another positive social impact, namely employees' work-life balance.

Promoting work-life balance benefits all employees, but female employees especially, since women bear more of the burden of domestic labor. To the extent that publicly traded firms are more likely than privately owned firms to implement policies and procedures promoting work-life balance, we predict:

> **Hypothesis 4**: Any observed gender gap in mentioning work-life balance is smaller for workers in publicly traded firms than in privately owned firms.

Context matters: firm size. Organization size can affect relationships between gender and employee discourse about work-life balance in several ways. First, larger firms tend to be more formalized: they have more written rules and procedures than smaller firms (Blau and Scott, 1962; Blau and Schoenherr, 1971). Such rules and procedures can help employees indirectly by constraining managerial discretion (Reskin and McBrier, 2000) and so preventing work-life conflict, *and* because they can affirmatively facilitate work-life balance. Second, larger firms are more visible than smaller ones and so more sensitive to negative news – including employee complaints about work-life conflict posted to job-search platforms. Visibility makes larger firms more likely to help employees achieve work-life balance. Third, larger firms tend to have more slack resources that they can devote to paying for programs and policies that help employees achieve work-life balance. Large firms grow large because they performed well in the past; good performance generates slack resources (Cyert and March, 1963; Daniel, Lohrke, Fornaciari, and Turner, 2004).

Although larger firms could commit more resources and develop more formal policies and procedures to the issue of work-life balance, they may not do so. But our expectation is bolstered by research showing that larger firms are more likely to adopt civil-rights grievance procedures for employees (e.g., Edelman, 1990; Edelman, Uggen, and Erlanger, 1999), more likely to have corporate social responsibility policies (e.g., Gallo and Christiansen, 2011) and more likely to donate more to charities (e.g., Galaskiewicz, 1997; Brammer and Millington, 2006). Such behaviors generally have positive social impacts, so they suggest that it is reasonable to assume a positive association between organizational size and efforts to promote another social benefit, employees' work-life balance.

Promoting work-life balance benefits all employees, but especially female employees, since women bear more of the burden of domestic labor. To the extent that larger firms are more likely to implement policies and procedures promoting work-life balance, we expect:

> **Hypothesis 5**: Any observed gender gap in mentioning work-life balance is smaller for workers in larger firms than in smaller firms.

What does work-life balance mean? Work-life balance encompasses several different phenomena, including (but not limited to) taking care of domestic chores, children, and elders; having time to support social, religious, or political causes; being able to enjoy leisure pursuits like gardening, golf, hiking, theater, and book clubs; and participating in religious services and outreach. Therefore, it is important to distinguish among different interpretations of work-life balance. In previous research, the two most common interpretations have been (i) family responsibilities and (ii) time and location, meaning when and where work is done (e.g., Hochschild, 1997; Kelly and Moen, 2020). Some research also discusses quality of life and work-life balance in abstract terms (e.g., Elizur and Shye, 1990; Williams, 2018; Kelliher, Richardson, and Boiarintseva, 2019). In this section, we consider all three interpretations, and predict which employees are most likely to invoke which meaning(s) of this umbrella concept.

We begin with family responsibilities. As explained above, women bear the brunt of family and childcare tasks. Therefore, when employees discuss work-life balance, we expect family responsibilities to be more salient for women than for men:

> **Hypothesis 6a:** Women are more likely than men to mention *family responsibilities*.

To help employees balance work and life outside work, many companies have implemented flexible work arrangements that both women and men can take advantage of. Women, particularly those with young children, express stronger preferences for flexible work arrangements, like telecommuting, than men do (Mas and Pallais, 2017; Kelly and Moen, 2020). Recent experience with remote work during the Covid pandemic shows that despite the additional unpaid domestic labor that women have taken on when their jobs go remote, such as childcare, housework, and handling schedule disruptions and emergencies, women are still more likely than men to express a desire for flex-time and flex-location work opportunities (Lyttelton, Zang and Musick, 2022).[5] We expect, therefore:

> **Hypothesis 6b:** Women are more likely than men to mention *flexibility* concerning when and where work is done.

Last, consider abstract conceptions of work-life balance denoted by use of terms like “quality of life” and “work-life balance” itself. It is not clear how workers interpret these phrases – they could use them while conceiving of them in terms of family responsibilities, flexibility, or some other concept. Therefore, we make no prediction about who uses these abstract terms or in which contexts; we merely assess associations empirically.

What do employees say about work-life balance?

Opinions about work-life balance shape recruiting, employee performance, and retention (James, 2014; Kelly and Moen, 2020), thereby affecting employees, their families, and

---

[5] Yet women are penalized more than men for taking advantage of these supports, creating a “flexibility stigma.” Women who take advantage of flexible work accommodations are seen as less devoted to their jobs (Williams, Blair-Loy, and Berdahl, 2013; Padavic, Ely, and Reid, 2020); therefore, they are less likely to be promoted and are paid less (Glass, 2004).

firms. Recruiting and retention are especially important for firms operating in tight labor markets, such as tech. Accordingly, we pose three questions. First and most basically, we ask *how positive or negative are employee opinions about work-life balance*? Second, *what topics do employees associate with work-life balance?* In particular, we probe for gendered differences in discourse about work-life balance (Keene and Quadagno, 2004; Williams, 2018). Third, *which employees in which firms express more positive opinions about work-life balance?*

We can investigate employee opinions about work-life balance only when employees mention it – when it is salient. The relationship between the salience of work-life balance and opinions about it is complicated. It may be that the employees who care more about work-life balance are more likely to discuss it when they are concerned about it, which would suggest a *negative relationship* between salience and opinion. Instead, however, it may be that employees are more likely to discuss work-life balance when they are satisfied with it, especially when and where there is a perceived talent shortage, so employees may be motivated to promote their firms to job seekers, in order to help their firms and improve their own work lives. That would translate to a *positive relationship* between salience and opinion.[6] Because there are two opposing possibilities, we develop competing hypotheses about relationships between gender and age, on the one hand, and employee opinions about work-life balance, on the other:

> **Hypothesis 1a**: Women are more likely than men to express *negative* sentiments about work-life balance.

> **Hypothesis 1b**: Women are more likely than men to express *positive* sentiments about work-life balance.

---

[6] The relationship between salience and opinions is further complicated by our focus on the tech sector, where successful firms offer employees high pay and lavish perquisites to attract and retain scarce talent (Luckerson, 2014; Saiidi, 2016; Tian, 2016; Thompson, 2020). Most germane to this analysis is that many tech firms offer paid parental leave and subsidies for expenses related to childcare (e.g., Tian, 2016; Lotze, 2019). If pay and perks compensate for the lack of work-life balance, then there may be no strong relationship between the salience of work-life balance and sentiment toward it. Any relationship would appear only after partialling out the effects of compensation and perks, so we make sure to control for these factors.

> **Hypothesis 2a**: Workers in the middle of the age range are more likely than younger or older ones to express *negative* sentiment about work-life balance.

> **Hypothesis 2b**: Workers in the middle of the age range are more likely than younger or older ones to express *positive* sentiment about work-life balance.

Context moderates relationships between age and gender, on the one hand, and opinions about work-life balance, on the other. As we explained above, we expect work-life balance to be most salient to female workers with young children – those in the middle of the age range. Similarly, as we explained above, we expect work-life balance to be less salient to workers in older firms and publicly traded firms. Therefore we predict:

> **Hypothesis 3a**: Any observed gender gap in expressed sentiment about work-life balance is stronger for workers in the middle of the age range than for younger or older workers.

> **Hypothesis 4a**: Any observed gender gap in sentiment about work-life balance is weaker for workers in publicly traded firms than in privately owned firms.

> **Hypothesis 5a**: Any observed gender gap in sentiment about work-life balance is weaker for workers in larger firms than in smaller firms.

Research Methods

Research site: Tech firms

To investigate employee discourse about work-life balance, we study tech firms, meaning those operating in industries with computing technology at their core, such as computer hardware and software, networking systems, and video games.[7] Tech firms have long had a gender problem, with women underrepresented in the technical and managerial ranks. Beyond mere numbers, the ideal tech worker is male, so women find it difficult to fit into tech firms (Wynn and Correll, 2018; Luhr, 2020). Women in tech have complained loudly,

---

[7] Concretely, we studied firms in nine industries: computer hardware and software manufacturing, electronics manufacturing, enterprise software and network systems, information-technology services, internet-based services, internet-service providers, telecommunications manufacturing, telecommunications services, and video games. We selected those industries because these are the domain of the large publicly traded tech firms in the San Francisco Bay Area that dominate this sector (e.g., the Silicon Valley 150 [Lonergan, 2020]).

clearly, and frequently about discrimination and harassment, and the “bro” culture that permeates the tech sector (e.g., Fowler, 2017; Kolhatkar, 2017; Chang, 2018). Yet tech firms are often touted as offering better work-life balance than firms in other sectors (Thomas, 2021) because the supply of core tech employees often outstrips supply (O’Connor, 2011; Loten, 2019; Statistica, 2022). To attract scarce talent, many tech firms offer high compensation and excellent perks, some of which (e.g., subsidized child care, flextime) can help employees balance work and life outside work. In the end, it is not clear whether tech-firm employees will praise or complain about work-life balance.

Data sources

Our data consist of 948,785 employee reviews from Glassdoor.com from 2014 to 2020 for firms in the tech sector.[8] Glassdoor reviews contain ratings on a 1-5 scale. Overall ratings are required. Employees also have the option to provide ratings on specific topics: culture and values, diversity and inclusion, work-life balance, senior management, and career opportunities. The meat of these reviews is verbal descriptions of the pros and cons of their workplaces, plus advice to management. These descriptions are open-ended: employees present their own perceptions of and reactions to their workplaces. Figure 1 shows a typical recent review for Salesforce, a San Francisco-based tech firm that provides customer-relations software to large and small businesses. The bottom half of the figure shows the optional topical ratings.

[Figure 1 about here]

Glassdoor has several advantages for research on organizational culture and workplace practices. The company serves primarily as a job-search platform, so it attracts a large group of workers – about 17 million unique users every month. About half of job-seekers use Glassdoor (DeMers, 2014), so these reviews serve as signals to potential employees how they might feel

[8] As we explain below, missing data on several variables reduced the sample to 260,862 observations for the multivariate analysis of mentioning work-life balance, then to 65,943 for the multivariate analysis of sentiment about work-life balance.

about working for reviewed companies. Glassdoor has excellent coverage of many firms, especially large ones; for example, Apple has over 30,000 reviews on Glassdoor. Glassdoor allows people to review their employers anonymously, so reviews are not susceptible to bias stemming from fear of employer retribution (Marinescu et al., 2021). And Glassdoor reviews include both privately owned and publicly traded firms, which is unusual for large-N research.

Glassdoor reviews have three possible downsides, however. The first and second derive from sampling bias: rather than being written by a random sample of employees, the data are limited to those who contribute reviews. First, if people are more motivated to contribute reviews when they have strong emotional reactions to their employers, the data might be biased both positively *and* negatively. This would result in extreme bimodal distributions, with many one- and five-starred reviews. This is seen in review data from platforms like Yelp (retail businesses such as restaurants and plumbers) and Rotten Tomatoes (films and TV shows), where contributions are voluntary. But because Glassdoor has a "give-to-get" model that requires users to provide reviews before gaining unlimited access to data to aid their own job searches, these reviews are less likely to be biased in this way (Marinescu et al., 2021) and more likely to be representative of the experiences and perceptions of all employees. Indeed, the distribution of reviews in our sample is nearly uniform, with a slight positive skew: the mean is 3.48 out of 5.

Second, although reviews are anonymous, managers may encourage employees to contribute positive reviews. If so, some of what we observe may be biased. As mentioned above, there are slightly more positive reviews than negative ones. Given this, we conducted a robustness check by analyzing lower- and higher-scoring reviews separately, with lower meaning scores of 1 or 2 and higher meaning scores of 4 or 5. We discuss the results of this robustness check below.

The third downside may involve gender bias: reviews may disproportionately represent the perceptions and experiences of men. In our data, 38% of reviewers self-identified as male, 19% percent self-identified as female, and 42% did not report gender. Of those who reported gender, 33% were female, which is slightly more than the 28% of employees in the tech sector overall (McKinsey and Company, 2018). This suggests that these data do *not* disproportionately represent what male employees think.[9]

Measures

We used natural-language-processing (NLP) techniques to analyze employees' descriptions of their firms, based on the pros and cons sections of Glassdoor reviews. Our approach goes beyond most previous research on workplaces practices that used ethnographic methods (e.g., Barley, 1983; Eliasoph and Lichterman, 2003), in-depth interviews (e.g., Williams, 2018; Luhr, 2020), or surveys (e.g., Edelman, 1990; Reskin and McBrier, 2000; Dobbin, Schrage, and Kalev, 2015). Ethnographies and in-depth interviews provide rich details but cover only a small number of workers in one or a few organizations. Surveys can pick up variation across many workers in many organizations, but they rely on top-of-the head responses by survey participants, which may not mirror on-the-ground reality. In contrast, our approach captures on-the-ground reality coming straight from employees themselves, does so in rich detail, and covers a large number of workers in many different organizations.

Applying NLP techniques requires thinking through the rich and complex nature of language – spelling, punctuation, grammar and syntax, word forms (morphology), semantics, symbolic representations (e.g., using all capitals or italics for emphasis), and special characters (e.g., $, ☺) (Jurafsky and Martin, 2021). NLP techniques must be able to determine what constitutes a word and what constitutes a sentence, differentiate meaning based on context (e.g., negations, intensifying adjectives and adverbs), and be flexible about the unit of analysis (a word, a phrase, a sentence, etc.). Because of the complexity of language, NLP techniques require thought and effort to prepare text data for analysis. The appendix explains the steps we took in data preparation.

---

[9] Because these data come from a web portal, some might expect them to over-represent male voices. Instead, they confirm the findings of a recent survey, which showed that women are almost as likely as men to contribute to social-media sites (Statistica, 2018).

Measuring the salience of work-life balance. We began by creating a lexicon for such terms based on prior research (e.g., Williams, Berdahl, and Vandello, 2016; Padavic, Ely, and Reid, 2020; Kelly and Moen, 2020) and the first author's reading of hundreds of employee reviews. We then extended this list inductively to include terms that were closely related to the original set, using an NLP technique called word embeddings.[10] We used pre-trained word embeddings produced by the most popular model, word2vec, trained on some 100 billion words scraped from Google News articles published by American media. These training data target the general public, which is similar to the intended audience for the Glassdoor reviews we analyze. The final set of terms is listed in Table 1.

[Table 1 about here]

We created a binary indicator variable set equal to one when an employee mentioned *any* term in Table 1 and zero otherwise. To probe the different meanings of work-life balance, we divided the terms in Table 1 into three groups: family responsibilities, flexibility, and general. The *family responsibilities* measure was an indicator set equal to one when an employee mentioned the terms under the "family care" column in Table 1 and zero otherwise. The *flexibility* measure was an indicator set equal to one when an employee mentioned the terms under the "time and location" column in Table 1 and zero otherwise. The *general* measure was an indicator set equal to one when an employee mentioned the terms under the "general" column in Table 1 and zero otherwise. As a robustness check, we created a second measure, *general-restricted*, by excluding all terms in the "general" column that contained "family" or "families."

[10] Word-embedding models map words onto high-dimensional vector spaces (typically 100-300 dimensions) and represent semantic relations between words as geometric relations in those spaces (Mikolov, Sutskever, Chen, Corrado, and Dean, 2013). Words that occupy similar contexts (i.e., are collocated with the same other words) are positioned near each other in vector space and are oriented toward shared meanings. Words that occupy very different contexts (i.e., are collocated with different other terms) are positioned far apart and are oriented toward very different meanings. Thus, word-embedding models can be used to uncover words that are semantically close to focal words, even if they are not morphemes, synonyms, or antonyms of antonyms.

Measuring opinions about work-life balance. We used a common NLP technique, *sentiment analysis*, to capture employees' opinions about work-life balance.[11] As its name suggests, sentiment analysis captures the feelings, attitudes, and emotions of a text (Pang and Lee, 2008; Liu, 2020; Jurafsky and Martin, 2021). Much like oral language, written language has a variety of paralinguistic features that express or modify emotion, such as punctuation (e.g., an exclamation mark captures more intensity than a period), capitalization (e.g., "The food was GROSS." is more negative than "The food was gross."), intensifying adjectives and adverbs (e.g. very, extremely, a little, sort of), superlatives (e.g., best, worst), negation (e.g., "The food wasn't really that good"), sarcasm (e.g., "What a great meal! It made me sick to my stomach."), and emphasis through lengthened spelling (e.g., "The service was sloooow."). Sentiment-analysis algorithms are optimized to take these conditions into account to determine, for example, whether texts under study reveal specific emotions (e.g., admiration, envy, surprise, or joy) or the overall polarity (positive to negative) of the emotion expressed in texts.

Because we were concerned with capturing sentiment about a specific topic, we conducted *aspect sentiment analysis* (Liu, 2020). We know that the meanings, including the sentiment polarity, of words and phrases depends on their context; i.e., the surrounding words (Harris, 1954; Firth, 1957). Therefore, we used a well validated algorithm – VADER sentiment analysis algorithm in the NLTK (Natural Language Tool Kit) platform – to measure the sentiment in the words surrounding terms associated with work-life balance. We used a window of three words before and three words after each term in Table 1.[12] VADER works well on short texts like the review snippets we analyzed. We analyzed the pros and cons sections of employee reviews separately because we expected them to focus on very different things. Indeed, pros section sentiment was generally more positive than cons section sentiment, although there were many instances of positive sentiment in the cons section of reviews. For each review, we

[11] The appendix explains sentiment analysis and details the specific techniques we used.

[12] We also tried a window of five words before and five words after. The results were similar, so we focus on the measures based on three-word windows.

measured the sentiment polarity of each word surrounding every instance of each term in Table 1 and took the mean across words in the focal window to get its sentiment polarity. We then summed sentiment polarity scores across the windows surrounding all work-life balance terms in the focal review to yield *total WLB sentiment* of that review.

When the words in windows surrounding work-life balance terms were not in the VADER sentiment lexicon, the algorithm returned values of zero sentiment. This happened for 11.5% of reviews that had data on all covariates and that mentioned work-life balance (7,601/65,943 [see comment]). For the multivariate analysis, we dropped those reviews because they were uninformative. In a robustness check, we re-estimated all multivariate models including those reviews. Those results, which assume that zero sentiment scores indicate neutrality about work-life balance, are very similar, so we show results on the dataset that omits these observations.

We validated *total WLB sentiment* by correlating it with employee ratings of work-life balance. As explained above, this rating is optional, but 90% of Glassdoor reviews included it. The correlation between this rating and total WLB sentiment was 0.40 – moderate but in the expected direction. To further validate this measure, we read hundreds of employee reviews. Table 2 shows examples of reviews that scored high, neutral, and low on sentiment about work-life balance. The top panel shows snippets from reviews scoring among the highest on *total work-life balance sentiment* (strongly positive); the middle, reviews scoring neutral (zero); the bottom, reviews scoring among the lowest (strongly negative). The snippets shown include sentences that contain terms denoting work-life balance; those terms are in bold. The top panel clearly includes the most positive opinions about the ability to achieve work-life balance; the bottom, the most negative opinions. For example, one of the reviews with the highest score explained that "quality of life has increased exponentially" at their firm, and another reviewer indicated "For a parent … you are welcomed into a role that allows you to keep a healthy work/life balance with your family". In contrast, reviews with the lowest scores explicitly expressed "poor work-life balance" even in a family emergency, and that "you're expected to work from home" with no respect for one's free time.

[Table 2 about here]

Measuring explanatory variables. Before we describe these measures, we should comment on missing data. Unfortunately, many employee reviews were missing data on gender (42%) or age (71%). Reviews with missing data may be written by respondents who paid little attention to filling out Glassdoor's online form, just going through the motions in order to get full access to Glassdoor's site. If so, reviews that contain full demographic information may be of higher quality. But t tests revealed that reviews with information on gender were significantly more likely to mention work-life balance and to discuss it in more positive terms. Moreover, reviews where gender was *not* reported tended to be for smaller firms and privately owned (rather than publicly traded) firms. (Note that privately owned firms tend to be smaller than publicly traded firms.) Employees in smaller firms may be more reluctant to reveal their demographic attributes because it is easier to identify them, even though reviews are posted anonymously. To investigate the possibility of selection bias in reviews of small firms, we conducted a robustness check where we dropped reviews from firms in the bottom quartile for size.

We measured our key variable, *employee gender*, with a binary indicator variable equal to one for female and zero for male.[13] We measured *employee age* by subtracting birth year from the year the focal review was written; we logged age to normalize its distribution.[14] We measured *ownership type* with a trichotomous variable: publicly traded, privately owned, and other. Our focus is on differences between public and private firms; the category "other" is merely a control. It includes non-profits, government agencies, educational institutions, and

[13] None of the people who wrote the reviews in our sample listed any other descriptor for gender than male or female.

[14] We eliminated observations with outliers on age: 2,678 observations where age was over 65 and 108 where age was under 18. Most of the former indicated birth year as 1900 and so were typographical errors. Together, outliers on age constituted 1% of observations in the dataset where age was recorded.

subsidiaries of for-profit firms. We could not determine whether subsidiaries were owned by privately owned or publicly traded firms, so this is a heterogeneous category. But it is small: it constitutes less than 7% of observations. For *firm size*, we counted the number of employees and logged it to normalize its distribution. (We dropped 406 reviews that indicated employers had zero employees because these were most likely typographical errors.)

We included controls for region and industry to capture similarities among firms and help deal with non-independence of observations across reviews. Organizations that operate in the same industry or region may be similar to each other due to coercive forces (state legal regimes), normative forces (industry standards or regional cultural values), or mimetic forces (local role models) (DiMaggio and Powell, 1983), so their employees' descriptions may be similar. In the tech sector, where worker mobility rates have always been high, mimesis is even more likely than in other sectors. For *region*, we relied on the U.S. Bureau of Economic Analysis definition, which takes into consideration the homogeneity of states in terms of economic, demographic, social, and cultural characteristics (U.S. Bureau of Economic Analysis, 2014). There are eight regions: New England, Mideast, Southeast, Great Lakes, Plains, Southwest, Rocky Mountain, and Far West.[15] When state was not recorded, we created an "unknown" category. We measured *industry* using the nine categories listed above in footnote 6.

Finally, we controlled for attributes of workplaces and jobs that might compensate for lack of work-life balance or that might be associated with mentions of that topic. Based on preliminary analysis, we focused on *compensation* (e.g., "pay," "salary," "compensation") and several different kinds of perks: *food and drink* (e.g., "food," "lunch," "snacks"), *retirement and*

---

[15] New England includes Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. The Mideast contains Delaware, District of Columbia, Maryland, New Jersey, New York, and Pennsylvania. The Southeast includes Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The Great Lakes comprises Illinois, Indiana, Michigan, Ohio, and Wisconsin. The Plains includes Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota. The Southwest contains Arizona, New Mexico, Oklahoma, and Texas. The Rocky Mountain region includes Colorado, Idaho, Montana, Utah, and Wyoming. Last, the Far West comprises Alaska, California, Hawaii, Nevada, Oregon, and Washington.

time off (e.g., “vacation,” “pto” [paid time off], “holidays”), *health and wellness* (e.g., “gym,” “healthcare,” “dental”), and *miscellaneous* (e.g., “swag,” “pets,” “laundry,” “perks”). For each category, we included words that appeared at least 100 times in the full dataset. We created two binary indicator variables. The first was set equal to one if any word about compensation was mentioned in the focal review, and zero otherwise; the second was set equal to one if any word about other perks was mentioned in the focal review, and zero otherwise.[16]

Methods of analyses

First, we analyzed *whether or not employees discussed work-life balance* using the binary indicator variable equal to one when an employee review mentioned any terms in the work-life-balance lexicon and zero otherwise. We estimated logistic regressions on the 260,862 observations that were not missing data on any of the variables of interest.

Second, to assess employee’s opinions about work-life balance, we analyzed the distribution of *total work-life balance sentiment* across the entire sample. For this outcome, our analysis was limited to the 65,943 reviews that mentioned work-life balance. Because total WLB sentiment is a continuous variable, we used OLS regression. (The distribution of this variable, as we explain below, is close to normal, with a slight positive skew.)

To probe what employees talk about when they talk about work-life balance, we inductively coded a random sample of 240 reviews (about 0.4% of the total), stratified by *total work-life balance sentiment*: one-third positive, one-third neutral, and one-third negative. We used MAXQDA, a flexible software application, to code this sample of reviews. MAXQDA is a mixed-methods data analysis software that is used to qualitatively code text and multimedia data. We began by coding reviews using themes identified from the literature and our research questions (e.g., around-the-clock work expectations, flexibility in location and scheduling, parental leave), then added additional codes that emerged as we read through the reviews.

[16] Lexicons for compensation and perks are available from the first author.

Results

As mentioned above, 25.3% of reviews mentioned work-life balance overall, 14.0% mentioned family responsibilities specifically, 9.6% mentioned flexibility, and 9.8% mentioned quality of life or work-life balance in general (9.7% when we used the restricted set of general terms). To put this in perspective, 32.7% of reviews mentioned compensation, while 10.3% mentioned perks. Thus work-life balance is a moderately common topic in employee descriptions of their firms.

Who discusses work-life balance?

Table 3 shows univariate statistics and correlations for the theoretical variables in this analysis; to save space, we do not show correlations with the dummy variables for industry, region, or year. There are no strong correlations between the dependent variable and any explanatory variables. However, there is a positive correlation between firm size and the public firm dummy (r= 0.704). This makes sense because privately owned companies tend to be smaller than publicly traded ones.

[Table 3 about here]

Table 4 shows logistic regression results in two parts, 4a and 4b. In Table 4a, model 1 is a baseline with fixed effects for industry, region, and firm type; a continuous variable for firm size; and dummy variables for discussion of compensation and benefits. The industry (cable, internet, and telephone provider), region (Far West), and firm type (privately owned) categories that accounted for the largest number of observations are reference categories. As expected, there were differences between both industries and regions in the likelihood of employees mentioning work-life balance in reviews. Employees were more likely to mention work-life balance in their reviews the more they discussed compensation, and less likely the more they discussed other perks. Employees in publicly traded companies were significantly less likely than those in privately owned firms to mention work-life balance in their reviews, while employees in larger firms were significantly more likely than those in smaller firms.

[Table 4 about here]

Model 2 adds gender and shows that women were one-third (exp[0.295] = 1.34) more likely to mention work-life balance than men. This result supports hypothesis 1. Model 3 adds employee age and reveals that younger employees were more likely to discuss work-life balance than older ones. Model 4 incorporates both variables; the results do not change. Model 5 adds a squared term for employee age and shows a positive coefficient on the linear term for age and a negative coefficient on the squared term. The peak of this inverted-U pattern is around 34 years, near the mean for age. This means that employees in the middle of the age range were the most likely to mention work-life balance. Employees ten years younger were 21% less likely to mention work-life balance, while those ten years older were 10% less likely. These results support hypothesis 2.

Table 4b shows the second part of the results, focusing on interactions with gender. Model 6 includes the gender/age interaction (both linear and squared terms). It shows a negative (constrained) main effect for gender, an inverted-U-shaped (constrained) main effect for age, and an inverted-U-shaped effect for the interaction between gender and age. For women, the peak of the inverted-U was at 35 years, which is just above the mean; for men, it was at 34 years, which is at the mean. Across the observed range for age (18-65 years), women were always more likely to mention work-life balance. As Figure 2 shows, the gender gap was widest at age 42, which is at the 75th percentile of the age distribution. Between ages 34 and 52, the gender gap was always 40% or more. At age 24, it was 25%; at age 60, it was 35%. Overall, this pattern of results supports hypothesis 3.

[Figure 2 about here]

Model 7 shows the gender/firm type interactions. The coefficient on gender is positive; the coefficient on publicly traded firms is negative. The coefficient on the interaction is negative, indicating that the gender gap in mentioning work-life balance was less in publicly traded companies (33%) than in privately owned firms (39%). This result supports hypothesis 4. Model 8 shows the gender-firm size interaction, which is negative. Combined with the positive (constrained) main effects on gender and firm size, this indicates that the gender gap in mentioning work-life balance was slightly less in larger firms. The gender gap was 10% smaller in firms at the mean plus one standard deviation than those at the mean. This pattern supports hypothesis 5, although the difference is small: a gender gap of 36% vs. 38% in the odds of mentioning work-life balance. Finally, model 9 shows all interactions in the same model. Previous results hold except that the coefficient on the gender/firm size interaction is positive but very tiny and nonsignificant, indicating no difference in the gender gap in discussing work-life balance across firms of different size.

Table 5 shows the logistic regression results for different aspects of work-life balance: family responsibilities (model 1), time and location (model 2), and general (model 3). Women were significantly more likely than men to mention family responsibilities when discussing work-life balance in reviews, supporting hypotheses 6a and 6b. Although we did not offer a hypothesis about whether women were also more likely to discuss work-life balance in general, model 3 shows that this is true. This result holds in model 4, where we drop terms for general work-life balance that contain "family" or "families." The gender gap was similar across all models, ranging from 35% in model 1 to 41% in model 3. Thus tech employees were about equally likely to discuss all components of work-life balance.

[Table 5 about here]

Robustness checks. We conducted four supplementary analyses to assess how sensitive our multivariate results are to model specification and sampling. First, our data cover almost seven years, from the beginning of 2014 to October 2020. To control for temporal shifts in attitudes toward and interest in work-life balance, we re-estimated all models adding fixed effects for year, with 2020 as the reference category. The models with year fixed effects, which are not shown here to save space, are almost identical to those without year fixed effects.

Second, the multivariate analyses we presented include only those reviews that have data on *both* employee gender and age. But data on gender are available for many reviews that lack data on age. To assess the severity of sample-selection bias due to missing data, we re-estimated most models without age for all reviews that contain data on gender (N=545,849). The results of the analysis of whether employees mention work-life balance, which are shown in Table 6, show that female employees were more likely than their male counterparts to discuss work-life balance. Furthermore, female employees in publicly traded firms and larger firms were less likely to discuss work-life balance than those in privately owned firms and smaller firms. These results bolster our confidence in the findings shown in Table 4.

[Table 6 about here]

Third, we re-estimated the main models for mentioning work-life balance after dropping data on firms in the bottom quartile of the size range. We did this to investigate the possibility that employees in smaller firms are more reluctant to reveal their demographic attributes because it is easier to identify them. In this subsample of data, 75% of reviews reported gender and 21% reported age, slightly more than the 58% that reported gender and 19% that reported age in the full dataset. Table 7 shows these multivariate results. The results are very similar to those found from analysis of the full dataset, further bolstering our confidence in the findings shown in Table 4.

[Table 7 about here]

Fourth and finally, to assess possible bias due to employees being pressured to write positive reviews of their firms, we re-estimated the main models for mentioning work-life balance in two separate samples: reviews with lower overall ratings (1 or 2 stars) and reviews with higher overall ratings (4 or 5 stars). Table 8 shows these analyses: models 1 to 3 are based on lower-rated reviews; models 4 to 6 on higher-rated reviews. While the results are very similar to those found in our multivariate analysis of the full dataset, there are slight differences in results in low-rating versus high-rating reviews. With regard to the main effects of gender and age, findings from both lower- and higher-rated reviews parallel the results from the full dataset. With regard to interactions with gender, some in the low-rating subsample are not statistically significant. This may be due to its small sample size (n=68,395 for low ratings;

n=145,766 for high ratings). Overall, these results indicate that higher-rated reviews are *not* different from the patterns found in all reviews.

[Table 8 about here]

What do employees say about work-life balance?

We begin by asking simply whether tech workers seemed happy with work-life balance. Figure 3 shows the distribution of sentiment scores for all reviews that mentioned work-life balance. Figure 3a shows the distribution for the pros section of the reviews, Figure 3b for the cons section, and Figure 3c for the two sections combined (i.e., their scores added together.)

[Figure 3 about here]

Not surprisingly, sentiment about work-life balance expressed in the pros section of reviews was generally positive (mean = 0.546), with only 1.9% of reviews expressing negative sentiment. In contrast, sentiment about work-life balance expressed in the cons section of reviews was more evenly distributed: the mean was -0.014, with 9.8% positive and 11.2% negative. Tech employees were more likely to discuss work-life balance in the pros section than in the cons section of reviews. When they did discuss work-life balance in the cons section, most tech employees expressed neutral sentiment, reflecting the fact that tech employees often discussed work-life balance in simple factual terms, not sentiment-laden terms. Finally, the combined work-life balance sentiment measure was generally positive, with a mean of 0.532, 75.8% positive, 12.3% negative, and 11.8% neutral.

Despite the generally positive sentiment about work-life balance, there was considerable variation. The standard deviation for total WLB sentiment was almost as large as the mean for the pros section and the combination of pros and cons: 0.505 and 0.803, respectively. For the cons section, where there were the most negative scores and many zero (neutral) scores, the standard deviation was far larger than the mean: 0.569.

Testing hypotheses about work-life balance sentiment. Table 9 shows correlations for the variables of theoretical interest in this analysis; again, to save space, we do not show correlations for the dummy variables for industry, region, or year. As before, there are no strong correlations between the dependent variable and any explanatory variables. Again, there is a positive correlation between firm size and the public firm dummy (r= 0.697).

[Table 9 about here]

Table 10 shows the results of the multivariate (OLS regression) analysis of sentiment about work-life balance. This table is split into two parts; we discuss each in turn. In Table 10a, model 1 contains control variables while model 2 adds gender. Model 2 shows a null relationship between gender and opinions about work-life balance: the coefficient is near zero and not statistically significant. This null relationship, which fails to support either hypothesis 1a or hypothesis 1b, holds across all subsequent models. Model 3 substitutes age for gender, model 4 adds gender back in, and model 5 adds the squared term for age. Model 5 shows an inverted-U-shaped relationship between age and sentiment about work-life balance. The peak of this inverted U is 33 years, very close to the middle of the age range. These results support hypothesis 2b rather than hypothesis 2b.

[Table 10 about here]

Table 10b focuses on gender interactions with age, firm type and firm size, respectively. In model 6, both gender/age interactions are near-zero and nonsignificant, failing to support hypothesis 3a. In model 7, both gender/firm type interactions are near zero and nonsignificant, which fails to support hypothesis 4a. In model 8, the gender/firm size interaction is negative, but near zero. This indicates that the gender gap in satisfaction with work-life balance is slightly greater in larger firms, which fails to support hypothesis 5a.

What employees say about work-life balance: beyond sentiment. As we mentioned above, tech employees generally discussed work-life balance in factual terms; for example, by just mentioning a term denoting work-life balance without elaborating. To investigate themes that appeared when employees did elaborate, we coded a random sample of 240 reviews stratified by the *total WLB sentiment* score (one-third positive, one-third neutral, one-third negative). This analysis revealed that employees often emphasized how their firms prioritized their health, families, and overall well-being. For example, one female employee noted that her company offers flexible schedules, "always understanding that your health and family come first." A 25-year-old female employee at a different firm underscored that her company "not only discourages overworking but they ENCOURAGE that family and your happiness is first and foremost." Such themes were not limited to female employees. For instance, a 43-year-old male employee noted that "employee health and balance is the most important thing to everyone" at his firm. While our Glassdoor data include reviews through the early months of the Covid-19 pandemic, the reviews highlighted here were written before the pandemic, indicating a priority on holistic health and wellness as a key component of work-life balance even before the pandemic.

In addition, many employees underscored how work-life balance practices vary based according to work team, manager, and geographic location. A 26-year-old female employee noted, "There is not much structure or accountability, which is good if you want to take unlimited vacation days or 'work' from home a lot, but not so good if you're the one picking up the slack. In general, there's a lack of accountability in most departments." This lack of accountability and uniform structure from management results in de facto work-life balance, such as being able to take paid time off or work remotely, but only with managerial approval, rather than through formalized rules or policies that apply uniformly and automatically to all employees. As 52-year-old male employee highlighted this variability, indicating that work-life balance is generally good, but "certain groups can have poor work-life balance." A 38-year-old female employee shared his sentiment, acknowledging, "Work/life balance honestly depends on which group you're in. Some groups are more lenient than others." There is some evidence that differences also occur across work sites for firms that have operations nationally and globally. As a 35-year-old male employee described, "Management is completely different between sites; I originally started working in MX [Mexico] then moved to CA and finally ended in TX. Management is the worst area…, they don't care about people's lives and wellbeing, they think they own you which is wrong…. I would say the best was in CA, there was more work-life balance."

Our findings echoed those of a small-N (interview-based) study (Williams, 2018), which found that firms with less formalized policies left accommodations like flexible work schedules to the discretion of supervisors. This often disadvantaged female employees, who were more likely to have to negotiate these accommodations to handle childcare or eldercare responsibilities. Our large-N analysis, which covers all firms in the tech sector across the country over seven years, similarly revealed that managers often decide which employees, in which work groups and sites, can structure their jobs in ways that allow them to balance work and life outside of work. The lack of uniform policies concerning work-life balance increases the likelihood of gender-based discrimination (Reskin and McBrier, 2000) because women tend to do more domestic labor. This may be the reason why work-life balance is raised as a topic of discussion more by female employees than their male counterparts.

Discussion and Conclusion

Analyses of Glassdoor reviews provide a window into the gendered experiences of work-life balance among tech employees. As we expected, we found that the characteristics of tech employees and their firms mattered a great deal. Female tech employees were more likely than their male counterparts to mention work-life balance in their descriptions of their firms. The salience of work-life balance is greater for women than men because women tend to be responsible for most domestic duties. We also found that employees in the middle of the age range were more likely than their younger or older counterparts to discuss work-life balance. Employees in the middle of the age range are the most likely to have young children, who need hands-on care and direct supervision, which explains why this topic is most salient for them. Moreover, we found that women in the middle of the age range were the most likely to bring up work-life balance when they described their firms. While younger and older women have to juggle the demands of work and family life more than their same-age male counterparts, it is women in the middle of the age range who are likely to bear the heaviest child-care burden. Finally, we found that the gender gap in mentioning work-life balance was smaller in larger and publicly traded firms, which we argued was because of those firms' greater resources, visibility, and formalization.

Tech workers generally expressed satisfaction with work-life balance, although there was considerable variation in how satisfied they were, and a substantial minority felt dissatisfied. We did not, however, find much evidence that tech employees' opinions (positive, neutral, or negative) could be predicted by personal characteristics, including gender, or workplace context. The only statistically significant finding about sentiment was that employees in the middle of the age range – those who are most likely to have young children – were more likely to have positive opinions about work-life balance. Digging into a stratified random sample of reviews using qualitative coding methods revealed that work-life balance was often not contingent on standardized policies applied uniformly to all employees; rather, it depended on managerial approval and discretion. These findings fit with small-N, interview-based research that revealed flexible work arrangements to be perks that workers must negotiate with their managers (Kelly and Kalev, 2014; Williams, 2018). Yet, as with the quantitative sentiment analysis, this qualitative thematic analysis did not reveal any gender differences.

A caveat about causality: These data generally contain one observation (one review) per employee, so we cannot pinpoint causality. To use statistical techniques to identify causation, we would need to aggregate data to the firm in the time period. That requires matching firm names (the raw data do not contain firm identification numbers), which can be done for a majority of observations because the Glassdoor interface prompts employees who start entering their firm name with a list of possible responses. But a sizable fraction of reviews has idiosyncratic or very rare names, so we would have to employ fuzzy-matching techniques to prepare to aggregate the data. That task is on our agenda for the future. That said, our arguments focus on regressors that are unlikely to be affected by whether employees discuss work-life balance when they describe their firms, and if so, what their opinion about it is.

One more caveat bears mention: Our analysis did not control for job level or job type. Higher-level positions tend to have more flexibility than lower-level ones (Gerstel and Clawson, 2014; Boushey, 2016). And in the tech sector, technology-focused jobs (such as software engineer, data scientist, and systems administrator) are accorded higher status than non-technical jobs (such as in sales representative and financial analyst). Therefore, future research might usefully distinguish between managerial and non-managerial positions (or draw more fine-grained distinctions where possible) and between technical and non-technical positions. It would also be useful to expand the scope of the analysis beyond the tech sector and to compare sectors where women are more likely to thrive (e.g., retail sales) with those where they are less likely to thrive (e.g., financial services).

Looking at the big picture, our analysis reveals considerable variation in the salience of this topic across workers: salience was greater for female employees and employees who are (based on age) most likely to have young children. This suggests that employers might benefit from designing cafeteria-style work-life balance policies, with workers able to choose the options that fit their particular situations. For example, flexibility in time and place means more to female than male employees, while support for activities outside of the family sphere (e.g., volunteering in their local communities) may resonate more with younger employees. Future research (e.g., survey and field experiments) could test this argument.

We also found substantial differences between firms – larger vs. smaller, publicly traded vs. privately owned – which we argued were due to differences in resources, formalization, and visibility. These findings indicate that while smaller and privately owned firms may not have as many resources to allocate to work-life balance policies, it is important to workers (25.3% mentioned it when asked to describe their firms). Therefore, even in firms that are resource-constrained, designing jobs and workplace practices around work-life balance is crucial to attracting, motivating, and retaining talent.

References

Acker, Joan. 1990. Hierarchies, jobs, bodies: A theory of gendered organizations. *Gender and Society*, 4: 139-158.

Anderson Stella E., Betty S. Coffey, and Robin T. Byerly. 2002. Formal organizational initiatives and informal workplace practices: Links to work-family conflict and job-related outcomes. *Journal of Management*, 28: 787-810.

Barley, Stephen R. 1983. Semiotics and the study of occupational and organizational cultures. *Administrative Science Quarterly*, 28: 393-413.

Baron, James N., Alison Davis-Blake, and William T. Bielby. 1986. The structure of opportunity: How promotion ladders vary within and among organizations. *Administrative Science Quarterly*, 31: 248-273.

Berdahl, Jennifer L., Marianne Cooper, Peter Glick, Robert W. Livingstone, and Joan C. Williams. 2018. Work as a masculinity contest. *Journal of Social Issues*, 74 (3): 422-448.

Bielby, William T., and James N. Baron. 1986. Men and women at work: Sex segregation and statistical discrimination. *American Journal of Sociology*, 91: 759-799.

Blau, Peter M., and W. Richard Scott. 1962. *Formal Organizations: A Comparative Approach*. San Francisco, CA: Chandler Publishing Company.

Blau, Peter M., and Richard Schoenherr. 1971. *The Structure of Organizations*. New York: Basic Books.

Boushey, Heather. 2016. *Finding Time*: *The Economics of Work-Life Conflict*. Cambridge, MA: Harvard University Press.

Brammer, Stephen, and Andrew Millington. 2006. Firm size, organizational visibility, and corporate philanthropy. *Business Ethics*, 15 (1): 6-18.

Britton, Dana M. 2000. The epistemology of the gendered organization. *Gender and Society*, 14 (3): 418-434.

Carley, Kathleen, and Michael Palmquist. 1992. Extracting, representing, and analyzing mental models. *Social Forces*, 70 (3): 601-636.

Castilla, Emilio J. 2015. Accounting for the gap: A firm study manipulating organizational accountability and transparency in pay decisions. *Organization Science*, 26 (2): 311-333.

Cha, Youngjoo, and Kim A. Weeden. 2014. Overwork and the slow convergence in the gender gap in wages. *American Sociological Review*, 79 (3), 457-484.

Chang, Emily. 2018. *Brotopia: Breaking up the Boys' Club of Silicon Valley*. New York: Portfolio/Penguin.

Choper, Joshua, Daniel Schneider, and Kristen Harknett. 2022. Uncertain time: Precarious schedules and job turnover in the US service sector. *Industrial Relations*, in press.

Correll, Shelley J., Erin L. Kelly, Lindsey Trimble O'Connor, and Joan C. Williams. 2014. Redesigning, redefining work. *Work and Occupations*, 41 (1): 3-17.

Correll, Shelley J., Katherine R. Weisshaar, Alison T. Wynn, JoAnne Delfino Wehner. 2020. Inside the black box of organizational life: The gendered language of performance assessment. *American Sociological Review*, 85 (6): 1022-1050.

Cuddy, Amy J.C., Susan T. Fiske, and Peter Glick. 2004. When professionals become mothers, warmth doesn't cut the ice. *Journal of Social Issues*, 60 (4): 701-718.

Cyert, Richard M., and James G. March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.

Dalton, Dan R., and Debra J. Mesch. 1990. The impact of flexible scheduling on employee attendance and turnover. *Administrative Science Quarterly*, 35: 370-387.

Daminger, Alison. 2019. The cognitive division of household labor. *American Sociological Review*, 84 (4): 609-633.

Daniel, Francis, Franz T. Lohrke, Charles J. Fornaciari, and R. Andrew Turner. 2004. Slack resources and firm performance: A meta analysis. *Journal of Business Research*, 57 (6): 565-574.

Davies, Andrea Rees, and Brenda D. Frink. 2014. The origins of the ideal worker: The separation of work from home in the United States from the market revolution to 1950. *Work and Occupations*, 41 (1): 18-39.

DeMers, Jayson. 2014. How negative online company reviews can impact your business and recruiting. *Forbes*, September 9th. (https://www.forbes.com/sites/jaysondemers/2014/09/09/how-negative-online-company-reviews-can-impact-your-business-and-recruiting/?sh=2b604f1b1d9b; viewed 2020-08-02)

DiMaggio, Paul. 1997. Culture and cognition. *Annual Review of Sociology*, 23: 263-287.

DiMaggio, Paul J., and Walter W. Powell. 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48: 147-160.

Dobbin, Frank R., Daniel Schrage, and Alexandra Kalev. 2015. Rage against the iron cage: The varied effects of bureaucratic personnel reforms on diversity. *American Sociological Review*, 80 (5): 1014-1044.

Edelman, Lauren B. 1990. Legal environments and organizational governance: The expansion of due process in the American workplace. *American Journal of Sociology*, 95: 1401-1440.

Edelman, Lauren B., Christopher Uggen, and Howard S. Erlanger. 1999. The endogeneity of legal regulation: Grievance procedures as rational myth. *American Journal of Sociology*, 105: 406-454.

Eliasoph, Nina, and Paul Lichterman. 2003. Culture in interaction. *American Journal of Sociology*, 108 (4): 735-94.

Elizur, Dov, and Samuel Shye. 1990. Quality of work life and its relation to quality of life. *Applied Psychology: International Review*, 39 (3): 275-291.

Firth, John R. 1957. A synopsis of linguistic theory 1930-1955. In John R. Firth, ed., *Studies in Linguistic Analysis* (Special volume of the Philological Society): 1-32. Oxford: Basil Blackwell.

Fowler, Susan. 2017. Reflecting on one very strange year at Uber. (viewed 12 June, 2019)

Galaskiewicz, Joseph. 1997. An urban grants economy revisited: Corporate charitable contributions in the Twin Cities, 1979-81, 1987-89. *Administrative Science Quarterly*, 42 (3): 445-471.

Gallo, Peter Jack, and Lisa Jones Christiansen. 2011. Firm size matters: An empirical investigation of organizational size and ownership on sustainability-related behaviors. *Business and Society*, 50 (2): 315-349.

Gerstel, Naomi, and Dan Clawson. 2014. Class advantage and the gender divide: Flexibility on the job and at home. *American Journal of Sociology*, 120 (2): 395-431.

Glass, Jennifer. 2004. Blessing or curse? Work-family policies and mother's wage growth over time. *Work and Occupations*, 31 (3): 367-394.

Gorman, Elizabeth H. 2005. Gender stereotypes, same-gender preference, and organizational variation in the hiring of women: Evidence from law firms. *American Sociological Review*, 70: 702-728.

Harris, Zellig S. 1954. Distributional structure. *Word*, 10 (2-3): 146-162.

Hochschild, Arlie Russell. 1989. *The Second Shift: Working Parents and the Revolution at Home*. New York: Avon Books.

Hochschild, Arlie Russell. 1997. *The Time Bind: When Work Becomes Home and Home Becomes Work*. New York: Metropolitan Books.

Hunzaker, M.B. Fallin, and Lauren Valentino. 2019. Mapping cultural schemas: From theory to method. *American Sociological Review*, 84 (5): 950-981.

Jacobs, Jerry A., and Kathleen Gerson. 2004. *The Time Divide: Work, Family, and Gender Inequality*. Cambridge, MA: Harvard University Press.

James, Al. 2014. Work-life 'balance,' recession and the gendered limits to learning and innovation (Or, why it pays employers to care). *Gender, Work, and Organizations*, 21 (3): 273-294.

Jurafsky, Dan, and Martin. 2021. *Speech and Language Processing, 3rd Edition Draft*. (https://web.stanford.edu/~jurafsky/slp3/)

Kanter, Rosabeth Moss. 1977. *Men and Women of the Corporation*. New York: Basic Books.

Keene, Jennifer Reid, and Jill Quadagno. 2004. Predictors of perceived work-family balance: Gender difference or gender similarity? *Sociological Perspectives*, 47 (1): 1-23.

Kelliher, Clare, Julia Richardson, and Galine Boiarintseva. 2019. All of work? All of life? Reconceptualizing work-life balance for the 21st century. *Human Resources Management Journal*, 29: 97-112.

Kelly, Erin L., and Frank Dobbin. 1999. Civil rights law at work: Sex discrimination and the rise of maternity leave policies. *American Journal of Sociology*, 105: 455-492.

Kelly, Erin L., and Alexandra Kalev. 2006. Managing flexible work arrangements in US organizations: Formalized discretion or "a right to ask." *Socio-Economic Review*, 4: 379-416.

Kelly, Erin L., and Phyllis Moen. 2020. *Overload: How Good Jobs Went Bad and What We Can Do About It*. Princeton, NJ: Princeton University Press.

King, Julia. 1998. Long hours, cold dinners, and vacation interrupts. *Computerworld*, 32 (41): 1. (http://www.cnn.com/TECH/computing/9810/16/techslave.idg/; viewed 2020-10-08)

Kolhatkar, Sheela. 2017. The tech industry's gender-discrimination problem. *New Yorker*, Nov. 13. (https://www.newyorker.com/magazine/2017/11/20/the-tech-industrys-gender-discrimination-problem; viewed 2020-01-04)

Liu, Bing. 2020. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge: Cambridge University Press.

Lonergan Partners. 2020. *The Lonergan SV 150: List of the Top Public Companies in Technology in Silicon Valley*. (https://lonerganpartners.com/2020-lonergan-sv150-insights; retrieved 2021-06-10)

Loten, Angus. 2019. America's got talent, just not enough in IT. *Wall Street Journal*, Oct. 15. (https://www.wsj.com/articles/americas-got-talent-just-not-enough-in-it-11571168626?mod=hp_featst_pos1; viewed 2022-07-29)

Lotze, Kristen. 2019. 12 tech companies with impressive employee benefits. *Tech Republic*, March. (https://www.techrepublic.com/article/12-tech-companies-with-impressive-benefits/; viewed 2022-04-05)

Luhr, Sigrid W. 2020. Engineering inequality: Rethinking Inclusion in the San Francisco Bay Area Tech Industry. Unpublished dissertation, U.C. Berkeley.

Luckerson, Victor. 2014. 10 most lavish job perks in Silicon Valley. *Time*, 14 Oct. (https://time.com/3506815/10-best-job-perks/; viewed 2022-02-07)

Lyttleton, Thomas, Emma Zang, and Kelly Musick. 2022. Telecommuting and gender inequalities in parents' paid and unpaid work before and during the COVID-19 pandemic. *Journal of Marriage and Family*, 84: 230-249.

Marinescu, Ioana, N. Klein, Andrew Chamberlain, and Morgan Smart. 2021. Incentives can reduce bias in online reviews. *Journal of Experimental Psychology: Applied,* 27 (2): 393-407.

Mas, Alexandre, and Amanda Pallais. 2017. Valuing alternative work arrangements. *American Economic Review*, 107 (12): 3722-3759.

McKinsey and Company. 2018. *Women in the Workplace 2018*. (https://womenintheworkplace.com/2018; retrieved 2020-10-22)

McNall, Laurel A., Aline D. Masuda, and Jessica M. Nicklin. 2009. Flexible work arrangements, job satisfaction, and turnover intentions: The mediating role of work-to-family enrichment. *Journal of Psychology*, 144 (1): 61-81.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems*, Lake Tahoe, CA (December 2013): 3111-3119.

Moen, Phyllis. 2003. *It's About Time: Couples and Careers*. New York: Cornell University Press.

Mogenroth, Thekla, Michelle K. Ryan, Floor Rink, and Christopher Begeny. 2021. The (in)compatibility of identities: Understanding gender differences in work-life conflict through fit with leaders. *British Journal of Social Psychology*, 60: 448-469.

Nippert-Eng, Christena E. 1996. *Home and Work: Negotiating Boundaries through Everyday Life*. Chicago, IL: University of Chicago Press.

O'Connor, Fred. 2011. IT worker demand remains strong despite slow economy. InfoWorld. (https://www.infoworld.com/article/2621679/it-worker-demand-remains-strong-despite-slow-economy.html; viewed 2022-08-22)

Padavic, Irene, Robin J. Ely, and Erin M. Reid. 2020. Explaining the persistence of gender inequality: The work-family narrative as a social defense against the 24/7 work culture. *Administrative Science Quarterly*, 65 (1): 61-111.

Pang, Bo, and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*: 2 (1-2): 1-135.

Pedulla, David S., and Sarah Thébaud. 2015. Can we finish the revolution? Gender, work-family ideals, and institutional constraint. *American Sociological Review*, 80 (1): 116-139.

Perlow, Leslie A. 2012. *Sleeping with Your Smartphone: How to Break the 24/7 Habit and Change the Way You Work*. Cambridge, MA: Harvard Business Review Press.

Rangarajan, Sinduja. 2018. Here's the clearest picture of Silicon Valley diversity yet: It's bad. But some companies are doing less bad. (https://www.salon.com/2018/06/29/heres-the-clearest-picture-of-silicon-valleys-diversity-yet-it-is-bad_partner/; viewed 13 June, 2020)

Reskin, Barbara F. 1993. Sex segregation in the workplace. *Annual Review of Sociology*, 19: 241-270.

Reskin, Barbara F., and Debra Branch McBrier. 2000. Why not ascription? Organizations' employment of male and female managers. *American Sociological Review*, 65: 210-233.

Rho, Hye Jin, Hayley Brown, and Shawn Fremstad. 2020. *A Basic Demographic Profile of Workers in Frontline Industries*. Washington, DC: Center for Economic and Policy Analysis. (https://www.cepr.net/a-basic-demographic-profile-of-workers-in-frontline-industries/; retrieved 2022-01-05).

Sapir, Edward. 1958. *Culture, Language and Personality*. Berkeley: University of California Press.

Saiidi, Uptin. 2016. Want a back rub? Tech doles out perks for talent. *Yahoo! finance*, 21 Feb. (https://finance.yahoo.com/news/want-back-rub-tech-doles-140000861.html; viewed 2022-02-02)

Schor, Juliet B. 1992. *The Overworked American*. New York: Basic Books.

Stainback, Kevin, Thomas N. Ratliff, and Vincent J. Roscigno. 2011. The context of workplace sex discrimination: Sex composition, workplace culture, and relative power. *Social Forces*, 89 (4): 1165-1188.

Statistica. 2018. Most popular media activities carried out on a regular basis among consumers in the United States as of January 2018, by gender. (https://www.statista.com/statistics/816398/popular-media-activities-conduced-regularly-us-by-gender/; viewed 2021-08-25)

Statistica. 2022. Distribution of tech jobs that are hard to fill from 2019 to 2021, by country. (https://www.statista.com/statistics/1222470/tech-job-supply-demand-balance-country-year/; viewed 2022-08-22)

Stryker, Sheldon. 1980. *Symbolic Interactionism: A Social Structural Version*. Menlo Park, CA: The Benjamin/Cummings Publishing Company.

Thomas, Lauren. 2021. Burnout on the rise: Work-life balance in the U.S. during COVID-19. Glassdoor Economic Research report. (https://www.glassdoor.com/research/burnout-work-life-balance-covid-19/; viewed 2022-08-08)

Thompson, Clive. 2020. *Coders: The Making of a New Tribe and the Remaking of the World*. New York: Penguin.

Tian, Gloria. 2016. *The Best Industries for Benefits*. Glassdoor Economic Research report. (https://www.glassdoor.com/research/best-industries-for-benefits/; viewed 2022-02-03).

U.S. Bureau of Economic Analysis. 2014. Regional Maps, GDP by State 2014. https://www.bea.gov/news/2015/gross-domestic-product-state-advance-2014-and-revised-1997-2013/regional-maps

U.S. Bureau of Labor Statistics. 2019. American Time Use Survey, "Average hours per day parents spent caring for and helping household children as their main activity." (https://www.bls.gov/charts/american-time-use/activity-by-parent.htm; viewed 2022-02-06)

Venkatesh, A., and N. Vitalari. 1992. An emerging distributed work arrangement: An investigation of computer-based supplemental work at home. *Management Science*, 38: 1687-1706.

Whorf, Benjamin L. 1956. Science and linguistics. In John B. Carroll, ed., *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.

Williams, Christine L. 2018. The gendered discourse of work-family balance in the oil and gas industry. *Social Currents*, 5 (2): 120-139.

Williams, Joan C., Mary Blair-Loy, and Jennifer L. Berdahl. 2013. Cultural schemas, social class, and the flexibility stigma. *Journal of Social Issues*, 69: 209-234.

Williams, Joan C., Jennifer L. Berdahl, and Joseph A. Vandello. 2016. Beyond work-life "integration." *Annual Review of Psychology*, 67: 515-539.

Wynn, Alison T., and Shelley J. Correll. 2018. Puncturing the pipeline: Do technology companies alienate women in recruiting sessions? *Social Studies of Science*, 48 (1): 149-164.

Zerubavel, Eviatar. 1991. *The Fine Line: Making Distinctions in Everyday Life*. Chicago: University of Chicago Press.

Figure 1: Example Employee Review from Glassdoor.com

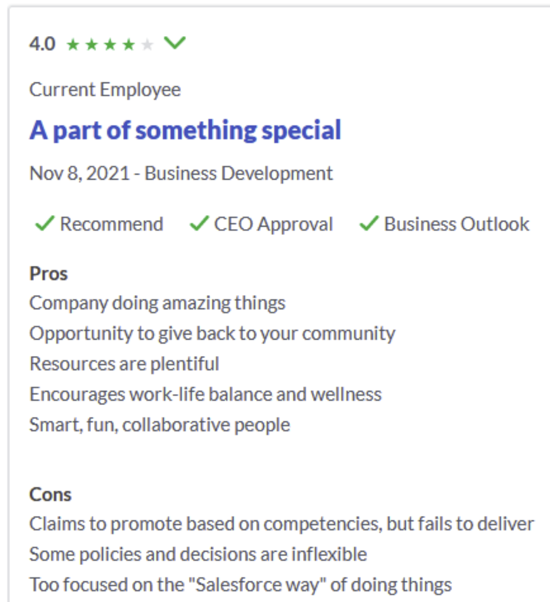

Figure 1a: Example Employee Review from Glassdoor.com: Optional Ratings on Topics

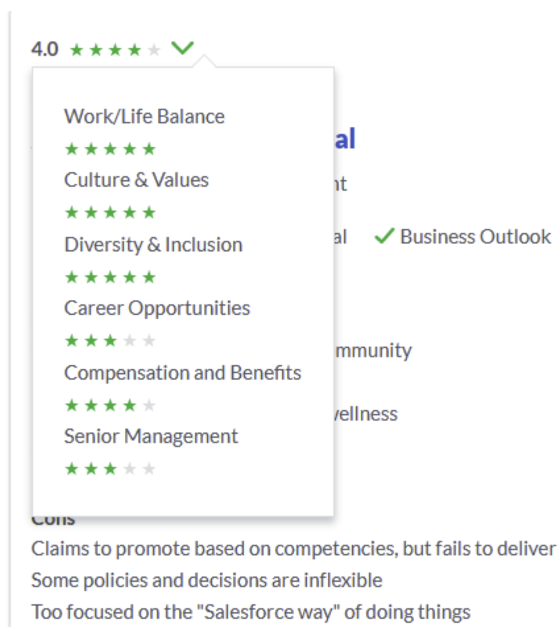

Table 1: Work-Life Balance Lexicon

Time and Location	Family Care	General
flex time	child care	quality [of] life
flextimes	childcare	work life
flex times	dependent care	worklife
flextime	elder care	work families
flexibility	family leave	workfamilies
flexible	family leaves	work family
telecommute	maternity leave	workfamily
telecommuter	maternity leaves	balance
telecommuters	parental leave	balances
telework	parental leaves	balancing
teleworker	paternity leave	
teleworkers	paternity leaves	
work [from] home	baby	
work [at] home	babies	
	care home	
	family	
	families	
	nursery	
	nurseries	
	parent	
	parents	

Source: Compiled by the first author based on Kelly and Moen (2020) and other research on work-life balance. Extended using a pre-trained word-embedding model (word2vec) to uncover semantically similar terms. Terms in square brackets are stop words, which we removed from the corpus during pre-processing.

Table 2:
Example Reviews with High, Neutral, and Low Total Work-Life Balance Sentiment Scores

Review Number	Total WLB Sentiment Score	
Positive Reviews		
64883	4.218	I have a kid and with the home schooling going on they have made sure I understood my priority should be my **family** and kiddo. This has reduced so much stress and allowed me the freedom to be the best person I can be in both **work** [and] **home** life. I am so happy I've made the jump to Paylocity. My **quality** [of] **life** has increased exponentially and I can't imagine going back to where I was before. I haven't had any yet! The PTO doesn't accrue as quickly as some previous jobs I've had but I also haven't had to use it as often due to the fantastic **work/life balance**. Keep doing what you're doing!
30205	3.903	Great f**lexibility** and **work** [from] **home** are the primary reasons why people stay at Lionbridge. For a **parent**, like me, or someone who wants a **flexible** schedule, you are welcomed into a role that allows you to keep a healthy **work/life balance** with your **family**.
Neutral Reviews		
24	0	Nice buildings, plenty of perks here and there. **Work** [from] **home** was really not an option but since the offices were nice, it wasn't a big deal.
238	0	Lots of money to be made. Profit sharing. Yearly minimum 3 percent bonus. Yearly 3 percent increase in salary. The benefits are the cheapest and most comprehensive I've ever heard of. In retail, the stores close at 9PM; so it can be difficult to work the job and manage a **family** with **children** as well.
Negative Reviews		
54713	-2.404	Poor **work-life balance**. No **work** [from] **home**. Even my **family** emergency is not understandable here. I once requested for a time off for my kid fever, and my director instead asked me back for my spouse's support. And I only have 3 of these requests in the whole 2013!
32629	-2.111	Sick days: You are not allowed to be sick. If you're sick, you're expected to **work** [from] **home**. No respect of your free time: you have to be available over email and chat long after work hours, on weekends, holidays, sick days, and when you're off. The **work-life-balance** is bad - lot's of unpaid overtime.

Table 3: Descriptive Statistics for Logistic Regression Variables

	Variable	1	2	3	4	5	6
	Mean	0.2528	0.4564	0.0664	7.5438	0.3455	3.5148
	Standard Deviation	0.4346	0.4981	0.2490	3.2019	0.4755	0.2780
	Minimum	0	0	0	0	0	2.8904
	Maximum	1	1	1	13.122	1	4.1744
	Number of Observations	260,862	260,862	260,862	260,862	260,862	260,862
1	Discuss WLBal						
2	Publicly Held	-0.0040					
3	Other Ownership	-0.0114	-0.2444				
4	Log(Firm Size)	-0.0000	0.7039	-0.0205			
5	Gender (F=1)	0.0589	-0.0139	-0.0032	-0.0136		
6	Log(Age)	-0.0150	0.0138	0.0111	0.0250	-0.0292	

Table 4a: Logistic Regression Analysis of Mentioning Work-Life Balance

	(1)	(2)	(3)	(4)	(5)
Computing HW & SW	0.376***	0.385***	0.376***	0.384***	0.386***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Electrical Mfg	0.119***	0.136***	0.122***	0.139***	0.162***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Bus SW & NW Solns	0.354***	0.358***	0.355***	0.359***	0.360***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
IT Svcs	0.238***	0.249***	0.241***	0.251***	0.251***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Internet	0.231***	0.212***	0.223***	0.205***	0.203***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Telecomm Mfg	0.121	0.139*	0.125*	0.143*	0.169**
	(0.062)	(0.062)	(0.062)	(0.063)	(0.063)
Telecomm Svcs	0.158***	0.152***	0.157***	0.151***	0.149***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Video Games	0.189***	0.214***	0.186***	0.211***	0.181***
	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)
Great Lakes	-0.029	-0.037	-0.032	-0.039	-0.027
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Mideast	-0.058**	-0.062**	-0.060**	-0.064***	-0.055**
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Mountains	-0.026	-0.022	-0.027	-0.023	-0.023
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
New England	0.044	0.041	0.043	0.040	0.062*
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Plains	0.060	0.049	0.059	0.048	0.054
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Southeast	-0.006	-0.024	-0.005	-0.023	-0.013
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Southwest	-0.048*	-0.056**	-0.048*	-0.056**	-0.050**
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Unknown region	-0.138***	-0.161***	-0.142***	-0.165***	-0.157***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)
Discuss Compensn	0.485***	0.488***	0.483***	0.486***	0.474***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Discuss Benefits	-2.357***	-2.584***	-2.395***	-2.618***	-2.584***
	(0.179)	(0.180)	(0.180)	(0.181)	(0.181)
Publicly Held	-0.065***	-0.060***	-0.064***	-0.060***	-0.060***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Other Ownership	-0.094***	-0.091***	-0.092***	-0.090***	-0.088***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Log(Firm Size)	0.009***	0.009***	0.009***	0.010***	0.012***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Gender (F=1)		0.295***		0.294***	0.306***
		(0.009)		(0.009)	(0.009)
Log(Age)			-0.104***	-0.096***	13.899***
			(0.017)	(0.017)	(0.427)
Log(Age)^2					-1.970***
					(0.060)

Note: N = 260,862 *p<0.05; **p<0.01; ***p<0.001

Table 4b: Logistic Regression Analysis of Mentioning Work-Life Balance

	(6)	(7)	(8)	(9)
Computing HW & SW	0.387***	0.386***	0.386***	0.387***
	(0.027)	(0.027)	(0.027)	(0.027)
Electrical Mfg	0.161***	0.161***	0.162***	0.160***
	(0.033)	(0.033)	(0.033)	(0.033)
Bus SW & NW Solns	0.361***	0.360***	0.360***	0.361***
	(0.028)	(0.028)	(0.028)	(0.028)
IT Svcs	0.252***	0.252***	0.251***	0.252***
	(0.028)	(0.028)	(0.028)	(0.028)
Internet	0.204***	0.203***	0.203***	0.204***
	(0.028)	(0.028)	(0.028)	(0.028)
Telecomm Mfg	0.169**	0.170**	0.169**	0.170**
	(0.063)	(0.063)	(0.063)	(0.063)
Telecomm Svcs	0.149***	0.149***	0.149***	0.149***
	(0.030)	(0.030)	(0.030)	(0.030)
Video Games	0.182***	0.183***	0.182***	0.184***
	(0.050)	(0.050)	(0.050)	(0.050)
Great Lakes	-0.028	-0.026	-0.027	-0.027
	(0.021)	(0.021)	(0.021)	(0.021)
Mideast	-0.055**	-0.054**	-0.055**	-0.054**
	(0.019)	(0.019)	(0.019)	(0.019)
Mountains	-0.022	-0.022	-0.022	-0.022
	(0.027)	(0.027)	(0.027)	(0.027)
New England	0.062*	0.062*	0.063*	0.062*
	(0.026)	(0.026)	(0.026)	(0.026)
Plains	0.053	0.054	0.054	0.053
	(0.031)	(0.031)	(0.031)	(0.031)
Southeast	-0.013	-0.012	-0.012	-0.013
	(0.018)	(0.018)	(0.018)	(0.018)
Southwest	-0.050**	-0.050**	-0.050**	-0.050**
	(0.019)	(0.019)	(0.019)	(0.019)
Unknown region	-0.157***	-0.157***	-0.157***	-0.157***
	(0.013)	(0.013)	(0.013)	(0.013)
Discuss Compensn	0.474***	0.473***	0.473***	0.473***
	(0.010)	(0.010)	(0.010)	(0.010)
Discuss Benefits	-2.584***	-2.580***	-2.581***	-2.580***
	(0.181)	(0.181)	(0.181)	(0.181)
Publicly Held	-0.060***	-0.031*	-0.060***	-0.029
	(0.014)	(0.016)	(0.014)	(0.017)
Other Ownership	-0.088***	-0.084***	-0.088***	-0.082**
	(0.020)	(0.025)	(0.020)	(0.025)
Log(Firm Size)	0.012***	0.012***	0.015***	0.011***
	(0.002)	(0.002)	(0.002)	(0.003)
Gender (F=1)	-4.653**	0.342***	0.368***	-4.473**
	(1.550)	(0.014)	(0.024)	(1.552)
Log(Age)	12.850***	13.903***	13.903***	12.887***
	(0.543)	(0.427)	(0.427)	(0.543)
Log(Age)^2	-1.830***	-1.970***	-1.970***	-1.835***
	(0.076)	(0.060)	(0.060)	(0.076)

Table 4b: Logistic Regression Analysis of Mentioning Work-Life Balance (cont'd)

	(6)	(7)	(8)	(9)
Gender x Log(Age)	2.671**			2.585**
	(0.876)			(0.877)
Gender x $Log(Age)^2$	-0.356**			-0.344**
	(0.123)			(0.123)
Gender x Publicly Held		-0.078***		-0.081**
		(0.019)		(0.028)
Gender x Other Ownership		-0.009		-0.012
		(0.040)		(0.041)
Gender x Log(Firm Size)			-0.008**	0.001
			(0.003)	(0.004)

Note: N = 260,862 *p<0.05; **p<0.01; ***p<0.001

Figure 2: The Gender Gap in Discussing Work-Life Balance Over the Age Range

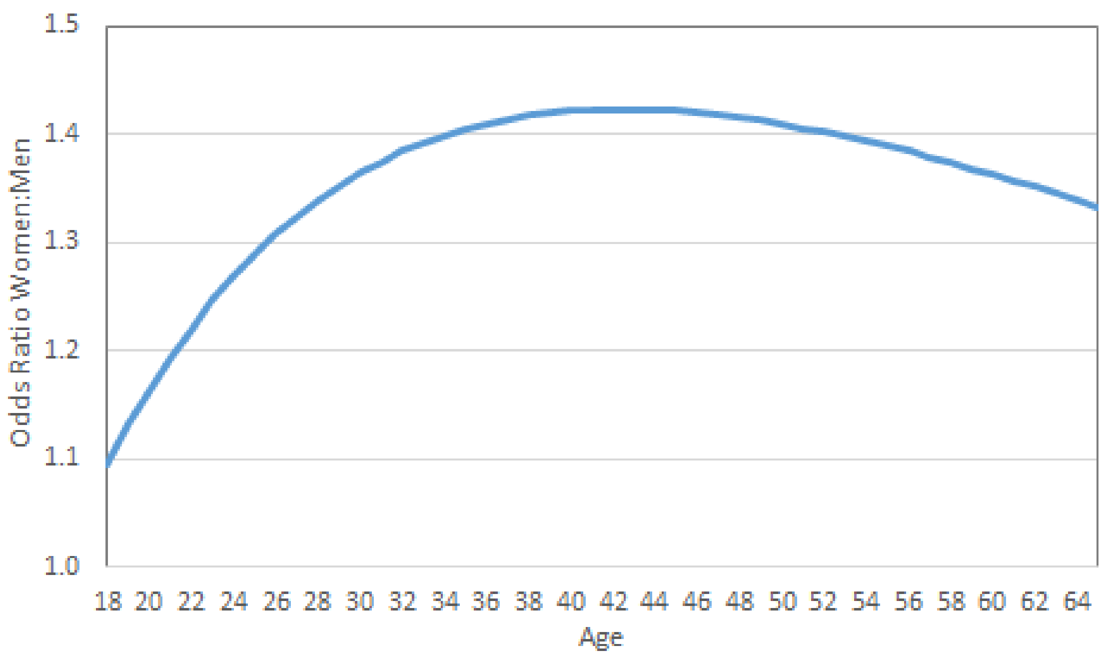

Table 5: Logistic Regression Analysis of Mentioning Work-Life Balance Component

	(1)	(2)	(3)	(4)
Computing HW & SW	0.407***	0.059	0.746***	0.774***
	(0.034)	(0.030)	(0.044)	(0.045)
Electrical Mfg	0.346***	-0.144***	0.496***	0.514***
	(0.041)	(0.039)	(0.052)	(0.053)
Bus SW & NW Solns	0.383***	0.062	0.767***	0.793***
	(0.035)	(0.032)	(0.046)	(0.047)
IT Svcs	0.213***	0.040	0.473***	0.496***
	(0.035)	(0.031)	(0.046)	(0.046)
Internet	0.194***	-0.007	0.519***	0.549***
	(0.036)	(0.032)	(0.046)	(0.047)
Telecomm Mfg	0.299***	-0.180*	0.524***	0.536***
	(0.075)	(0.078)	(0.093)	(0.095)
Telecomm Svcs	0.196***	-0.075*	0.336***	0.351***
	(0.038)	(0.035)	(0.049)	(0.050)
Video Games	0.420***	-0.098	0.683***	0.711***
	(0.058)	(0.061)	(0.071)	(0.072)
Great Lakes	-0.113***	0.191***	-0.267***	-0.268***
	(0.026)	(0.026)	(0.030)	(0.030)
Mideast	-0.164***	0.160***	-0.228***	-0.229***
	(0.023)	(0.023)	(0.026)	(0.026)
Mountains	-0.178***	0.193***	-0.297***	-0.308***
	(0.033)	(0.033)	(0.039)	(0.040)
New England	-0.050	0.215***	-0.019	-0.025
	(0.032)	(0.032)	(0.036)	(0.036)
Plains	-0.118**	0.337***	-0.234***	-0.236***
	(0.038)	(0.036)	(0.044)	(0.044)
Southeast	-0.159***	0.242***	-0.357***	-0.365***
	(0.021)	(0.022)	(0.026)	(0.026)
Southwest	-0.171***	0.182***	-0.317***	-0.322***
	(0.023)	(0.023)	(0.027)	(0.027)
Unknown region	-0.279***	0.049**	-0.300***	-0.302***
	(0.015)	(0.016)	(0.017)	(0.017)
Discuss Compensn	0.393***	0.574***	0.371***	0.374***
	(0.012)	(0.012)	(0.014)	(0.014)
Discuss Benefits	-2.756***	-1.990***	-2.929***	-2.970***
	(0.224)	(0.213)	(0.260)	(0.262)
Publicly Held	0.008	-0.140***	0.101***	0.103***
	(0.017)	(0.017)	(0.019)	(0.019)
Other Ownership	-0.050*	-0.108***	-0.001	-0.001
	(0.024)	(0.024)	(0.029)	(0.029)
Log(Firm Size)	0.014***	0.015***	0.061***	0.062***
	(0.003)	(0.003)	(0.003)	(0.003)
Gender (F=1)	0.197***	0.360***	0.136***	0.129***
	(0.012)	(0.011)	(0.013)	(0.013)
Log(Age)	20.725***	4.187***	21.937***	22.052***
	(0.548)	(0.504)	(0.647)	(0.651)
Log(Age)^2	-2.917***	-0.621***	-3.123***	-3.143***
	(0.077)	(0.071)	(0.091)	(0.092)

Note: N = 260,862 *p<0.05; **p<0.01; ***p<0.001

Table 6: Logistic Regression Analysis of Mentioning Work-Life Balance without Age

	(1)	(2)	(3)	(4)
Computing HW & SW	0.367***	0.366***	0.377***	0.366***
	(0.018)	(0.018)	(0.018)	(0.018)
Electrical Mfg	0.073**	0.072**	0.076***	0.072**
	(0.022)	(0.022)	(0.022)	(0.022)
Bus SW & NW Solns	0.358***	0.358***	0.369***	0.358***
	(0.019)	(0.019)	(0.019)	(0.019)
IT Svcs	0.205***	0.206***	0.217***	0.206***
	(0.019)	(0.019)	(0.018)	(0.019)
Internet	0.192***	0.192***	0.200***	0.192***
	(0.019)	(0.019)	(0.019)	(0.019)
Telecomm Mfg	0.079	0.079	0.088*	0.079
	(0.043)	(0.043)	(0.043)	(0.043)
Telecomm Svcs	0.127***	0.127***	0.136***	0.127***
	(0.020)	(0.020)	(0.020)	(0.020)
Video Games	0.232***	0.234***	0.224***	0.233***
	(0.037)	(0.037)	(0.037)	(0.037)
Great Lakes	-0.022	-0.022	-0.019	-0.021
	(0.015)	(0.015)	(0.015)	(0.015)
Mideast	-0.056***	-0.056***	-0.054***	-0.056***
	(0.013)	(0.013)	(0.013)	(0.013)
Mountains	-0.015	-0.015	-0.013	-0.015
	(0.019)	(0.019)	(0.019)	(0.019)
New England	0.070***	0.070***	0.071***	0.070***
	(0.018)	(0.018)	(0.018)	(0.018)
Plains	0.061**	0.061**	0.063**	0.061**
	(0.021)	(0.021)	(0.021)	(0.021)
Southeast	-0.025*	-0.024	-0.022	-0.024*
	(0.012)	(0.012)	(0.012)	(0.012)
Southwest	-0.064***	-0.063***	-0.063***	-0.063***
	(0.013)	(0.013)	(0.013)	(0.013)
Unknown region	-0.128***	-0.128***	-0.127***	-0.128***
	(0.009)	(0.009)	(0.009)	(0.009)
2014	0.069***	0.070***	0.067***	0.070***
	(0.017)	(0.017)	(0.017)	(0.017)
2015	-0.038*	-0.038*	-0.039*	-0.038*
	(0.016)	(0.016)	(0.016)	(0.016)
2016	-0.040**	-0.040**	-0.041**	-0.040**
	(0.015)	(0.015)	(0.015)	(0.015)
2017	-0.047**	-0.047**	-0.048**	-0.047**
	(0.015)	(0.015)	(0.015)	(0.015)
2018	-0.017	-0.017	-0.018	-0.017
	(0.016)	(0.016)	(0.016)	(0.016)
2019	0.021	0.021	0.020	0.021
	(0.016)	(0.016)	(0.016)	(0.016)
Discuss Compensn	0.496***	0.496***	0.496***	0.496***
	(0.007)	(0.007)	(0.007)	(0.007)
Discuss Benefits	-2.612***	-2.609***	-2.623***	-2.609***
	(0.124)	(0.124)	(0.124)	(0.124)

Table 6: Logistic Regression Analysis of Mentioning Work-Life Balance without Age (cont'd)

	(1)	(2)	(3)	(4)
Publicly Held	-0.055***	-0.030**		-0.024*
	(0.010)	(0.011)		(0.012)
Other Ownership	-0.094***	-0.089***		-0.087***
	(0.014)	(0.017)		(0.018)
Log(Firm Size)	0.007***	0.007***	0.003*	0.006**
	(0.001)	(0.001)	(0.001)	(0.002)
Gender (F=1)	0.303***	0.334***	0.347***	0.318***
	(0.007)	(0.009)	(0.017)	(0.018)
Gender x Publicly Held		-0.067***		-0.082***
		(0.014)		(0.020)
Gender x Other Ownership		-0.012		-0.018
		(0.028)		(0.028)
Gender x Log(Firm Size)			-0.006**	0.003
			(0.002)	(0.003)

Note: N = 545,849 *p<0.05; **p<0.01; ***p<0.001

Table 7: Logistic Regression Analysis of Mentioning Work-Life Balance without the Bottom Quartile of Firm Size

	(1)	(2)	(3)	(4)	(5)	(6)
Computing HW&SW	0.410***	0.409***	0.416***	0.416***	0.415***	0.415***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Electrical Mfg	0.145***	0.147***	0.175***	0.174***	0.174***	0.174***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Bus SW & NW Solns	0.392***	0.393***	0.401***	0.402***	0.401***	0.401***
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
IT Svcs	0.232***	0.235***	0.238***	0.239***	0.239***	0.238***
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Internet	0.173***	0.165***	0.167***	0.168***	0.167***	0.166***
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
Telecomm Mfg	0.197**	0.200**	0.234***	0.234***	0.235***	0.235***
	(0.068)	(0.068)	(0.068)	(0.069)	(0.069)	(0.069)
Telecomm Svcs	0.139***	0.138***	0.138***	0.138***	0.138***	0.139***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Video Games	0.162**	0.157**	0.127*	0.128*	0.129*	0.128*
	(0.059)	(0.059)	(0.059)	(0.059)	(0.059)	(0.059)
Great Lakes	-0.081***	-0.083***	-0.070**	-0.071**	-0.069**	-0.070**
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Mideast	-0.098***	-0.100***	-0.091***	-0.091***	-0.090***	-0.090***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Mountains	-0.070*	-0.071*	-0.070*	-0.069*	-0.069*	-0.069*
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
New England	0.067*	0.066*	0.091**	0.090**	0.091**	0.091**
	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)	(0.030)
Plains	0.027	0.026	0.032	0.030	0.032	0.032
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Southeast	-0.058**	-0.057**	-0.046*	-0.046*	-0.046*	-0.046*
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Southwest	-0.089***	-0.089***	-0.083***	-0.083***	-0.083***	-0.083***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Unknown region	-0.178***	-0.182***	-0.174***	-0.174***	-0.174***	-0.174***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)
2014	0.077**	0.078**	0.095***	0.095***	0.095***	0.095***
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
2015	0.004	0.003	0.021	0.021	0.021	0.021
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
2016	-0.003	-0.004	0.019	0.019	0.019	0.019
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
2017	0.004	0.003	0.029	0.028	0.028	0.029
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
2018	0.026	0.024	0.049	0.050	0.049	0.049
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
2019	0.047	0.045	0.063*	0.063*	0.063*	0.063*
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Discuss Compensn	0.460***	0.457***	0.445***	0.445***	0.445***	0.445***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)

Table 7: Logistic Regression Analysis of Mentioning Work-Life Balance without Bottom Quartile Firm Size (cont'd)

	(1)	(2)	(3)	(4)	(5)	(6)
Discuss Benefits	-2.498***	-2.531***	-2.488***	-2.489***	-2.488***	-2.492***
	(0.203)	(0.203)	(0.204)	(0.204)	(0.204)	(0.204)
Publicly Held	-0.080***	-0.079***	-0.083***	-0.083***	-0.044*	-0.083***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.017)	(0.015)
Other Ownership	-0.104***	-0.103***	-0.104***	-0.104***	-0.092**	-0.104***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.028)	(0.022)
Log(Firm Size)	0.021***	0.020***	0.024***	0.023***	0.023***	0.030***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Gender (F=1)	0.299***	0.298***	0.310***	-4.614**	0.372***	0.456***
	(0.011)	(0.011)	(0.011)	(1.772)	(0.018)	(0.041)
Log(Age)		-0.087***	14.684***	13.653***	14.684***	14.684***
		(0.019)	(0.487)	(0.619)	(0.487)	(0.487)
Log(Age)^2			-2.077***	-1.940***	-2.077***	-2.077***
			(0.069)	(0.087)	(0.069)	(0.069)
Gender x Log(Age)				2.642**		
				(1.000)		
Gender x Log(Age)^2				-0.351*		
				(0.141)		
Gender x Publicly Held					-0.103***	
					(0.023)	
Gender x Other Ownership					-0.029	
					(0.045)	
Gender x Log(Firm Size)						-0.016***
						(0.004)

Note: N = 197,947 *p<0.05; **p<0.01; ***p<0.001

Table 8: Logistic Regression Analysis of Mentioning Work-Life Balance by Overall Rating

Subsample	Lower-Rated Reviews (1 or 2 stars)			Higher-Rated Reviews (4 or 5 stars)		
	(1)	(2)	(3)	(4)	(5)	(6)
Computing HW&SW	0.314***	0.314***	0.314***	0.376***	0.376***	0.385***
	(0.046)	(0.046)	(0.046)	(0.041)	(0.041)	(0.041)
Electrical Mfg	0.059	0.059	0.058	0.223***	0.221***	0.221***
	(0.060)	(0.060)	(0.060)	(0.049)	(0.049)	(0.049)
Bus SW & NW Solns	0.309***	0.309***	0.309***	0.318***	0.318***	0.327***
	(0.049)	(0.049)	(0.048)	(0.042)	(0.042)	(0.042)
IT Svcs	0.154**	0.155**	0.156***	0.280***	0.280***	0.291***
	(0.047)	(0.047)	(0.047)	(0.042)	(0.042)	(0.042)
Internet	0.127*	0.127*	0.126*	0.206***	0.206***	0.211***
	(0.050)	(0.050)	(0.050)	(0.043)	(0.043)	(0.042)
Telecomm Mfg	0.107	0.108	0.107	0.149	0.148	0.153
	(0.114)	(0.114)	(0.114)	(0.093)	(0.093)	(0.093)
Telecomm Svcs	0.090	0.090	0.090	0.145**	0.146**	0.152***
	(0.054)	(0.054)	(0.054)	(0.046)	(0.046)	(0.045)
Video Games	0.101	0.102	0.098	0.198**	0.201**	0.190**
	(0.100)	(0.100)	(0.100)	(0.068)	(0.068)	(0.068)
Great Lakes	-0.009	-0.009	-0.008	0.009	0.011	0.013
	(0.044)	(0.044)	(0.044)	(0.028)	(0.028)	(0.028)
Mideast	0.036	0.036	0.037	-0.072**	-0.071**	-0.068**
	(0.039)	(0.039)	(0.039)	(0.024)	(0.024)	(0.024)
Mountains	0.093	0.093	0.093	-0.034	-0.034	-0.032
	(0.056)	(0.056)	(0.056)	(0.035)	(0.035)	(0.035)
New England	0.101	0.101	0.101	0.035	0.035	0.036
	(0.057)	(0.057)	(0.057)	(0.034)	(0.034)	(0.034)
Plains	0.087	0.087	0.088	0.120**	0.120**	0.122**
	(0.064)	(0.064)	(0.064)	(0.041)	(0.041)	(0.041)
Southeast	-0.028	-0.028	-0.027	0.026	0.026	0.030
	(0.037)	(0.037)	(0.037)	(0.023)	(0.023)	(0.023)
Southwest	-0.060	-0.060	-0.059	0.011	0.012	0.012
	(0.040)	(0.040)	(0.040)	(0.025)	(0.025)	(0.025)
Unknown region	-0.018	-0.018	-0.017	-0.167***	-0.167***	-0.166***
	(0.027)	(0.027)	(0.027)	(0.016)	(0.016)	(0.016)
2014	0.072	0.072	0.071	0.049	0.049	0.048
	(0.045)	(0.045)	(0.045)	(0.030)	(0.030)	(0.030)
2015	-0.020	-0.020	-0.020	-0.032	-0.032	-0.033
	(0.044)	(0.044)	(0.044)	(0.029)	(0.029)	(0.029)
2016	-0.031	-0.031	-0.031	0.015	0.015	0.014
	(0.044)	(0.044)	(0.044)	(0.028)	(0.028)	(0.028)
2017	-0.044	-0.044	-0.045	0.026	0.027	0.027
	(0.044)	(0.044)	(0.044)	(0.028)	(0.028)	(0.028)
2018	-0.013	-0.013	-0.014	0.055	0.055	0.055
	(0.045)	(0.045)	(0.045)	(0.028)	(0.028)	(0.028)
2019	-0.031	-0.031	-0.032	0.090**	0.090**	0.090**
	(0.046)	(0.046)	(0.046)	(0.029)	(0.029)	(0.029)
Discuss Compensn	0.514***	0.514***	0.514***	0.589***	0.588***	0.589***
	(0.021)	(0.021)	(0.021)	(0.014)	(0.014)	(0.014)
Discuss Benefits	-4.312***	-4.310***	-4.316***	-2.818***	-2.811***	-2.830***
	(0.424)	(0.424)	(0.424)	(0.234)	(0.234)	(0.233)

Table 8: Logistic Regression Analysis of Mentioning Work-Life Balance by Overall Rating (cont'd)

Subsample	Lower-Rated Reviews (1 or 2 stars)			Higher-Rated Reviews (4 or 5 stars)		
	(1)	(2)	(3)	(4)	(5)	(6)
Publicly Held	-0.026	-0.023	-0.026	-0.064***	-0.021	-0.064***
	(0.028)	(0.032)	(0.028)	(0.018)	(0.021)	(0.018)
Other Ownership	-0.017	-0.028	-0.017	-0.105***	-0.101**	-0.105***
	(0.038)	(0.048)	(0.038)	(0.028)	(0.035)	(0.028)
Log(Firm Size)	0.037***	0.037***	0.037***	-0.004	-0.004	0.000
	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)
Gender (F=1)	0.280***	0.281***	0.285***	0.338***	0.388***	0.422***
	(0.019)	(0.027)	(0.050)	(0.013)	(0.017)	(0.031)
Log(Age)	14.247***	14.247***	14.247***	14.414***	14.422***	14.422***
	(0.930)	(0.930)	(0.930)	(0.560)	(0.560)	(0.560)
Log(Age)^2	-2.031***	-2.031***	-2.031***	-2.022***	-2.023***	-2.023***
	(0.130)	(0.130)	(0.130)	(0.079)	(0.079)	(0.079)
Gender x Publicly Held		-0.009			-0.116***	
		(0.040)			(0.026)	
Gender x Other Ownership		0.028			-0.007	
		(0.074)			(0.056)	
Gender x Log(Firm Size)			-0.001			-0.012**
			(0.006)			(0.004)

Note: N = 68,395 (models 1-3) N = 145,766 (models 4-6) *p<0.05; **p<0.01; ***p<0.001

Figure 3a: Distribution of Employee Opinions about Work-Life Balance (<u>Pros</u> Section of Reviews)

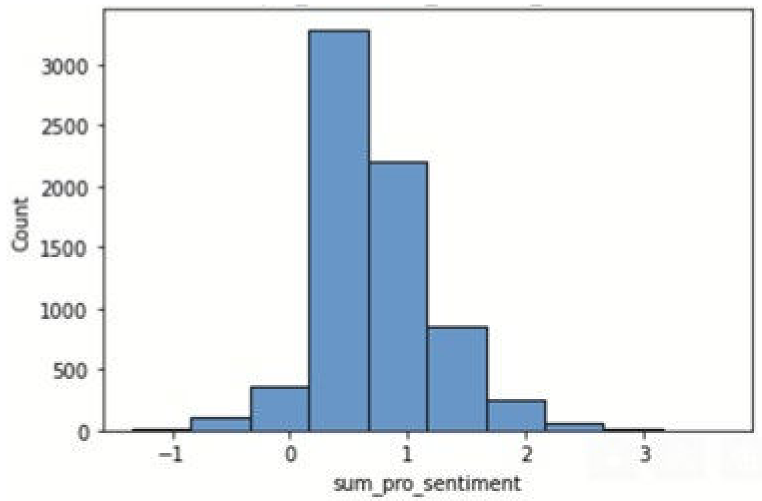

Figure 3b: Distribution of Employee Opinions about Work-Life Balance (<u>Cons</u> Sections of Reviews)

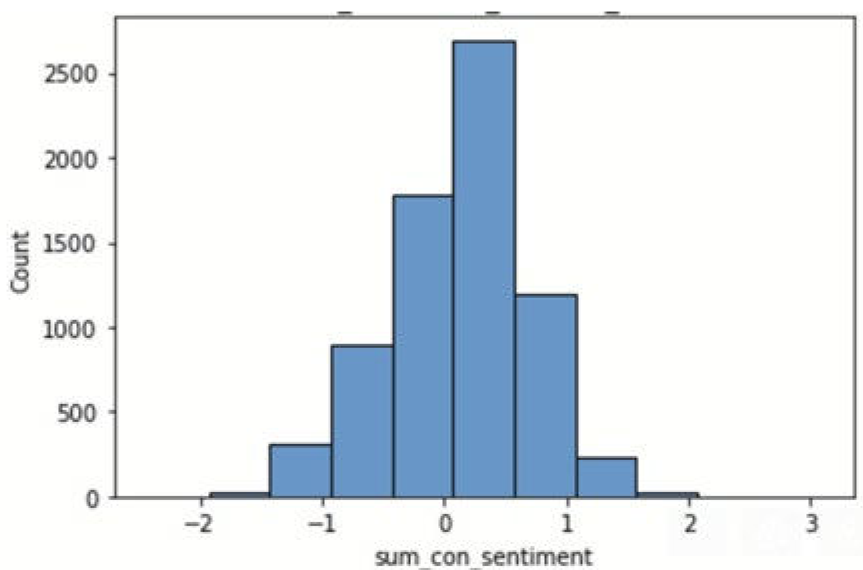

Figure 3c: Distribution of Employee Opinions about Work-Life Balance (Pros + Cons Sections Combined)

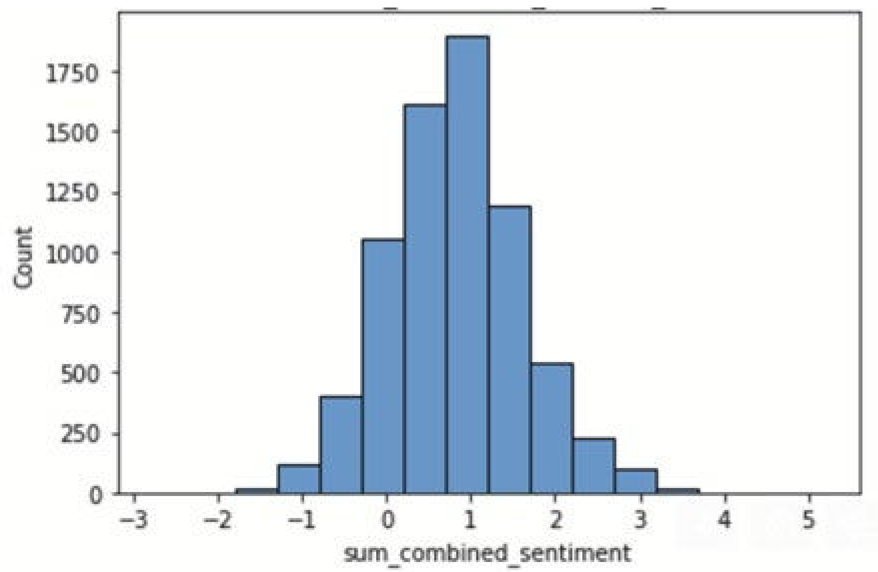

Table 9: Descriptive Statistics for OLS Regression Variables

	Variable	1	3	4	5	6	7
	Mean	0.5365	0.4529	0.0616	7.5434	0.3936	3.5076
	Standard Deviation	0.6505	0.4978	0.2403	3.2046	0.4886	0.2596
	Minimum	-2.4043	0	0	0	0	2.8904
	Maximum	4.4358	1	1	13.122	1	4.1744
	Number of Observations	65,943	65,943	65,943	65,943	65,943	65,943
1	WLBal Sentiment						
2	Publicly Held	-0.0212					
3	Other Ownership	-0.0055	-0.2330				
4	Log(Firm Size)	-0.0452	0.6972	-0.0380			
5	Gender (F=1)	-0.0021	-0.0290	-0.0005	-0.0246		
6	Log(Age)	-0.0040	0.0265	0.0103	0.0456	-0.0145	

Table 10a: OLS Regression Analysis of Opinions about Work-Life Balance

	(1)	(2)	(3)	(4)	(5)
Computing HW & SW	0.198***	0.198***	0.198***	0.198***	0.198***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Electrical Mfg	0.125***	0.125***	0.125***	0.125***	0.127***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Bus SW & NW Solns	0.231***	0.231***	0.231***	0.231***	0.232***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
IT Svcs	0.164***	0.164***	0.164***	0.164***	0.165***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Internet	0.122***	0.122***	0.122***	0.121***	0.121***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Telecomm Mfg	0.188***	0.188***	0.189***	0.189***	0.192***
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
Telecomm Svcs	0.006	0.006	0.006	0.006	0.006
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Video Games	0.154***	0.155***	0.154***	0.154***	0.152***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Great Lakes	-0.057***	-0.057***	-0.057***	-0.057***	-0.056***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Mideast	-0.043***	-0.043***	-0.043***	-0.043***	-0.043***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Mountains	0.003	0.003	0.003	0.003	0.003
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
New England	-0.009	-0.009	-0.009	-0.009	-0.007
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Plains	-0.017	-0.017	-0.017	-0.017	-0.017
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Southeast	-0.083***	-0.083***	-0.083***	-0.083***	-0.082***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Southwest	-0.057***	-0.057***	-0.057***	-0.057***	-0.057***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Unknown region	-0.062***	-0.062***	-0.062***	-0.062***	-0.062***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Discuss Compensn	0.005	0.005	0.005	0.005	0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Discuss Benefits	0.003	0.003	0.003	0.003	0.004
	(0.108)	(0.108)	(0.108)	(0.108)	(0.108)
Publicly Held	0.020**	0.020**	0.020**	0.020**	0.020**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Other Ownership	0.009	0.009	0.009	0.009	0.010
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Log(Firm Size)	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Gender (F=1)		0.005		0.005	0.005
		(0.005)		(0.005)	(0.005)
Log(Age)			-0.012	-0.012	1.427***
			(0.010)	(0.010)	(0.241)
Log(Age)^2					-0.203***
					(0.034)

Note: N = 65,943 *p<0.05; **p<0.01; ***p<0.001

Table 10b: OLS Regression Analysis of Opinions about Work-Life Balance

	(6)	(7)	(8)	(9)
Computing HW & SW	0.198***	0.198***	0.198***	0.198***
	(0.015)	(0.015)	(0.015)	(0.015)
Electrical Mfg	0.127***	0.128***	0.127***	0.128***
	(0.019)	(0.019)	(0.019)	(0.019)
Bus SW & NW Solns	0.232***	0.232***	0.232***	0.232***
	(0.016)	(0.016)	(0.016)	(0.016)
IT Svcs	0.165***	0.165***	0.165***	0.165***
	(0.016)	(0.016)	(0.016)	(0.016)
Internet	0.122***	0.122***	0.121***	0.122***
	(0.016)	(0.016)	(0.016)	(0.016)
Telecomm Mfg	0.193***	0.193***	0.192***	0.193***
	(0.035)	(0.035)	(0.035)	(0.035)
Telecomm Svcs	0.006	0.006	0.006	0.006
	(0.017)	(0.017)	(0.017)	(0.017)
Video Games	0.152***	0.152***	0.152***	0.153***
	(0.028)	(0.028)	(0.028)	(0.028)
Great Lakes	-0.057***	-0.056***	-0.056***	-0.057***
	(0.012)	(0.012)	(0.012)	(0.012)
Mideast	-0.043***	-0.043***	-0.043***	-0.042***
	(0.010)	(0.010)	(0.010)	(0.010)
Mountains	0.004	0.003	0.003	0.004
	(0.015)	(0.015)	(0.015)	(0.015)
New England	-0.007	-0.007	-0.007	-0.007
	(0.014)	(0.014)	(0.014)	(0.014)
Plains	-0.017	-0.017	-0.017	-0.017
	(0.017)	(0.017)	(0.017)	(0.017)
Southeast	-0.082***	-0.082***	-0.082***	-0.082***
	(0.010)	(0.010)	(0.010)	(0.010)
Southwest	-0.057***	-0.057***	-0.057***	-0.057***
	(0.011)	(0.011)	(0.011)	(0.011)
Unknown region	-0.062***	-0.062***	-0.062***	-0.062***
	(0.007)	(0.007)	(0.007)	(0.007)
Discuss Compensn	0.005	0.005	0.005	0.005
	(0.006)	(0.006)	(0.006)	(0.006)
Discuss Benefits	0.003	0.005	0.004	0.006
	(0.108)	(0.108)	(0.108)	(0.108)
Publicly Held	0.020**	0.022**	0.020**	0.014
	(0.007)	(0.009)	(0.007)	(0.010)
Other Ownership	0.010	-0.010	0.010	-0.013
	(0.011)	(0.014)	(0.011)	(0.014)
Log(Firm Size)	-0.008***	-0.008***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Gender (F=1)	-1.061	0.006	0.029*	-1.016
	(0.871)	(0.007)	(0.013)	(0.871)
Log(Age)	1.205***	1.428***	1.428***	1.213***
	(0.309)	(0.241)	(0.241)	(0.309)
$Log(Age)^2$	-0.174***	-0.203***	-0.203***	-0.175***
	(0.044)	(0.034)	(0.034)	(0.044)

Table 10b: OLS Regression Analysis of Opinions about Work-Life Balance (cont'd)

	(6)	(7)	(8)	(9)
Gender x Log(Age)	0.567			0.554
	(0.492)			(0.492)
Gender x Log(Age)^2	-0.075			-0.072
	(0.069)			(0.069)
Gender x Publicly Held		-0.007		0.013
		(0.011)		(0.015)
Gender x Other Ownership		0.050*		0.058*
		(0.022)		(0.022)
Gender x Log(Firm Size)			-0.003	-0.004*
			(0.002)	(0.002)

Note: N = 65,943 *p<0.05; **p<0.01; ***p<0.001

Appendix: Data Preparation and Analysis

Here we describe how we prepared the Glassdoor data for analysis and provide more details on the NLP techniques we used.

Data preprocessing

We used standard NLP techniques (i.e., regular expressions) to pre-process the text data. The dataset originally contained 1,128,375 reviews from 2008 to (October) 2020. To remove reviews written in other languages, we used lists of common words in English texts (stop words, which add little informational value to the texts in which they appear, such as "the," "too," "and," and "of"). For this, we used the intersection of three commonly used stop-word lists (from the popular Python libraries NLTK, scikit-learn, and spaCy) and eliminated reviews that contained none of those words.

We transformed some symbols into words: "&" became "and," "%" became "percent," and "@" became "at." We removed most punctuation marks and symbols. We retained the symbols "/" and ":" when they were surrounded by numbers, and the symbol "$" when it was followed by numbers. We replaced hyphens with "to" when surrounded by numbers and retained them otherwise. We also normalized the text – i.e., we turned upper-case letters into lower-case ones – which increased the frequencies of rare words. Finally, we removed excess white space (blanks).

Next, we removed stop words if they did not reduce the informational content, such as "the," "and," "or," and "if." We began with the union of stop-word lists from three popular Python libraries, NLTK, scikit-learn, and spaCy. After removing words on this union list, we compared the filtered strings to the original strings in two ways. First, we conducted visual analysis, examining each filtered, original string pair to determine if the current stop-word list was removing essential meaning. For this, we made value judgments based on the important keywords and information persisting after filtering. In particular, degrees of intensity (e.g., "more", "any," "very," and "always"), negations (e.g., "not," and "without"), and negative contractions (e.g., "don't" and "can't"). Second, we conducted data-driven analysis, generating a table to compare stop words with their frequency of occurrence in the entire dataset and the average shortening of each review when they were removed. These analyses led us to drop about 30 words from the union stop-word list. We then added some non-specific location words (e.g., “here” and “there”) and pronoun contractions (e.g., “I’m” and “we’re”) to the stop-word list. The end result was a list containing 242 words; it is available from the first author upon request.

We then eliminated short reviews (e.g., “Great firm! Excellent benefits”), which are often perfunctory and so not very informative about organizational culture and practices. Specifically, we eliminated 41,444 reviews containing 13 or fewer words (the tenth percentile of reviews by length). The final dataset contains 1,079,978 reviews from 2008 to 2020. When we limit the analysis to reviews posted 2014 onward, we are left with 948,785 reviews. There are 210,422 unique words in this dataset.

Sentiment analysis of work-life balance

To capture how employees felt about their firms’ norms and practices that affected work-life balance, we deployed NLTK, the most comprehensive Python library for NLP, which contains over 50 corpora and lexical resources (Bird, Klein, and Loper, 2009). NLTK is generally best for analyzing short texts like tweets, so it is appropriate for the short windows of words surrounding terms that denote work-life balance that we study here. We used VADER (Valence Aware Lexicon and sEntiment Reasoner), a lexicon and rule-based sentiment analysis tool in NLTK. VADER is sensitive to both polarity (i.e., positive or negative) and intensity (i.e., how positive or negative) of terms that indicate sentiment (e.g., good vs. great vs. bad). VADER performs just as well as, and in some cases better than, human coders in correctly labeling the sentiment of tweets (Hutto and Gilbert, 2014). VADER was created by drawing on existing dictionaries like General Inquirer (Stone, Bales, Bales, Namenwirth, and Ogilvie, 1962), Affective Norms for English Words (Bradley and Lang, 1999), and Linguistic Inquiry Word Count (Tausczik and Pennebaker, 2009; Pennebaker, Booth, Boyd, and Francis, 2015; LIWC).

VADER incorporates three key lexical features: a full list of Western-style emoticons (e.g., “:)” denotes a smiley face), sentiment-laden acronyms and initialisms (e.g., “LOL” denotes “laugh[ing] out loud”), and commonly used slang (e.g., “nah”, “meh”). To validate each feature’s sentiment score, ten human raters independently evaluated each feature on a scale from -4 to +4, where -4 stands for the most negative sentiment, +4 for the most positive sentiment, and 0 is neutral or neither or N/A. The scores for each rater were compared and reconciled. This procedure resulted in over 7,500 lexical features with validated valence scores that indicated both sentiment polarity and intensity (-4 to +4). For example, "okay" has a valence of 0.9, "good" 1.9, and "great" 3.1, whereas "horrible" has a valence of -2.5, the frowning emoticon :(-2.2, and "sucks" and "sux" -1.5 (Hutto and Gilbert, 2014).

To begin, VADER assigns each lexical unit (e.g., words, emoticons, initialisms) to one category: positive, negative, or neutral. Next, VADER makes use of five heuristic rules to calculate continuous valence (sentiment polarity) scores: (1) punctuation, namely the exclamation point; (2) capitalization, specifically all caps; (3) degree modifiers like extremely, slightly, and very; (4) polarity shift due to conjunctions like "but"; and (5) more complex polarity negation syntax, such as "The weather isn't really that hot." VADER's polarity scores consist of four values: negative (neg), neutral (neu), positive (pos), and compound. Neg, neu, and pos are ratios for the proportions of a text that fall in each category, which sum to 1. The compound score is computed by summing the valence scores of each word in a text, adjusted according to the five rules above, and normalizing the sum to create a measure that runs from -1 (most negative) to +1 (most positive). Thus, individual words have a sentiment score between -4 and +4, while texts have compound sentiment scores between -1 and +1.

Using VADER takes into account punctuation, capitalization and intensifiers (NLTK documentation), as Table A.1 below demonstrates. For example, in the first row, the sentence "The movie was bad," has a compound score of -0.5423, a negative score of 0.538, a neutral score of 0.462, and a positive score of 0.0. Our analysis began with windows of three words before and three words after each word or phrase denoting work-life balance. We then experimented with windows of five words but found no significant differences.

Table A.1: Example VADER Scores for Sentences

	Compound	Negative	Neutral	Positive
The movie was bad.	-0.5423	0.538	0.462	0.0
Very bad movie.	-0.5849	0.655	0.345	0.0
VERY bad movie.	-0.6732	0.694	0.306	0.0
VERY BAD movie!	-0.7616	0.735	0.265	0.0

References for Appendix[17]

Bradley, Margaret M., and Peter J. Lang. 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings. (https://www.semanticscholar.org/paper/Affective-Norms-for-English-Words-(ANEW)%3A-Manual-Bradley-Lang/c765eb0a31849361d829b24e173a37bab0919892)

Bird, Steven, Ewan Klein, and Edward Loper. 2009. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. Sebastopol, CA: O'Reilly Media, Inc.

Hutto, C., and Eric Gilbert. 2014. VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and social media*, 8(1): 216-225.

Pennebaker, James W., Roger J. Booth, and Martha E. Francis. 2015. *Linguistic Inquiry and Word Count: LIWC 2015*. Austin, TX: Pennebaker Conglomerates.

Stone, Phillip J., Bales, Robert F. Bales, J. Zvi Namenwirth, and Daniel M. Ogilvie. 1962. The General Inquirer: A computer system for content analysis and retrieval based on the sentence as a unit of information. *Behavioral Science*, 7 (4): 484-498.

Tausczik, Yla R., and James W. Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29 (1): 24-54.

[17] This list includes only those references that were not in the main reference list.