



OPEN

## Quantifying the impact of positive stress on companies from online employee reviews

Sanja Šćepanović<sup>1</sup>, Marios Constantinides<sup>1</sup>, Daniele Quercia<sup>1,2</sup>✉ & Seunghyun Kim<sup>3</sup>

Workplace stress is often considered to be negative, yet lab studies on individuals suggest that not all stress is bad. There are two types of stress: *distress* refers to harmful stimuli, while *eustress* refers to healthy, euphoric stimuli that create a sense of fulfillment and achievement. Telling the two types of stress apart is challenging, let alone quantifying their impact across corporations. By leveraging a dataset of 440 K reviews about S & P 500 companies published during twelve successive years, we developed a deep learning framework to extract stress mentions from these reviews. We proposed a new methodology that places each company on a stress-by-rating quadrant (based on its overall stress score and overall rating on the site), and accordingly scores the company to be, on average, either a *low stress*, *passive*, *negative stress*, or *positive stress* company. We found that (former) employees of positive stress companies tended to describe high-growth and collaborative workplaces in their reviews, and that such companies' stock evaluations grew, on average, 5.1 times in 10 years (2009–2019) as opposed to the companies of the other three stress types that grew, on average, 3.7 times in the same time period. We also found that the four stress scores aggregated every year—from 2008 to 2020—closely followed the unemployment rate in the U.S.: a year of positive stress (2008) was rapidly followed by several years of negative stress (2009–2015), which peaked during the Great Recession (2009–2011). These results suggest that automated analyses of the language used by employees on corporate social-networking tools offer yet another way of tracking workplace stress, allowing quantification of its impact on corporations.

According to the American Institute of Stress, 40% of workers consider their jobs to be stressful; a number that has significantly increased during the COVID-19 pandemic<sup>1</sup>. The World Health Organization treats stress as the number one health threat in the U.S., with more than 60% of doctor visits being due to a stress-related issue<sup>2</sup>. Workplace stress is often linked to lower motivation, poor performance, and decline in employees' well-being<sup>3</sup>, while it is estimated to amount to 190 billions in healthcare costs in the U.S. alone<sup>4</sup>. To currently track how its employees deal with stress, a large company would typically recruit consultants who would then administer surveys tailored to the company's situation, which typically end up being costly<sup>5</sup>, and restricted to a limited pool of self-selected participants<sup>6–8</sup>. The current situation points to the need of more research.

Stress is defined as “a set of physical and psychological responses to external conditions or influences, known as stressors”<sup>9</sup>. According to Lazarus<sup>10</sup>, “any change, either good (eustress) or bad (distress), is stressful, and whether it is a positive or a negative change, the physiological response is the same.” To cope with workplace stress, an employee has to cognitively acknowledge that a situation causes stress before even attempting to manage it<sup>11</sup>. Kobasa's framework of psychological hardiness offers three main categories of coping strategies<sup>12</sup>: *commitment* (having an active involvement in one's own work with a sense of purpose), *control* (believing and acting instead of feeling helpless in front of adversity), and *challenge* (believing that change could be a source of improvement). Kobasa posited that these categories could help individuals face challenges and, as such, individuals could turn stressful events into opportunities for personal growth<sup>12</sup>. However, despite having explored the relation between stress and job performance for decades, researchers have not yet established whether stress and performance are in a negative linear relation, a positive linear relation, or an inverted-U relation<sup>13</sup>.

To tackle this gap, we draw upon two streams of previous literature, that is, literature about stress in the corporate context, and literature on how to gauge stress from online data. In the literature about stress in the corporate context, stress is often being portrayed as negative<sup>14</sup>; and as a leading cause of death, poor work performance, and diminishing well-being<sup>3</sup>. More recently, however, researchers have advocated that there exist another type of

<sup>1</sup>Nokia Bell Labs, Cambridge, UK. <sup>2</sup>CUSP, Kings College London, London, UK. <sup>3</sup>Georgia Tech, Atlanta, USA. ✉email: quercia@cantab.net

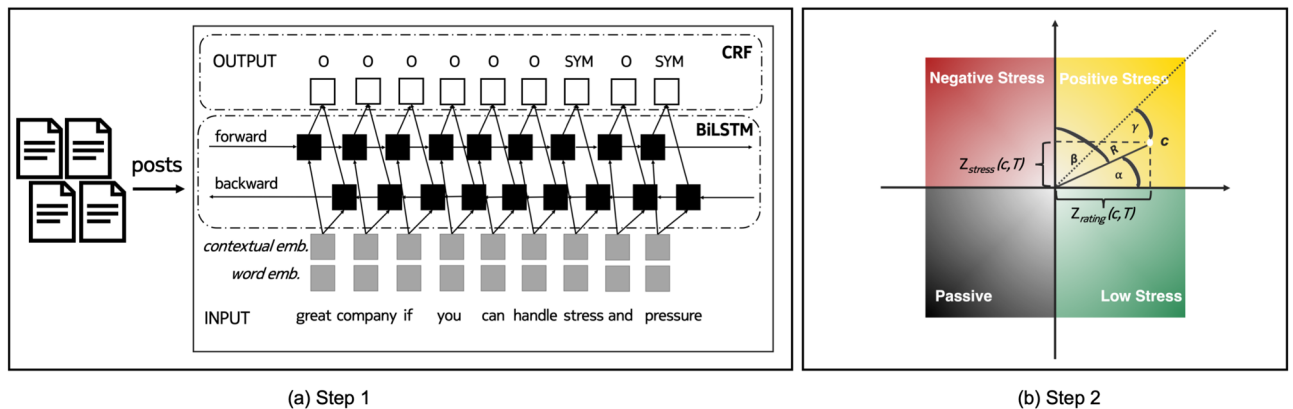
Stress Type	Review excerpt
Low stress	My company walks its talk. It [ <i>the company</i> ] takes care of customers and employees.
Negative stress	There is a feeling of scarcity due to the constant reorganizations, pressure, and surprise layoffs. [...] You could imagine how toxic the environment is.
Passive stress	There is no regard for how the remaining work will get done, just how the bottom line looks at that moment in time. People are not treated as respected contributors to the organization. [...] This is a very unstable, unhealthy, volatile, stressed out environment, with incredibly poor leadership.
Positive stress	You have to be a very driven and self-motivated person to be successful here. If you are willing to commit and put in the extra effort and hard work, it will be extremely worth it. [...] Every day is very busy and it can be stressful at times but its all worth it!

**Table 1.** Example reviews of companies of each stress type.

stress: *positive stress*. The idea is that whether stress is positive or negative depends on how an individual reacts to a stressor<sup>15</sup>. ‘One’s stress mindset can be conceptualized as the extent to which one holds the belief that stress has enhancing consequences for various stress-related outcomes such as performance and productivity, health and well-being, and learning and growth, or holds the belief that stress has debilitating consequences for those outcomes’<sup>15</sup>. Of prime importance is to distinguish appraisal from stress mindset. Stress mindset describes the evaluation of the nature of stress itself as positive or negative (i.e., enhancing or debilitating)<sup>15</sup>, whereas appraisal is about the evaluation of a particular stressor as more or less stressful<sup>16</sup>. For example, one may appraise a difficult task as highly stressful and have a stress debilitating mindset, which, in turn, leads the individual to experience the situation as draining (negative stress). By contrast, another individual may again consider the task as highly stressful but have a stress enhancing mindset, leading the individual to experience the situation as an opportunity for growth and development (positive stress). While stress is often linked to depression<sup>17,18</sup>, several accounts posit that certain stressful experiences may fundamentally change individuals for the better—a phenomenon referred to as stress-related growth<sup>15</sup>. The experience of stress can enhance the development of mental toughness, greater appreciation for life, and an increased sense of meaningfulness<sup>19,20</sup>. However, as Crum et al.<sup>15</sup> pointed out, these conflicting findings in the stress literature suggest a nuanced view of stress. A view that recognizes the debilitating nature of stress on health and performance, but can also account for its enhancing nature in specific circumstances. We hypothesized that the presence of both positive and negative stress can be measured from digital data based on previous literature that has done just that with different techniques upon different datasets. Guntuku et al.<sup>21</sup> used the Linguistic Inquiry and Word Count (LIWC)<sup>22</sup> dictionary’s features (e.g., topical categories, emotions, parts-of-speech) to predict stress of social media (Facebook and Twitter) users. Saha and De Choudhury<sup>23</sup> did a similar study but on Reddit and did so in conjunction with gun violence events, and found specific stress markers to be associated with linguistic changes about “higher self pre-occupation and death-related discussion.” Similar to our study, Vedant et al.<sup>24</sup> showed that the use of language in employee reviews can be used to operationalize organizational culture: the collection of values, expectations, and practices that guide and inform employees’ actions within a company.

Based on these preliminary findings, we hypothesized that workplace stress is reflected in company reviews. To explore this hypothesis, we placed companies on a 2x2 coordinate system, based on their overall stress scores and overall ratings on the company review site. This stress-by-rating quadrant effectively divided companies in four stress types that we termed low stress, passive, negative stress, and positive stress (Table 1 shows example reviews of companies of each stress type). Low stress companies enjoy high overall ratings and low stress scores. These are usually established organizations that offer workplace flexibility, good pay, and bonuses. Passive companies are characterized by low overall ratings and a small proportion of posts mentioning stress. They tend to have high turnover, due to repetitive workload and non-motivated employees. Negative stress companies are characterized by low overall ratings and a high proportion of posts mentioning stress. Employees of these companies are particularly unhappy as, in addition to the unsatisfactory conditions, they also experience high pressure. Finally, positive stress companies enjoy high overall ratings but also high stress scores. These tend to be inspiring, reputable workplaces that attract employees because of the collaborative atmosphere and career prospects despite the pressure employees are subject to. The project website for our study is found on <https://social-dynamics.net/positive-stress>.

**Data.** After obtaining the U.S. unemployment rates between 2008 and 2020 from the U.S. Bureau of Labor Statistics<sup>25</sup> and the S & P 500 stock market data (including the 500 large capital U.S. companies with a cumulative market capitalization to be around 70-80% of the total in the country) from the Yahoo Finance portal<sup>26</sup>, we matched that data with our company reviews. More specifically, we obtained 440K geo-referenced posts on Glassdoor (<https://www.glassdoor.com>), a popular company reviewing site about the S & P 500 companies published during twelve successive years, from 2008 to 2020. On this site, current and, more likely, former employees of companies write reviews about their own corporate experience, ranging from job interviews to salaries to workplace culture. The site provided an overall rating of each company based on employees’ reviews. As of 2021, there are 50M monthly visitors on the platform, and 70M reviews about 1.3M companies. To ensure quality reviews, the site employs the following three mechanisms. First, an automatic (proprietary) and manual content moderation system is paired with the possibility for users to flag content. Such a combined system partly prevents fake reviews (e.g., a company unfairly forcing employees to leave positive reviews). Second, every user wanting to browse others’ content has to take the time to write one review. This requirement encourages the



**Figure 1.** Placing companies on a stress-by-rating quadrant by detecting stress mentions in reviews about a company using a state-of-the-art NLP deep-learning framework (*Step 1*), placing the company in the rating-by-stress quadrant, and computing its association with its stress type (i.e., with its quadrant) (*Step 2*). To see how the association is computed, consider company  $c$  shown in (b) to be of positive stress.  $c$  is placed according to its  $z_{rating}(c, T)$  along the  $x$ -axis, and to its  $z_{stress}(c, T)$  along the  $y$ -axis.  $R$  is the radius from the center to  $c$ 's point;  $\alpha$  is the angle between the radius line and the  $x$ -axis;  $\beta$  is the angle between the radius line and the  $y$ -axis; and  $\gamma$  is the angle between the radius line and the diagonal shown as a dotted line. The function  $f(c, s, T)$  combines  $R$ ,  $\alpha$ ,  $\beta$ , and  $\gamma$ , and accordingly scores  $c$  to have a high association weight with positive stress  $s$  during period  $T$  (darker shades of colors), as  $c$  is close to the quadrant's diagonal, and distant from the intersection point.

presence of neutral reviews and partly prevents the so-called *non-response bias*, a situation in which the opinions of respondents are very different from those of non-respondents. Third, the site allows for a maximum of one review per employee per company per year, preventing any employee from contributing a disproportionate number of reviews, and, in so doing, discouraging *sampling bias*, a situation in which the opinions of some members are more represented in the data than those of others.

## Methods

We extracted mentions related to stress using an NLP deep-learning tool, which is trained to extract medical conditions from free-form text (Fig. 1a). We then designed a new methodology that placed the 500 S & P companies on a stress-by-rating quadrant based on their overall ratings on the reviewing site on one axis, and the presence of mentions related to stress in their reviews on the other axis (Fig. 1b). In so doing, we classified each company to be, on average, either a *low stress*, *passive*, *negative stress*, or *positive stress* company. We finally computed each company's strength of membership to its quadrant depending on whether the company is placed both close to the diagonal and far from the (0,0) intersection point. The function  $f(c, s, T)$ , which is expressed in Eq. (1) and graphically depicted in Fig. 1b, assigns a higher weight to company  $c$  of stress type  $s$ , if  $c$  is both closer to the quadrant's diagonal (i.e., it is farther from the remaining quadrants) and more distant from the two axes' intersection (i.e., it has higher absolute overall rating and stress score values). We call  $f(c, s, T)$  to be company  $c$ 's association with stress type  $s$  during  $T$  since the higher its value, the more  $c$  is associated with stress type  $s$  during  $T$ .

**Extracting stress mentions from posts** To extract stress mentions, we used the MedDL entity extraction module<sup>27</sup> (the left rectangle in Figure 1 (a)). MedDL uses *contextual embeddings* and a *deep BiLSTM-CRF sequence labeling architecture* (we used the default parameters as specified in<sup>27</sup>). The model was pre-trained and evaluated on a labeled dataset of Reddit posts called MedRed<sup>27</sup>. The MedRed dataset was split into train (50%), dev (25%), and test (25%) sets. We evaluated MedDL using the strict and relaxed *F1-scores*, two commonly used performance metrics for entity extraction models. The *strict F1-score* counts only the *exact* matches with the true labels as correct, while the *relaxed F1-score* takes into account the partially matching extracted entities. We provide formulae for the two scores in *Supplementary Information* (SI Eq. (1)). MedDL was compared against two well-known entity extraction tools: MetaMap (a well-established tool<sup>28</sup> and a de-facto baseline method for NLP studies related to health<sup>29</sup>) and TaggerOne (a machine learning tool using semi-Markov models and a medical lexicon<sup>30</sup>). The MedDL method achieved a strict/relaxed *F1-score* of .71/.85 when extracting symptoms (Figure S4), outperforming both MetaMap and TaggerOne by a large margin (the two have *F1-scores* of .17/.48 and .31/.58, respectively). Furthermore, MedDL has shown generalizability when applied on dataset (e.g., dream reports<sup>31</sup>) different than those it was trained on (i.e., Reddit data).

We extracted stress mentions in three steps (further detailed in *Supplementary Information*). First, we detected over 21K posts that mentioned over 5K unique medical conditions. Most frequent medical conditions identified include *stress*, *pain*, *headache*, and *depression*. Second, we inspected the top 200 most mentioned conditions and manually selected 31 of them that specifically reflect workplace stress (top 15 are shown in Table 2). Third, we extracted all reviews mentioning any of the 31 conditions. This resulted in 7,338 posts related to stress, which accounted for 1% of all posts. Despite this seemingly low number of posts, when aggregated, these posts returned statistically significant results for our metrics, which are described next.

**Associating stress types with companies** We placed each S & P 500 company on a stress-by-rating quadrant. More specifically, for each company  $c$ , we computed its average review rating and its stress score:

Condition related to stress	# mentions	Example mention
Stress	3473	"Great company to work for, if you can handle stress."
High stress	710	"High stress work environment, long work hours."
Pressure	447	"a lot of pressure to get things done."
Burnout	277	"[...], the ones who made the cut to stay are suffering from burnout."
Understaffing	99	"Somewhat job stability due to understaffing."
Heavy workload	58	"Lack of work/life balance, extremely heavy workload."
Exhaustion	58	"You will be pushed to the point of exhaustion [...]."
Stress levels	57	"[...] stress levels peak insanely when the store manager [...]."
Overworked	45	"At times, you can feel overworked and undervalued."
Tension	38	"There's a lot of tension between coworkers because of commission."
High workload	38	"[...] seeing many large set-backs which cause very high workload"
Extreme stress	33	"Beware: extreme stress and pressure."
Mental stress	23	"[...] ends up giving you a lot of mental stress."
Overload	17	"No work life balance [...], overloaded and benefits are not good."
Pressure to perform	9	"[...] a lot of pressure to perform, long working hours"

**Table 2.** Top15 most frequent stress-related mentions identified on a company review site, and their frequency counts.

$$\text{rating}(c, T) = c's \text{ average review rating during } T,$$

$$\text{stress}(c, T) = \frac{\# c's \text{ posts related to stress during } T}{\text{total } \# c's \text{ posts during } T}.$$

where  $T$  is set to initially include all the years under study (2009–2019). To then ease comparability, we z-scored these two values:

$$z_{\text{rating}}(c, T) = \frac{\text{rating}(c, T) - \mu_{\text{rating}}(T)}{\sigma_{\text{rating}}(T)},$$

$$z_{\text{stress}}(c, T) = \frac{\text{stress}(c, T) - \mu_{\text{stress}}(T)}{\sigma_{\text{stress}}(T)}.$$

where  $\mu_{\text{rating}}(T)$  and  $\sigma_{\text{rating}}(T)$  are the average and standard deviation of the review ratings for all companies (regardless of their stress types) during the whole period  $T$  (readily available on the company review site), and  $\mu_{\text{stress}}(T)$  and  $\sigma_{\text{stress}}(T)$  are the average and standard deviation of the stress scores for all companies during the whole period  $T$ .

Each S & P 500 company was assigned to one of the four quadrants based on the signs of their two z-scores (Fig. 1b). For example, a company  $c$  with a negative  $z_{\text{rating}}(c, T)$  and a positive  $z_{\text{stress}}(c, T)$  would be placed in the negative stress quadrant, while a company with a positive  $z_{\text{rating}}(c, T)$  and a positive  $z_{\text{stress}}(c, T)$  would be placed in the positive stress quadrant. The resulting quadrants are consequently four:

*Low Stress companies* These enjoy high overall ratings and low stress scores. Their employees tended to think very positively about their workplace experience with comparatively fewer posts mentioning stress conditions.

*Passive companies* These are characterized by low overall ratings and a small proportion of posts mentioning stress. Their employees were mostly not satisfied with their jobs, but they also showed comparatively fewer signs of stress in their reviews.

*Negative stress companies* These are characterized by high stress scores and low overall ratings. Their employees mentioned stress conditions, while also scoring their workplace experience low.

*Positive stress companies* These enjoy high ratings despite high stress scores. Their employees mentioned stress yet did so in the context of high-pressure and highly rewarding work environments.

Once a company  $c$  is placed in its quadrant (i.e., associated with its stress type  $s$ ), we needed to estimate its association with this quadrant, i.e., with  $s$ . For example, company  $c$  with  $(z_{\text{rating}}(c, T), z_{\text{stress}}(c, T))$  equal to (3,3) is more strongly associated with *positive stress*, than what a company with (0.5, 0.5) would be. To estimate  $c$ 's association with  $s$ , we combined  $c$ 's two z-scores concerning review rating and stress score as follows (and as depicted in Fig. 1b):

$$f(c, s, T) = \begin{cases} l(z_{rating}(c, T), z_{stress}(c, T)) = R/(\gamma + \pi), & \text{if } c \in s \text{ during } T; \\ 0, & \text{if } c \notin s, \text{ or } c \text{ has no review during } T; \end{cases}$$

where:

$$\begin{aligned} R &= \sqrt{z_{rating}(c, T)^2 + z_{stress}(c, T)^2}, \\ \gamma &= \max((\alpha - \pi/4), (\beta - \pi/4)), \\ \alpha &= \arccos(|z_{rating}(c, T)|/R), \\ \beta &= \arccos(|z_{stress}(c, T)|/R). \end{aligned} \quad (1)$$

where  $T$  is initially set to include all the years under study, from 2009 to 2019. To ease understanding of the above formula, consider that function  $l$ , on input of the two  $z$ -scores (i.e., the company's two coordinates in the quadrant), computes the extent to which company  $c$  is on the diagonal and far from the (0,0) intersection point (Fig. 1b). It gives higher weights to companies that are both closer to the quadrant's diagonal (i.e., which are farthest from the remaining quadrants) and more distant from the axes' intersection (i.e., which have higher absolute rating/stress score values).

**Computing stress scores over the years** For each year  $y$ , we quantified the amount of a given stress type  $s$  expressed in the posts produced in that year. More specifically, we computed:

$$m(s, y) = \sum_{c \in s} f(c, s, y) \times w(c, y, s), \quad (2)$$

For all the companies of stress type  $s$ , we summed each company's association  $f(c, s, y)$  with  $s$  during year  $y$  weighted by the presence of posts about the company during  $y$  (giving higher weights to companies whose employees contributed more reviews in that year):

$$w(c, y, s) = \begin{cases} \frac{\# c's \text{ posts in year } y}{\text{total } \# \text{ posts in year } y}, & \text{if } c \in s \text{ in year } y; \\ 0, & \text{if } c \text{ has no reviews in year } y. \end{cases} \quad (3)$$

**Associating topical categories with stress types** To identify relevant words for each stress type, we run BERTopic<sup>32</sup>, which is a state-of-the-art topic modeling algorithm. A topic modeling algorithm is an unsupervised technique to extract topics that appear frequently in a piece of text (in our case, a post). The algorithm works in four sequential steps:

1. converts each post into a 512-dimensional vector (called embedding) of numerical values using a pre-trained BERT-based sentence transformer<sup>33</sup> (in our case, we used the default model, that is, the "paraphrase-MiniLM-L6-v2"). BERT (Bidirectional Encoder Representations from Transformers) is a state-of-the-art transformer-based machine learning technique for natural language processing (NLP), which takes into account the context of each word.
2. reduces dimensionality using UMAP<sup>34</sup> (or Unification Map) for every embedding, as many clustering algorithms handle high dimensionality poorly. UMAP is arguably the best performing dimensionality reduction algorithm as it keeps significant portion of the high-dimensional structure in lower dimensionality.
3. uses HDBSCAN<sup>35</sup> for clustering with the "UMAP" embeddings, resulting in similar posts being clustered together. HDBSCAN is a density-based algorithm that works well with UMAP as the structure is preserved in a lower-dimensional space. Additionally, HDBSCAN does not force data points to clusters as it considers them outliers.
4. identifies keywords using the c-TF-IDF<sup>32</sup> score (Eq. 4), and using that score, derives topics from the identified keywords. To create a topic representation, we took the top 3 keywords per topic based on their c-TF-IDF scores. The higher the score, the more representative is as the score is a proxy of information density.

$$c\text{-TF-IDF}_l = \frac{k_l}{o_l} \times \frac{p}{\sum_j^q k_j} \quad (4)$$

where the frequency of each keyword  $k$  is extracted for each topic  $l$  and divided by the total number of keywords  $o$ . The total, unjoined, number of posts  $p$  is divided by the total frequency of keyword  $k$  across all topics  $q$ .

**Analysis plan.** Our analysis plan unfolded in three steps. First, as an initial validation step, we ascertained that stress was paraphrased in a company's reviews differently according to the company's stress type. Second, we tested whether the evolution of each stress score over the years tallied with large-scale exogenous events such as the Great Recession. Third, we tested that a company's stress type is partly associated with its stock growth.

## Results

**Topics discussed in reviews of companies of different stress types.** To ascertain whether the content of the reviews captured aspects specific to the four stress types, we identified the top relevant words for each type by running a topic modeling algorithm called BERTopic<sup>32</sup>, and did so on four distinct sets of reviews: each set contained reviews of all the companies of a given stress type. This algorithm found the emergent topics in the four sets, and Table 3 lists the the top three words for each topic. The top10 topics for each quadrant are

	Negative stress	Passive	Low stress	Positive stress
Career drivers	Overtime	Vacation	Scheduling	Teamwork
	Mandatory	pto	Flexibility	Supportive
	Shifts	Vacation/sick	Autonomy	Collaborative
Industry or benefits	Discounts	Insurance	401k	Gain
	Sale	Espp	Bonus	Billions
	Coupons	Hsa	Retirement	Software
Emotional aspects	Horrible	Repetitive	Fair	Prestige
	Terrible	Turnover	Friendly/good	Boost
	Awful	Workload	Pays	Reputation

**Table 3.** Three-word groups present in the reviews of companies of the four stress types. These groups were automatically found by BERTopic and speak to three main workplace characteristics: career drivers, industry and benefits, and emotional aspects. For each group, the top three words are shown together with their normalized word importance. Abbreviations of words describing monetary benefits include pto (paid time off); espp (employee stock purchase plan); hsa (health savings account); 401 k (a retirement savings and investing plan that employers offer).

statistically associated with the quadrant. That is, based on chi-square tests, each topic  $l$  associated with quadrant  $s$ : has frequency in  $s$  always above zero (is dependent on  $s$ ), and is independent of any quadrant other than  $s$ . As detailed in *Supplementary Information*, by inspecting these groups of words and corresponding representative reviews, six annotators identified the emergence of three workplace themes:

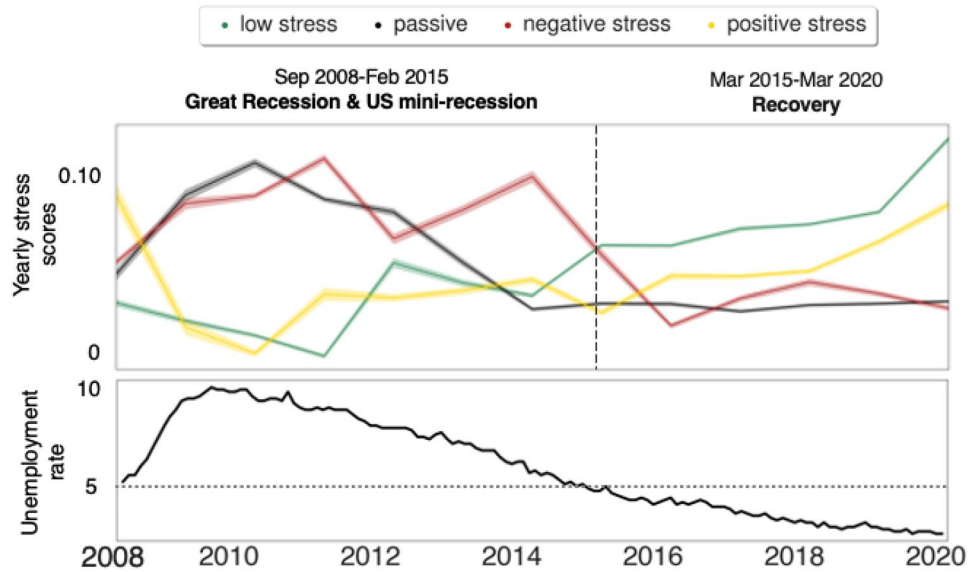
*Career drivers* (first set of rows in Table 3). Negative stress companies were associated with words such as ‘overtime’, ‘mandatory’, ‘shifts’, and the typical workplace described in the reviews, according to our annotators, was characterized by considerable emotional pressure. On the other hand, passive companies were associated with words such as ‘vacation’, ‘pto’, and ‘vacation/sick’, and the corresponding reviews tended to deflect from the day-to-day work and focus on activities outside work such as vacation and time off. Low stress companies were associated with words such as ‘scheduling’, ‘flexibility’, and ‘autonomy’, and the typical workplace described in the reviews was one in which employees cherished their sense of control over their work. Finally, positive stress companies were associated with words such as ‘teamwork’, ‘supportive’, and ‘collaborative’, and the typical workplace in the reviews was one with a collaborative and supportive culture.

*Industry or benefits* (second set of rows in Table 3). Negative stress companies were associated with words such as ‘discounts’, ‘sale’, ‘coupons’, while positive stress companies were associated with words such as ‘gain’, ‘billions’, and ‘software’. Their reviews were effectively mentioning the industry sectors they referred to: Consumer Discretionary (e.g., retail shops) for the reviews of negative stress companies, and Information Technology for those of positive stress ones. On the other hand, passive companies were associated with words such as ‘insurance’, ‘espp’, and ‘hsa’, and, similarly, low stress ones with words such as ‘401 k’, ‘bonus’, and ‘retirement’; the corresponding reviews indicated workplaces in which concerns about long-term financial benefits rather than the presence of implicit incentives in one’s own work were at the forefront.

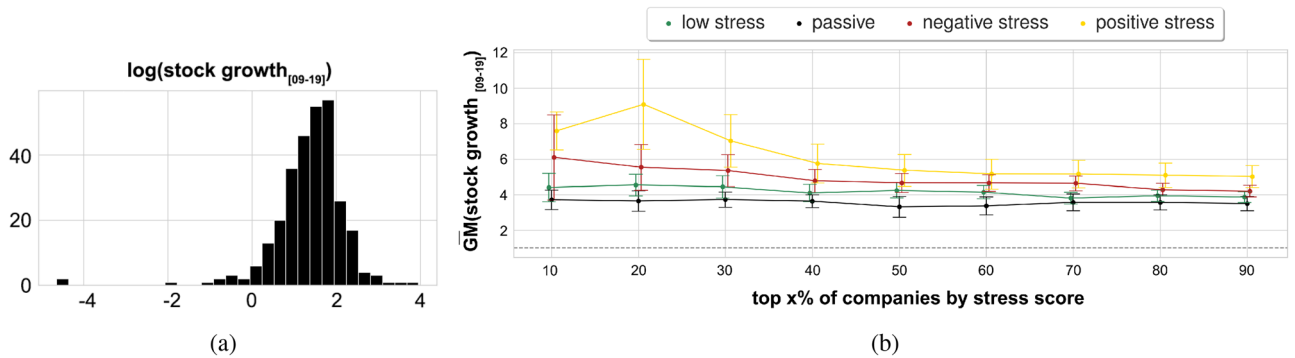
*Emotional Aspects* (third set of rows in Table 3). Negative stress companies were associated with words such as ‘horrible’, ‘terrible’, and ‘awful’, confirming, once again, the presence of emotional pressure. Passive companies were instead associated with words such as ‘repetitive’, ‘turnover’, and ‘workload’, confirming the tedious nature of those workplaces. Low stress companies were associated with words such as ‘fair’, ‘friendly/good’, and ‘pays’, and the corresponding reviews described a good work-life balance. Finally, positive stress companies were associated with words such as ‘prestige’, ‘boost’, and ‘reputation’, and their reviews described high performing, dynamic, and fast-paced workplaces.

**Evolution of stress types and the Great Recession.** After the preliminary validation step in which we ascertained that stress was paraphrased in reviews differently according to the stress type, we tested whether the evolution of each stress score over the years tallied with large-scale exogenous events such as the Great Recession. We plotted the amount  $m(s, y)$  of each stress score  $s$  in each year  $y$  (as per Equation (2)), from 2008 to 2020 (top panel in Fig. 2). The overall changes closely followed the unemployment rates from the U.S. Bureau of Labor (bottom panel in Fig. 2): a year of positive stress (2008) was rapidly followed by several years of negative stress (2009–2015), which peaked during the Great Recession (2009–2011) during which the U.S. went through a loss of over 7.5 million jobs and high unemployment rates<sup>36</sup>.

**Stock growth of companies of different stress types.** Finally, we hypothesized that a company’s way of dealing with workplace stress was partly associated with performance. Given our data, we cannot study whether stress causes (poor) performance but can only study how the two are associated. Also, there is no company performance measure that is solely affected by a company’s stress culture. As such, our stress scores are



**Figure 2.** The evolution of: (top) the four types of stress; and (bottom) the unemployment rate in the U.S., with the horizontal dashed line reflecting pre-recession rate. The stress score per year is calculated using Eq. (2), and its standard deviations are shown with shaded lines.



**Figure 3.** (a) Distribution across companies of the logarithm of stock growth values from the average stock price in 2009 and that of 2019 ( $stock\_growth_{09-19} = stock_{2019}/stock_{2009}$ ) showing the stock growth is log-normally distributed. The average stock price for year  $y$  ( $stock_y$ ) is calculated as the average of the daily Adjusted Closing Prices for the year. (b) Geometric mean of the stock growth values  $\bar{GM}(stock\_growth_{09-19})$  for increasing stress score percentiles for the companies of a given stress type. Error bars represent geometric standard error  $GSE(stock\_growth_{09-19}) = \bar{GM}(stock\_growth_{09-19}) / \sqrt{N} \cdot \sigma(\log(stock\_growth_{09-19}))$ .

unlikely to be predictive of any company-wide performance measure. We opted for long-term stock growth as our performance measure, not least because it is publicly available and standardized across companies. However, such a growth is partly affected by a company’s culture, and conflates endogenous factors (e.g., productivity) with exogenous ones (e.g., financial cycles). Yet we expected our stress measures to qualitatively describe different forms of financial success, at least in the long term. To that end, we computed stock growth during the full period of study, that is, between 2009 and 2019:

$$stock\ growth_{09-19}(c) = \frac{stock(c)_{2019}}{stock(c)_{2009}} \tag{5}$$

where  $stock^i$  is the average adjusted closing price of company  $c$ ’s stock in year  $i$ . We chose long-term growth instead of short-term one (e.g., that pertaining 2018-2019) to partly account for any potential influence of exogenous events (e.g., Great Recession, market manipulation, incidental growth/decline<sup>37</sup>). In *Supplementary Information*, we show that the results do not qualitatively change when considering the narrower 5-year period from 2014 to 2019. Given a stress type  $s$ , we computed company  $c$ ’s association  $f(c, s, T)$  with  $s$  during time period  $T$  (initially set to the whole period of study), consequently grouping all the companies of a given stress type into their stress score percentiles (Fig. 3b). As the distribution of stock growth values across companies is heavy-tailed (Fig. 3a), we used the geometric mean to average these values across companies. That is,

$GM(\text{stock growth}_{[09-19]}) = \Pi(\text{stock growth}_{[09-19]}(c))^{1/n}$ , where  $c$  is a company in a specific (*stress type, percentile*) bin, and  $n$  is the number of the companies in such a bin. Positive stress companies enjoyed the highest stock growth with an average value across all percentiles of  $GM(\text{stock growth}_{[09-19]}) = 5.07$  (Figure 3b), while the average stock growth across the other three types of companies was noticeably lower ( $GM(\text{stock growth}_{[09-19]}) = 3.70$ ), with passive stress companies exhibiting the lowest growth ( $GM(\text{stock growth}_{[09-19]}) = 3.42$ ). To ease the interpretation of such values, consider the example of Equinix, a digital infrastructure company headquartered in California, which our approach labeled to be a “positive stress” company. Its stock price traded at 61\$ in 2009 and its stock price climbed over 695% (i.e., its  $GM(\text{stock growth}_{[09-19]})$  was 7.95), trading at 485\$ ten years later.

## Discussion

**Limitations.** This work has five main limitations. The first concerns the inability of studying whether stress causes performance differences given the absence of cross-lag data that links performance to a stress-related company-wide indicator. Theoretically, we could run a lagged analysis as a linear regression where the dependent variable is the company’s growth at different time intervals. However, such an analysis is hard because of two main reasons: (a) no fine-grained temporal granularity for reviews is possible as reviews might be temporally misaligned since they could be posted after an employee leaves the company, and (b) many, mostly smaller, companies have joined the public reviewing site at later points in time, thus reviews will not cover all 12 years of analysis.

The second limitation is that the decreasing trend of stock growth may be dependent on the two main aspects: company ratings and industry sector. These two have little to do with the hypothesized relationship between stress and performance. We therefore repeated our analyses by considering a company’s overall website rating and its industry sector. As for ratings, we indeed found increasing stock growth with increasing review ratings; still, positive stress companies experienced the highest growth (Fig. S6 in *Supplementary Information*) compared to highly-rated companies. As for industry sectors, we showed that tech companies were over-represented in the positive stress set, and stock growth was partly driven by them (Fig. S7 in *Supplementary Information*). However, by separating companies by industry sector, we still observed that positive stress companies grew more than the other three types (Fig. S8 in *Supplementary Information*).

The third limitation concerns data biases related to temporal granularity and geographic representativeness. Upon new available data, future studies could study workplace stress outside US, allowing for cross-cultural comparisons.

The fourth limitation has to do with nuances when rating a company (e.g., being satisfied with the use of the overall company rating and not its composing dimensions). While on Glassdoor there are several rating fields available, only the overall rating field was mandatory and hence provided sufficient coverage for our analysis.

The fifth limitation is that the deep-learning model used to detect stress mentions in posts is not always accurate. Our medical entity extraction model has two main limitations. First, the model’s strict/relaxed accuracy is .71/.85, and, even though it outperformed competitive baselines by a large margin, it still is not perfect. To partly address this issue, our method limits itself to textual mentions pertaining stress at work. Second, entity extraction models such as ours are not always able to tell apart personal from figurative health mentions (e.g., ‘I felt pain’ vs. ‘He was such a pain to work with’). This is still an active area of research. Yet our model is relying on a large transformer model (e.g., contextual embeddings RoBERTa), and, as such, it is less likely to make such errors than a simpler, keyword-matching technique. Future studies could use some of the newly published social media datasets<sup>38</sup> to further train our model to distinguish between different *types of health mentions*.

**Implications.** To place our work in the broader context of the literature, we point out three main findings. Our first was that *company reviews contain linguistic markers of four stress types*. Previous work found that stress of social media users can be detected by analyzing their textual content, both on Twitter and Reddit<sup>21</sup>. Another study by Saha and De Choudhury found that high levels of stress markers were present in the use of language in Reddit comments posted by university students who experienced gun violence events at their campuses. This work showed that such linguistic changes are sufficiently subtle to reflect four *different types of stress*, that is, low, passive, positive, and negative stress. Our second finding was that *stress over the years tallied with large-scale exogenous events*. In particular, negative stress was the most prevalent among the four stress types in recession years (both great and mini recessions). This finding is in line with the literature linking economic downturn with stress and mental health issues caused by job instability<sup>39</sup>, and speaks to the presence of linguistic markers reflecting negative stress associated with country-level economic performance. Our third finding was that *company stock growth is associated with positive stress*. This is a new finding, not least because of lack of data in the past. While stock growth conflates endogenous factors (e.g., productivity) with exogenous ones (e.g., financial cycles), we found that positive stress companies enjoyed significantly stronger stock growth.

However, more work is needed to understand how to change a company’s culture into one in which stressors could be used for one’s growth and self-development. Given the recent wave of Great Resignation (i.e., the elevated rate at which U.S. workers have quit their jobs<sup>40</sup>), questions relating to corporate culture<sup>41</sup> and ways of retaining top talent are of utmost importance. A recent study from Mercer, an American asset management firm, found that elevated levels of employee turnover are not due to lack of engagement at work but attributed to workplace culture and heightened stressors. Therefore, organizations need to take immediate actions by (re) assessing their workplace culture first and by then shifting it when deemed appropriate, through training that fosters psychological safety and cultivates one’s mindset towards positive stress. Traditionally, employee well-being has been tracked with tailored surveys. Automated analyses of the language used by employees on corporate social-networking tools might offer yet another way of tracking workplace stress, which is sufficiently granular to assess the impact of interventions in a company. Beyond the immediate use of these findings for individual



companies, several other stakeholders could benefit from our methodology including government officials. As the performance of the S & P 500 companies affects the broader U.S. economy, recommended workplace practices could be established at state- or national-level to improve work conditions.

### Data availability

We made our code and data available in a readily usable format on GitHub ([https://github.com/sanja7s/positive\\_stress\\_in\\_companies](https://github.com/sanja7s/positive_stress_in_companies)) to allow for reproducibility. For each company, we shared the following attributes: company name, #total reviews, #stress reviews, average rating, rating of work-life balance, rating of career prospects, rating of the company, rating of the culture, rating of the management, stress type, strength of association with the stress type, stock values/growth for: 2009, 2012, 2014, 2019, and industry sector.

Received: 18 October 2021; Accepted: 20 December 2022

Published online: 28 January 2023

### References

- Sugar, A. Stay cool under pressure – without appearing cold. *Harvard Business Review* (2020).
- Nerurkar, A., Bitton, A., Davis, R. B., Phillips, R. S. & Yeh, G. When physicians counsel about stress: Results of a national study. *JAMA Internal Med.* **173**, 76–77 (2013).
- Cartwright, S. & Cooper, C. L. *Managing Workplace Stress* (Sage, 1997).
- Pal, P. Battling the Physical Symptoms of Stress. *Harvard Business Review* (2016).
- Vaske, J. J. Advantages and disadvantages of internet surveys: Introduction to the special issue. *Hum. Dimens. Wildl.* **16**, 149–153 (2011).
- Duda, M. D. & Nobile, J. L. The fallacy of online surveys: No data are better than bad data. *Hum. Dimens. Wildl.* **15**, 55–64 (2010).
- Gigliotti, L. M. Comparison of an internet versus mail survey: A case study. *Hum. Dimens. Wildl.* **16**, 55–62 (2011).
- Fricker, R. D. & Schonlau, M. Advantages and disadvantages of internet research surveys: Evidence from the literature. *Field Methods* **14**, 347–367 (2002).
- Selye, H. *The Stress of Life* (McGraw-Hill, 1956).
- Lazarus, R. S. *Toward Better Research on Stress and Coping*. American Psychological Association (2000).
- Colligan, T. W. & Higgins, E. M. Workplace stress: Etiology and consequences. *Workplace Behav. Health* **21**, 89–97 (2006).
- Kobasa, S. C. Stressful life events, personality, and health: An inquiry into hardiness. *Personality Soc. Psychol.* **37**, 1 (1979).
- Muse, L., Harris, S. & Feild, H. Has the inverted-U theory of stress and job performance had a fair test?. *J. Hum. Perform.* **16**, 349–364 (2003).
- Wallis, C., Mehrtens, R. & Thompson, D. Stress: Can we cope?. *Time* **121**, 48–54 (1983).
- Crum, A. J., Salovey, P. & Achor, S. Rethinking stress: The role of mindsets in determining the stress response. *Personality Soc. Psychol.* **104**, 716 (2013).
- Cohen, S., Kamarck, T. & Mermelstein, R. A global measure of perceived stress. *J. Health Soc. Behav.* **24**(4), 385–396 (1983).
- Hammen, C. Stress and depression. *Annu. Rev. Clin. Psychol.* **1**, 293–319 (2005).
- Wang, J. Work stress as a risk factor for major depressive episode (s). *Psychol. Med.* **35**, 865–871 (2005).
- Park, C. L. & Helgeson, V. S. Introduction to the special section: growth following highly stressful life events-current status and future directions. *J. Consult. Clin. Psychol.* **74**, 791 (2006).
- Tedeschi, R. G. & Calhoun, L. G. Posttraumatic growth: Conceptual foundations and empirical evidence. *Psychol. Inq.* **15**, 1–18 (2004).
- Guntuku, S. C., Buffone, A., Jaidka, K., Eichstaedt, J. C. & Ungar, L. H. Understanding and measuring psychological stress using social media. In *Proceedings of the International AAAI Conference on Web and Social Media* **13**, 214–225 (2019).
- Pennebaker, J. W., Francis, M. E. & Booth, R. J. Linguistic inquiry and word count: Liwc 2001. *Mahway Lawrence Erlbaum Assoc.* **71**, 2001 (2001).
- Saha, K. & De Choudhury, M. Modeling stress with social media around incidents of gun violence on college campuses. *Proc. ACM Hum. Comput. Interact.* **1**, 1–27 (2017).
- Das Swain, V. et al. Modeling organizational culture with workplace experiences shared on glassdoor. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI)*, 1–15 (2020).
- U.S. Bureau of Labor Statistics. (Seas) Unemployment Rate. <https://www.bls.gov> (2022).
- Yahoo Finance portal. Historical Data. <https://finance.yahoo.com> (2022).
- Šćepanović, S., Martín-López, E., Quercia, D. & Baykaner, K. Extracting medical entities from social media. In *Proceedings of the ACM Conference on Health, Inference, and Learning (CHIL)*, 170–181 (2020).
- Aronson, A. R. & Lang, F.-M. An overview of metamap: Historical perspective and recent advances. *J. Am. Med. Inform. Assoc.* **17**, 229–236 (2010).
- Tutubalina, E., Miftahutdinov, Z., Nikolenko, S. & Malykh, V. Medical concept normalization in social media posts with recurrent neural networks. *J. Biomed. Inform.* **84**, 93–102 (2018).
- Leaman, R. & Lu, Z. Taggerone: Joint named entity recognition and normalization with semi-Markov models. *Bioinformatics* **32**, 2839–2846 (2016).
- Šćepanović, S., Aiello, L. M., Barrett, D. & Quercia, D. Epidemic dreams: Dreaming about health during the covid-19 pandemic. *R. Soc. Open Sci.* **9**, 211080 (2022).
- Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794 (2022).
- Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018).
- McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426 (2018).
- McInnes, L., Healy, J. & Astels, S. hdbscan: Hierarchical density based clustering. *J. Open Sour. Softw.* **2**, 205 (2017).
- Grusky, D. B., Western, B. & Wimer, C. The consequences of the great recession. *Great Recess.* 3–20 (Russell Sage Foundation, 2011).
- Brown, A. Gamestop shares surge 100% in a day, Reddit group rejoices. *Forbes*. <https://www.forbes.com/sites/abrambrown/2021/02/24/gamestop-shares-surge-100-in-a-day-reddit-group-rejoices> (2021).
- Naseem, U., Kim, J., Khushi, M. & Dunn, A. G. Identification of disease or symptom terms in reddit to improve health mention classification. *Proc. ACM Web Conf.* **2022**, 2573–2581 (2022).

39. Mehta, K. *et al.* Depression in the us population during the time periods surrounding the great recession. *J. Clin. Psychiatry* **76**, 4221 (2015).
40. How to manage the great resignation. *The Economist*. <https://www.economist.com/business/2021/11/27/how-to-manage-the-great-resignation> (2021).
41. Reading corporate culture from the outside. *The Economist*. <https://www.economist.com/business/2022/07/07/reading-corporate-culture-from-the-outside> (2022).

### Author contributions

S.S., M.C. and D.Q. wrote the main manuscript text. S.K. did the preliminary experimental setup. S.S. did the main experiments and produced the main results. All authors reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-26796-6>.

**Correspondence** and requests for materials should be addressed to D.Q.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023