Article:

Symitsi, E and Stamolampros, P orcid.org/0000-0001-8143-7244 (2021) Employee sentiment index: Predicting stock returns with online employee data. Expert Systems with Applications, 182. 115294. ISSN 0957-4174

https://doi.org/10.1016/j.eswa.2021.115294

Employee Sentiment Index: Predicting Stock Returns with Online Employee Data

Employee Sentiment Index: Predicting Stock Returns with Online Employee Data[1]

First Draft: 05/01/2019

Accepted Version: 09/01/2021

Efthymia Symitsi[2]
Leeds University Business School, University of Leeds, United Kingdom
Location: 2.13 Maurice Keyworth Building
Email: e.symitsi@leeds.ac.uk
Phone: +44(0)113 343 6854

Panagiotis Stamolampros
Leeds University Business School, University of Leeds, United Kingdom
Location: 1.14 Charles Thackrah
Email: p.stamolampros@leeds.ac.uk
Phone: +44(0)113 343 6850

[1]We thank Shima Amini, Panagiotis Andrikopoulos, Konstantinos Bozos, George Daskalakis, Darren Duxbury, Phil Holmes Lazaros Symeonidis, Vasileios Kallinterakis, Barbara Summers, participants at the Behavioural Finance Working Group Conference (Queen Mary University of London), and anonymous reviewers for helpful comments and feedback on this paper.

[2]Corresponding author.

Abstract

We propose an aggregate measure of employee sentiment based on millions of employee online reviews and we test whether big employee data embedded in expert financial models can improve stock return predictability. In line with behavioral finance theory, our results document that the collective employee sentiment is a strong predictor of stock market returns with lower future returns following high employee sentiment. This predictive power is more pronounced when the employee sentiment index is constructed using the expectations of employees about the near-term business outlook of their employer. Our market-wide sentiment measure has superior performance compared to existing proxies of investor sentiment and commonly-studied macroeconomic variables. The forward-looking property of this data is also evident in predicting industry returns or portfolio returns sorted on characteristics, such as size, age, risk, profitability, dividend payout, tangibility, financial constraints and growth opportunities. Importantly, market-wide employee sentiment has relative power in predicting future asset returns after controlling for firm-level employee sentiment. The predictive power of aggregate employee online data is explained by investors' biased beliefs about expected cash flows and volatility. These results indicate that financial models can be enriched with sentiment factors derived from various big data sources and stakeholders, providing insights into mispriced assets and assisting investment decisions.

Keywords: Employee sentiment, online big data, voluntary information disclosure, business outlook predictions, return predictability, expert financial models

JEL classifications: G12, G17, G40

1. Introduction

The impact of market-wide sentiment in investment decisions has been widely advocated in behavioral finance. Noise-trader models propose that sentiment-driven investors temporarily drive prices away from their fair values (Shleifer and Summers, 1990; De Long et al., 1990; Barberis et al., 1998; Campbell and Kyle, 1993; Daniel et al., 1998, 2001; Grossman and Stiglitz, 1980; Hong and Stein, 1999; Kogan et al., 2006). In their seminal papers, Baker and Wurgler (2006; 2007) construct an aggregate sentiment index and demonstrate cross-sectional and market return predictability stressing that *"descriptively accurate models of prices and expected returns need to incorporate a prominent role for investor sentiment"* (Baker and Wurgler, 2006, p.1677). Although several sentiment proxies have been proposed (Lemmon and Portniaguina, 2006; Huang, Jiang, Tu and Zhou, 2015; Da et al., 2015; Gao et al., 2019; Tetlock, 2007; Oliveira et al., 2017), with sentiment being unobservable, the need for accurate measures calls for the examination of new sources of sentiment that complement existing metrics and improve the power of financial models. Market sentiment constructed using data from surveys, news, social media, financial reports, web searches have focused on market participants, such as investors, managers and consumers. However, data limitations have prevented so far the measurement of the market sentiment of employees, an important group of stakeholders.

The advent of online platforms, such as *Glassdoor*, allows employee voices to be heard offering new data in large scales. Recent studies investigate firm-level employee sentiment in stock returns, though, an aggregate measure of employee sentiment and its effect on predicting asset prices is yet to be studied. There exists an increasing consensus in the literature that there are delineated effects from market-wide and firm-level sentiment (Mahmoudi et al., 2020; Aboody et al., 2018; Ahmad et al., 2016; Tetlock et al., 2008). While Kaniel et al. (2008) find low cross-sectional correlation in investors' sentiment, other studies including Kumar and Lee (2006) and Barber et al. (2008) report that, on the aggregate, retail investor trading decisions are correlated. In the same spirit, Kothari et al. (2006) and Hirshleifer et al. (2009) report a positive association when testing firm-level variables, while they document a negative relationship for their aggregated counterpart. Such findings are also in line with Engelberg et al. (2019) who document that most cross-sectional predictors are not necessarily good time-series predictors.

In this paper, we propose and test an aggregate employee sentiment measure in predicting stock market returns and returns in the cross-section in a comprehensive empirical analysis. To this end, we aggregate opinions of employees disclosed voluntarily and anonymously on *Glassdoor*. We form, hence, three variant measures of employee sentiment. The first measure uses the overall rating of employers. The second measure aggregates both structured (numerical ratings in various job aspects) and unstructured (free text with the positive and negative job aspects) data. The third measure is an expected employee sentiment index, that utilizes a forward-looking feature of *Glassdoor* which allows employees to rate the 6-month ahead business outlook of their employer. The employee sentiment (ES) measures are constructed at monthly frequency aggregating opinions from millions of employees from thousands of employers (both public and private companies) operating in all sectors of the US economy rendering a well-representative market-wide measure of employee sentiment.

With employees being an important group of stakeholders, market-wide employee sentiment from staff across the labor market is worth investigating for the following reasons. First, studies have provided empirical evidence that employees' information is valuable and, to some extent, their information set differs from that of top managers. For instance, using Employee Stock Purchase Plans (ESPPs), Babenko and Sen (2015) report that nonexecutive employee purchases altogether predict their employer's stock returns. An increasing number of studies also relates online employee data with firm-level profitability and stock performance (Huang, Li, Meschke and Guthrie, 2015; Symitsi et al., 2018; Green et al., 2019; Hales et al., 2018; Stamolampros et al., 2019; Symitsi et al., 2021). Thus, employees have information advantage over outside investors, incremental information over their top managers and, similarly to external investors and managers, are not immune to behavioral biases. This implies that their opinions for their employer are of particular interest and will reflect, in addition to any private information they have, their optimism or pessimism about the firm.

Second, various studies have demonstrated that firm-specific information is reflected in non-traditional news outlets and investors increasingly resort to them to make inferences for the expected cash flows of firms and, make accordingly stock investment decisions (Miller and Skinner, 2015). Websites, social media and review platforms (Campbell et al., 2017; Tang, 2018), such as online bulletin boards (Antweiler and Frank, 2004), *Seeking Alpha* (Chen et al., 2014), *Estimize* (Jame et al., 2016), *Facebook* (Siganos et al., 2014), *Twitter* (Bartov et al., 2017; Oliveira et al., 2017), *Amazon* (Huang, 2018) and *Glassdoor* (Green et al., 2019), consist examples of non-traditional information sources for investors. Shiller and Pound (1989) explicitly refer to the particular attention of investors to opinions of employees in target firms and stocks. It is reasonable, thus, to expect that information revealed by employees on *Glassdoor* will influence those investors' trading decisions who seek to exploit this relatively new source of information in their attempt to gain competitive trading advantage. This is in line with media articles reporting that practitioners and investment professionals, such as hedge funds, private equity, and venture capitals, are increasingly consulting information from *Glassdoor* as part of their investment due diligence process[3] and for identifying acquisition targets[4]. This information is also likely to reach investors, who do not observe directly employee opinions shared online, through traditional word-of-mouth channels.

Third, employees typically invest in their employer. According to the US National Center for Employee Ownership, approximately 30 million Americans in 2014 held their employer stocks through Employee Stock Ownership Plans (ESOPs), ESPPs, 401 (K) plans, and stock options, with the median financial stake being 23% of their salary.[5] The participation in these schemes and, in turn, the proportion of employees' wealth allocated in employers' stocks depend on the expectations for their employer prospects. Therefore, the business outlook predictions capture the sentiment of those employee-investors that post their opinions online and should predict their trading behavior. We argue, though, that on average, the aggregate opinions may also reflect the sentiment and trading patterns of those employee-investors who do not share their opinions as their decision to invest into their employers' stock could be similar to that of their colleagues. This is supported by a strong positive correlation between the investment decisions of individuals with those of their co-workers found by Hvide and Östberg (2015) while investigating social interaction effects at work. Finally, employees' opinions and predictions, in addition to firm-specific knowledge, can be determined by the overall economic conditions. This suggests that even if employees do not possess superior information, or they do not invest in stock markets, or even if investors do not observe, or are not influenced by this information, still

[3]See "Hedge funds and private equity tap *Glassdoor* for investment tips" by Madison Marriage, *Financial Times*, January 21, 2017. Accessed 21/01/2017.

[4]See "Salesforce uses *Glassdoor* like Yelp for billion-dollar buyout decisions" by Joon Ian Wong, *Quartz*, October 19, 2016 (https://qz.com/813671/salesforce-crm-treats-glassdoor-like-yelp-for-billion-dollar-buyout-decisions/). Accessed 21/10/2016.

[5]https://www.nceo.org/articles/widespread-employee-ownership-us. Accessed 20/05/2019.

employee opinions will measure market-wide sentiment shaped by the general market conditions.

The abovementioned argumentation treats employees' opinions as a constellation of knowledge, employer expertise, and sentiment. While these can be informative of firm-specific returns, is yet to determine whether the aggregated employee sentiment could work as a mispricing factor. With sentiment being a societal rather than an individual process that spreads across the market, there exists potential to affect individuals' consumption and investment decisions. This means that whereas at the firm level such "private" information and sentiment cues could decrease mispricing of asset returns, at an aggregate level they may increase systematic mispricing and correlated judgement errors that propagate across assets (Daniel et al., 2001; Peng and Xiong, 2006; Daniel and Hirshleifer, 2015). Thus, aggregating firm-level employee expectations will remove the idiosyncratic elements and maintain the average employee sentiment across the market. Mahmoudi et al. (2020) present a theoretical model that accommodates both firm-level and market-level sentiment positing that the former could better capture heterogeneous sentiment across firms, while the latter could better capture correlated judgement errors. Altogether, based on the theoretical and empirical underpinnings of Daniel et al. (2001), Mahmoudi et al. (2020) and Engelberg et al. (2019), we argue that the aggregate employee sentiment may work in a distinct way than firm-level employee sentiment, justifying the context of this analysis.

Motivated by the theoretical predictions of noise-trader sentiment models (De Long et al., 1990), we test the value of aggregated employee sentiment embedded in financial models in predicting stock market returns. We subsequently perform multivariate regressions controlling for a large set of economic variables and examining whether market-wide employee sentiment complements existing well-established market sentiment proxies. We also test if an aggregate employee sentiment index captures sentiment effects that manifest in the cross-section, considering a large set of characteristic- and industry-sorted portfolios. Importantly, to the best of our knowledge, we are the first to explore whether aggregated employee sentiment predicts asset returns beyond firm-level employee sentiment. Last, we investigate the economic drivers that explain the predictability of the *ES*.

Our findings suggest that collectively business outlook forecasts from employees capture investor sentiment and have predictive power in stock markets. There exists a strong negative effect on 1-, 3-, 6-, 9- and 12-month ahead excess returns that remains significant after controlling for economic and sentiment variables and maximizes for the 9-month period. We also report a positive, though, statistically insignificant association of our index with contemporaneous excess returns. Not surprisingly, sentiment betas vary significantly across industries, characteristics and horizons. In line with Baker and Wurgler (2007), the findings suggest that employee sentiment changes capture systematically a stronger effect on portfolios at the top decile with the most volatile stocks compared to the bottom decile with the least risky stocks. We document also a strong market-wide employee sentiment ability in predicting asset returns controlling for firm-level employee sentiment and market returns. It appears that cash flow is the channel through which employee sentiment alters expectations for assets. Moreover, our results uncover a positive association with volatility, consistent with behavioral finance predictions.

Our paper contributes to several streams in the literature. We add to the literature on how market sentiment impacts asset prices with a growing number of studies in stock return predictability (e.g., Baker and Wurgler, 2006, 2007; Baker et al., 2012; Hribar and McInnis, 2012; Arif and Lee, 2014; Huang, Jiang, Tu and Zhou, 2015; Giannini et al., 2018; Nguyen et al., 2015), and widely reported market anomalies (Stambaugh et al., 2012; Baker and Wurgler, 2006). Existing popular market-wide sentiment proxies use data for managers, investors, consumers or online users (Baker and Wurgler, 2006; Lemmon and Portniaguina, 2006; Huang, Jiang, Tu and Zhou, 2015; Da et al., 2015; Jiang et al., 2018; Oliveira et al., 2017). In this paper, for the first time, we explore the market-wide sentiment of employees, an unexplored group of stakeholders, for predicting stock returns at the market and the cross-section. Moreover, our employee sentiment measure aggregates non-financial online employee data rather than historical market data, financial reports, or corporate news investigating the incremental value of new data sources in generating advanced financial models with superior predictive performance. Importantly, no prior study examines empirically delineated effects between market-wide and firm-level sentiment in line with Mahmoudi et al. (2020).

This research is also related to the literature that studies the importance of social media in capital markets (Bartov et al., 2017; Tang, 2018; Chen et al., 2014; Campbell et al., 2017; Oliveira et al., 2017; Giannini et al., 2018; Nguyen et al., 2015; Weng et al., 2018). We provide evidence that employee online review platforms have information that is important not only for measuring accurately future prospects of individual firms, but also for the entire market.

A notable difference of aggregating data from *Glassdoor* platform is that reviews derived from all-level staff are less prone to biases compared to eponymous posts of top executives (Larcker and Zakolyukina, 2012) and involve better "insider" insights for the company perspectives than individuals not employed by the company.

We also expand the nascent literature that explores employee online reviews. This rather untapped source of information has received increasing attention mainly as an alternative and higher-frequency measure of employee satisfaction to predict firm and stock performance (Huang, Li, Meschke and Guthrie, 2015; Symitsi et al., 2018; Green et al., 2019; Huang et al., 2020; Symitsi et al., 2021). Only a limited number of studies employ business outlook expectations (Hales et al., 2018; Sheng, 2019; Huang et al., 2020). These studies, though, perform firm-level analyses. In this paper, we take a bottom-up approach investigating market-wide employee sentiment rather than firm-level employee sentiment and its ability in predicting both aggregate and firm-level returns.

Finally, we extend the literature on the information content of rank-and-file employee opinions and behaviors (Huddart and Lang, 2003; Babenko and Sen, 2015; Benartzi, 2001; Cohen, 2008). The findings of Babenko and Sen (2015) and Huddart and Lang (2003) corroborate the idea that altogether lower-level employees may have as precise information as the senior employees and face fewer constraints than the top senior staff. In contrast, other studies, including Benartzi (2001) and Cohen (2008), cast doubt on the propensity that lower-level employees possess private information to achieve superior returns. Our findings suggest that collectively employee business outlook forecasts, independently of their positions in the company hierarchy, can capture investor sentiment and have predictive power on stock returns.

The rest of the study is organized as follows: Section 2 discusses the related literature. Section 3 describes the data and the methodology. Section 4 presents the empirical results of market-wide employee sentiment predicting stock market, portfolio and asset returns. This part also examines the economic channels through which this predictability manifests. Section 5 discusses the contributions and the practical implications and Section 6 concludes.

2. Related Literature

The impact of investor sentiment on asset prices has long been debated among finance scholars and investment professionals (e.g., see Keynes, 1936; Nelson, 1902). According to traditional theory, assets and prices reach equilibrium after competition among rational investors (Gomes et al., 2003). Even if prices diverge from their fundamental values due to overly optimistic or pessimistic trading decisions of a subset of investors, arbitrageurs ensure that such deviations will only be short-lived. However, De Long et al. (1990) among others, present theoretically that sentiment-driven mispricing is possible with limits to arbitrage (see also Barberis et al., 1998; Campbell and Kyle, 1993; Daniel et al., 1998, 2001; Grossman and Stiglitz, 1980; Hong and Stein, 1999; Kogan et al., 2006). Empirical evidence supporting the predictions of noise-trader models is provided in Baker and Wurgler (2006) moving the academic research from questions beyond whether market sentiment affects assets to how it should be measured. In the absence of a perfect sentiment measure, Baker and Wurgler (2006) form a market-based sentiment measure combining several investor sentiment proxies.

In the same vein, several studies focus on developing appropriate measures to capture sentiment impact on asset prices both at the cross-section and market level. Based on psychological evidence that relates mood to judgement and decision making (e.g., Johnson and Tversky, 1983; Wright and Bower, 1992), sentiment is approximated on the basis of exogenous non-economic factors that affect investors' mood, such as weather conditions (Saunders, 1993; Hirshleifer and Shumway, 2003), temperature (Cao and Wei, 2005), daylight saving time (Kamstra et al., 2000), the Seasonal Affecting Disorder (SAD) (Kamstra et al., 2003), moon phases (Yuan et al., 2006), sport events (Edmans et al., 2007) and aviation disasters (Kaplanski and Levy, 2010). The findings support that changes in investor sentiment significantly impact asset prices.

Other proxies that have been proposed in the literature are based on survey, market, online search volume, media, and text data. For example, Lemmon and Portniaguina (2006) employ the Conference Board Consumer Confidence Index ($CBCI$) and the University of Michigan Consumer Sentiment Index ($UMCI$), and find that they predict future market and industry returns. Huang, Jiang, Tu and Zhou (2015) form the aligned investor sentiment index using an alternative methodology to Baker and Wurgler (2006) which increases the performance of the index in predicting stock market and cross-sectional returns both in-sample and out-of-sample.

Da et al. (2015) synthesize the *FEARS* index using the *Google* search volume for terms, such as "recession" and "unemployment". Their sentiment index, consistent with behavioural theories, predict short-term return reversals, volatility spikes, and equity fund flows to bonds. Gao et al. (2019) rely on aggregate household *Google* searches for economic and finance terms or other keywords that trigger sentiment, such as weather, disasters, and holidays, and investigate cross-country stock market predictability, concluding that sentiment prevails in international stock markets. Tetlock (2007) proposes a pessimism index, computed from the sentiment in a highly appreciated among investors *Wall Street Journal* column. The results show that negative market sentiment predicts low future returns at short horizons that reverse at longer horizons. In agreement with investor sentiment theories, extremely high or low pessimism is associated with high market trading volume. Jiang et al. (2018) build proxies of manager sentiment deriving the textual tone from statements and conference call transcripts. Their empirical analysis documents strong negative predictive power of market and cross-sectional returns. Oliveira et al. (2017) focus on extracting aggregate sentiment from social media and microblogs. Then, this is combined to several well-established sentiment measures including the *UMCI*, the Sentix, the American Association of Individual Investors (AAII) and Investors Intelligence (II) indexes to form a market-wide sentiment measure. Their index is validated against investor survey-based sentiment measures and is found to have incremental information in forecasting returns, volatility and trading volumes across indexes and portfolios.

Although a plethora of market-wide sentiment proxies has been employed in the literature across market participants, aggregating the sentiment of investors, managers, consumers, or on-line users, no study so far extracts the market-wide sentiment of employees. Answering to calls for accurate sentiment proxies, this paper fills this gap by constructing measures of market-wide employee sentiment using structured, unstructured and forward-looking information taken from an online employee platform. The appealing properties of *ES* makes it a good sentiment proxy over market sentiment indexes from alternative sources. Investor sentiment measures based on market data (e.g., Lee et al., 1991; Baker and Wurgler, 2006; Ben-Rephael et al., 2012) have the disadvantage of relying on historical information. As explained in Da et al. (2015, p.2), they *"are the equilibrium outcome of many economic forces other than investor sentiment"*. Alternative investor sentiment proxies based on search, media content and text data (Da et al., 2015;

Tetlock, 2007; Jiang et al., 2018) are particularly sensitive to the choice of terms considered to reveal sentiment, or are computationally demanding, requiring advanced text analytic techniques. Moreover, indexes that employ financial reports or transcript calls (Jiang et al., 2018) have the additional disadvantage that use information derived from eponymous announcements from managers. However, senior managers are subject to pressures due to their agency relationships, and their narratives are commonly designed to transfer particular sentiment cues to market participants and other key stakeholders (Larcker and Zakolyukina, 2012), even if they are not trying to purposely fool or manipulate investors. In the same spirit, Huddart and Lang (2003) show that option exercises by the most senior employees do not provide superior information than those by employees at relatively junior job roles, suggesting that it is not the information value between senior and junior staff that differs, but the fewer constraints the latter group faces. On the contrary, our proposed sentiment measures use structured and unstructured data that reflect the sentiment of employees for existing firm practices. Importantly, we estimate the expected market-wide employee sentiment aggregating employee expectations for the look-ahead prospects of their employers. Since employees across all levels of hierarchy disclose this information anonymously, therefore, their opinions and narratives are not subject to such pressures. Moreover, these opinions are explicitly stated by employees rather than inferred by textual sentiment analysis that could induce noise.

A number of studies use firm-level employee ratings for their employers in predicting asset returns. Edmans (2011) and Edmans et al. (2014) link employee sentiment to stock returns in the US and around the world using the "Best Places to Work" lists. However, online access to myriads employee reviews through employee social platforms, transformed the research landscape. Symitsi et al. (2018) use *Glassdoor* employee reviews to sort portfolios with the overall employee rating. Their findings suggest that employees possess valuable information for fundamentals that is not fully priced in stock markets. Green et al. (2019) also find that changes in employee satisfaction are associated with changes in firm profitability, sales growth and earning announcement surprises, as well as stock returns. Disaggregating the information content of different features from *Glassdoor* platform, they find that employees' beliefs of senior management and career opportunities are more informative than other work aspects. Symitsi et al. (2021) document that combining employee information from several job aspects, both structured and unstructured, offers significant benefits in firm profitability models. Sheng (2019), Huang et al. (2020) and Hales et al. (2018) use business outlook expectations and document a positive association with future asset returns and performance. While there is evidence which suggests that firm-level employee sentiment reveals information for fundamentals, there is a lack of understanding on how market-wide employee sentiment affects stock markets.

We fundamentally differ from prior studies in that we estimate market-wide employee sentiment, averaging the employee sentiment across employers in the US market. By aggregating firm-level employee sentiment the idiosyncratic components are suppressed leading to a common sentiment factor across employees. Mahmoudi et al. (2020) present a theoretical model that differentiates firm-specific from market-wide sentiment positing that they capture different elements of sentiment. Under this perspective, market-wide sentiment in an asset pricing model would reflect the common beliefs of investors and affect their decisions across a portfolios of assets. We investigate, thus, market-wide employee sentiment as a factor of mispricing both at the aggregate and cross-section level controlling also for firm-specific employee sentiment.

3. Data and Methodology

3.1. The Employee Sentiment Index

The information used to estimate proxies of employee sentiment are taken from *Glassdoor*.[6] *Glassdoor*'s platform, in addition to overall employer ratings (scales from 1 to 5), allows current and former employees to anonymously review several work aspects.[7] These aspects include career opportunities, compensation and benefits, senior leadership, work/life balance, and culture and values in structured ordinal scales. Employees can also explicitly state whether they approve the CEO of the company ("Approve", "Disapprove", "No Opinion"). Importantly, since June 2012 employees are opted to provide a personal view on the 6–month ahead outlook of their employer company varying between "better", "same" and "worse". Finally, employees accompany their ratings for the predetermined aspects with free narratives elaborating further on the advantages (Pros) and shortcomings (Cons) of working at a particular company. In line with Symitsi et al. (2018), we consider the assessments of employees who were employed at

[6] We would like to thank *Glassdoor* for providing the data for our research.

[7] Access to the platform from new users is granted after they complete a review. To safeguard the content and the quality of each employee rating from manipulation and fraud, *Glassdoor* follows a particular process to verify and check accounts and reviews employing both algorithm-based procedures and human inspections.

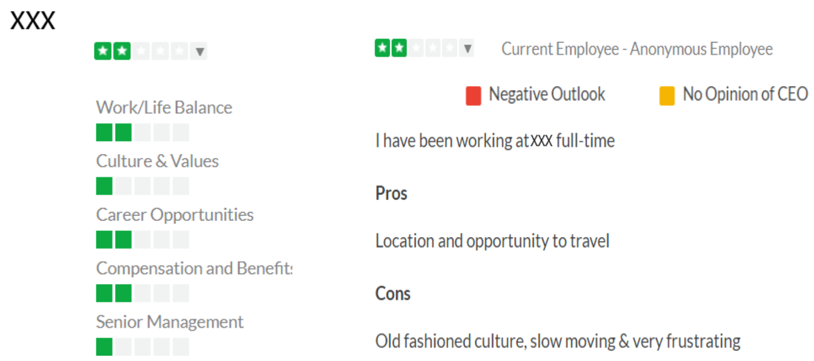

Figure 1. An example of an employee review on *Glassdoor* platform.

the company the time of the review post in order to ensure that our analysis is not biased by disgruntled former employees. The initial dataset contains a total of 2,778,343 online employee reviews from 225,748 private and public US employers spanning the period from June 2012 to July 2018. Figure 1 shows an example from an employee review for employer *XXX*.

We consider three variants of employee sentiment. The first index, ES_O, uses only the overall rating for each employer. The second employee sentiment, ES_{SU}, aggregates both structured and unstructured data. Finally, we build an expected sentiment index, ES_E, based only on the forward-looking opinions for the business outlook of companies. We follow a three-step process: In the first step, for each review, we re-assign all the ordinal scales to range between [-1,1]. Negative (positive) values reflect negative (positive) sentiment. For example, we assign -1, 0 and 1 values for "worse", "same" and "better" *Business Outlook*, respectively. In order to treat unstructured data, we extract the textual sentiment based on the narrative length in the Pros and Cons sections, estimated for each review as:

$$Text_{Sent} = \frac{(\text{Number of words in Pros} - \text{Number of words in Cons})}{(\text{Number of words in Pros} + \text{Number of words in Cons})}, \tag{1}$$

which ranges between [-1,1]. The overall review sentiment for each review for ES_{SU} is estimated by averaging the sentiment across all the criteria, S=(*Overall, CareerOpps, CompBenefits, SeniorLeadership, CultureValues, CEOApproval, BusinessOutlook,* $Text_{Sent}$).

In the second step, for each month t we compute the employee sentiment, $Firm_{Sent,it}$, for every employer i, as the sum of the sentiment values in all the respective reviews used scaled by the total number of firm reviews:

$$Firm_{Sent,it} = \sum Review_{Sent,it}/N_{it}, \tag{2}$$

N_{it} is the total number of reviews for the company i during month t. The $Review_{Sent,it}$, is the *Overall* rating, the sum of all the criteria in S defined above, and the *BusinessOutlook* for the ES_O, ES_{SU} and the ES_E, respectively.

In the third step, we obtain the employee sentiment time series, ES_t, by aggregating every month t the employee sentiment across firms:

$$ES_t = 100 \times \left(\sum Firm_{Sent,it}/M_t \right), \tag{3}$$

where M_t corresponds to the total number of firms with available review data in month t. We normalize the index to have a value of 100 in June 2012.

The *ES* exhibits several distinct features. First, biases resulting from firms with more frequent reviews are prevented. Second, it allows for equal representation of all companies irrespective of their size. Third, the index includes both private as well as public employers providing a quite diverse information source with approximately 32% of the sample consisting of reviews for listed firms.[8] Forth, the fact that the index aggregates information from a large cross-section of firms smooths out individual biases from employees who do not offer reliable evaluations of their employer. In particular, every month an average of 30,500 reviews from on average 13,860 firms are available through *Glassdoor* during the sample period. Moreover, the *ES* consists of a well-balanced set of reviews arriving indiscriminately from all industries, making it a good proxy for the aggregate market.

3.2. Other Data

3.2.1. Additional Sentiment Proxies

There is a possibility that the information content of our employee sentiment measures in predicting stock markets is subsumed by other sentiment proxies used in the literature. To investigate this, we consider the most well-established sentiment proxies in our analysis. We

[8] We examined in our analysis, an index derived from an alternative construction methodology that weighs firms with the volume of the reviews. The results remain qualitatively unchanged.

employ the University of Michigan Consumer Sentiment Index (*UMCI*) which considers the personal financial conditions, and the opinions of the future short- and long-term market and business conditions through a survey of 500 US households. We also control for the Conference Board Consumer Confidence Index (*CBCI*) which compiles the monthly consumer sentiment through a survey mailed to 5,000 households.[9] For both indexes, we take the expectation components which are based only on future-looking questions and have been shown to have greater forecasting power than the total components that include also questions related to the present conditions of participants (Bram et al., 1998; Ludvigson, 2004; Lemmon and Portniaguina, 2006).[10] The third sentiment proxy we use is the Purchasing Managers' Index (*PMI*), which is a business confidence index derived from a survey of up to 400 supply chain managers in the manufacturing sector. As described above, the number of opinions for the estimation of the *ES* every month are way beyond those used in the survey-based consumer and manager indexes. Furthermore, the *ES* has the advantage that is based on real-time information, while a publication delay exists for the *UMCI*, *CBCI*, and *PMI*.

Finally, we include the Baker and Wurgler (2006) investor sentiment index (*BWSI*). The updated data for this index is obtained from the webpage of Jeffrey Wurgler.[11] *BWSI* is constructed as the first principal component of the correlation matrix of stock market sentiment proxies, namely: the value-weighted dividend premium, the closed-end fund discount, the number of initial public offerings (IPOs) and their average first–day returns, and the equity share in new issues. Each of the proxy used for the construction of the *BWSI* index has been orthogonalized to a set of macroeconomic variables, which covers the growth rate in industrial production, the employment growth, the growth in real consumer durables, non-durables and services, and the NBER US recession dummy.

3.2.2. Equity Market Data and Economic Variables

In order to investigate the predictive ability of the employee sentiment index for stock market returns, we calculate monthly excess market returns, *ER*, by subtracting the one month US T-bill rate from the monthly S&P 500 return. Aggregate and firm-level stock market data are

[9] Approximately 3,500 responses are received each month regarding the perceived financial condition of the household and the general business conditions.

[10] We have also replicated the analysis using the total components of *UMCI* and *CBCI* and the results remain qualitatively unchanged.

[11] http://people.stern.nyu.edu/jwurgler/

taken from *CRSP*. [12]

Furthermore, we consider several control variables that are likely to capture variation in stock returns and changing economic conditions. More precisely, we employ the dividend-to-price ratio (DP), dividend yield (DY), earnings-to-price ratio (EP), book-to-market ratio (BM), stock market variance ($SVAR$), net equity expansion ($NTIS$), treasury bill rate (TBL), long-term return (LTR) and long-term yield (LTY) of US government bonds, default return spread between long-term corporate and government bond returns (DFR), default yield spread between the Moody's Baa and Aaa bonds (DFY), dividend payout ratio (DE), term spread between the long-term government bond and the treasury bill rate(TMS), inflation ($INFL$), industrial production growth (GIP), consumption of durables goods growth ($GCDG$) and consumption of non-durables goods growth ($GCNDG$), which are standard return predictors (e.g., see Welch and Goyal, 2008; Baker and Wurgler, 2006)[13] Data is obtained from the website of Amit Goyal[14] but the industrial production and consumption series which are sourced from the *Federal Reserve Bank of St. Louis*[15]. We use two lags for *INFL*, *GIP*, *GCDG*, *GCNDG*, and *GCS* to account for publication delay.

Table 1 presents summary statistics for the *ES* proxies, the other sentiment indexes and the economic variables described above. There is a positive correlation between the *ES* and the *UMCI*, *CBCI*, and *PMI* indexes.[16] The contemporaneous linear association is stronger for the two consumer indexes and lower for the *PMI*. Interestingly, the *ES* indexes have a negative association with the *BWSI*. Furthermore, the *ES* is positively correlated with several economic variables that reflect changing investment opportunities, such as the earnings to price ratio and the term spread. This is not surprising as it is expected that sentiment varies with the state of the economy (Garcia, 2013). In particular, Baker and Wurgler (2007) argue that contemporaneously the proxies of sentiment are subject to contamination from fundamentals. Though, changes in sentiment, used in our empirical analysis, are not significantly correlated with the tested variables.

[12] In the online Appendix, we assess the validity of our results when alternative equity market proxies and assets are used.

[13] For a detailed description of these variables, see the online Appendix.

[14] http://www.hec.unil.ch/agoyal/. The variables are updated up to December 2019. Accessed 30/10/2020.

[15] https://fred.stlouisfed.org/

[16] A correlation matrix is tabulated in the Online Appendix.

Table 1. Description of the Data

Variable Name	Description	Obs	Mean	Std. Dev.	Min	Max
Panel A: Employee Sentiment						
ES_O	Overall Rating	73	116.334	32.3555	64.5582	168.3681
ES_{SU}	Structured and Unstructured	73	120.7092	36.5747	65.7782	184.4506
ES_E	Business Outlook Predictions (Exp)	73	122.5334	32.69	74.4596	168.7642
Panel B: Other Data						
UMCI	University of Michigan Consumer Sentiment Index (Exp)	73	88.9233	8.0365	72.3	101.4
CBCI	Conference Board Consumer Confidence Index (Exp)	73	96.4082	19.8144	58.4	130
PMI	Purchasing Managers' Index	73	54.0055	3.3417	47.8	60.8
BWSI	Baker and Wurgler Investor Sentiment Index	73	-0.0454	0.119	-0.2671	0.1818
DP	Dividend price ratio	73	-3.9121	0.0526	-4.0481	-3.7888
DY	Dividend yield	73	-3.9022	0.0519	-4.0409	-3.7851
EP	Earning price ratio	73	-3.0348	0.1362	-3.2297	-2.7583
BM	Book to market	73	0.309	0.0356	0.2225	0.3516
SVAR	S&P 500 stock variance	73	0.0013	0.0011	0.0002	0.0058
NTIS	Net equity expansion	73	-0.0095	0.0137	-0.0325	0.0167
TBL	Treasury bill rate	73	0.0038	0.0052	0.0002	0.019
LTR	Long-term return	73	0.0018	0.0254	-0.0629	0.0709
LTY	Long-term yield	73	0.0271	0.0044	0.0175	0.0378
DFR	Default return spread	73	0.0018	0.0135	-0.0409	0.0426
DFY	Default yield spread	73	0.009	0.0024	0.0055	0.0149
DE	Dividend payout ratio	73	-0.8773	0.1363	-1.1126	-0.6704
TMS	Term spread	73	0.0267	0.0045	0.0173	0.0377
INFL	Inflation	73	0.0013	0.003	-0.0057	0.0082
GIP	Industrial production growth	73	0.0011	0.0049	-0.008	0.0151
GCDG	Consumer durables growth	73	0.0036	0.0083	-0.014	0.0259
GCNDG	Consumer non durables growth	73	0.0021	0.0071	-0.0188	0.0175
GCS	Consumer services growth	73	0.0037	0.0018	-0.002	0.0077
ER	S&P 500 Excess Stock Market Returns	73	0.0096	0.0275	-0.0647	0.0797

4. Empirical Analysis

4.1. Employee Sentiment and Stock Market Predictability

4.1.1. Predicting Stock Market Returns

To explore whether the aggregate employee outlook can predict excess stock market returns, we perform the following regression:

$$ER_{t+1:t+h} = \alpha + \beta \Delta ES_t + \varepsilon_{t+1:t+h}, \tag{4}$$

where $ER_{t+1:t+h}$ is the h-month ahead cumulative excess stock market return (i.e., the return on the S&P500 index in excess of the risk-free rate), where $h = 1, 3, 6, 9, 12$. $\varepsilon_{t+1:t+h}$ are the residuals. We also consider whether sentiment can predict contemporaneous returns. Δ indicates that the first differences of ES are considered (i.e., shocks to sentiment). Changes in sentiment have been associated with demand shocks that alter the levels of ownership between institutional and retail investors and, then, prices (DeVault et al., 2019). They also alleviate concerns for spurious results from OLS regression with highly correlated covariates and highly persistent and non-stationary predictors. Jiang et al. (2018) also comment on the spurious effects of highly persistent, correlated, and non-stationary covariates. They deal with this by using bootstrapped empirical values for the estimation of standard errors. We also employ empirically estimated robust standard errors from a bootstrapped procedure with 1,000 replications.[17] All the analyses are replicated with Newey-West heteroskedasticity and autocorrelation robust standard errors yielding similar results.

In line with behavioral predictions (e.g., De Long et al., 1990), we test the null hypothesis $H_0 : \beta = 0$ of no predictability, against the one-sided alternative $H_1 : \beta < 0$ that beta is negative and significant (Inoue and Kilian, 2005, recommend one-side tests to increase the power of the test). The explanatory variables are standardized to facilitate comparability across the variables as their scales differ. Therefore, the regression coefficients can be interpreted as the change in the aggregate stock returns from a one standard deviation change in the tested variables.

Table 2

Employee Sentiment and Excess Stock Market Return Predictability

	(1)	(2)	(3)	(4)	(5)	(6)
Horizon	t+0	t+1	t+3	t+6	t+9	t+12
Panel A: Employee Sentiment Index - Overall Rating						
ΔES_{OR}	−0.0005	−0.0016	−0.0067	−0.0086*	−0.0149**	−0.0132*
	(0.0040)	(0.0035)	(0.0055)	(0.0056)	(0.0080)	(0.0103)
Constant	0.0096***	0.0098***	0.0277***	0.0521***	0.0762***	0.0983***
	(0.0034)	(0.0032)	(0.0046)	(0.0058)	(0.0073)	(0.0087)
R^2	0.0003	0.0034	0.0280	0.0276	0.0534	0.0285
Panel B: Employee Sentiment Index - Structured and Unstructured Data						
ΔES_{SU}	0.0002	−0.0026	−0.0079*	−0.0084*	−0.0143**	−0.0131
	(0.0038)	(0.0034)	(0.0051)	(0.0053)	(0.0081)	(0.0102)
Constant	0.0096***	0.0098***	0.0277***	0.0521***	0.0762***	0.0983***
	(0.0034)	(0.0032)	(0.0046)	(0.0058)	(0.0073)	(0.0087)
R^2	0.0001	0.0091	0.0393	0.0263	0.0497	0.0282
Panel C: Employee Sentiment Index - Business Outlook Expectations						
ΔES_{E}	0.0023	−0.0065**	−0.0108**	−0.0107**	−0.0243***	−0.0211**
	(0.0031)	(0.0031)	(0.0047)	(0.0052)	(0.0084)	(0.0099)
	0.0096***	0.0098***	0.0277***	0.0521***	0.0762***	0.0983***
	(0.0034)	(0.0031)	(0.0046)	(0.0058)	(0.0070)	(0.0086)
R^2	0.0070	0.0557	0.0726	0.0422	0.1431	0.0730
Obs.	73	73	73	73	73	73

Note: This table presents the estimation results from regressing the h-month returns of the S&P 500 index in excess of the risk-free rate on the monthly differences of the three proxies of *ES* built with aggregate overall rating (Panel A), both structured and unstructured data (Panel B), and 6-month ahead business outlook (Panel C) opinions. Bootstrapped robust standard errors are reported in parentheses. Column 1 shows the results of the contemporaneous regression (i.e., same time t), while columns 2–6 display the results of h–day ahead cumulative excess return prediction (h=1,3,6,9,12). The sample period is from June 2012 to July 2018. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively, using a one-tailed test.

Table 2 presents the results from the predictive regressions with bootstrapped standard

[17] A large number of papers including Kothari and Shanken (1997), Nelson and Kim (1993), Baker and Stein (2004), Horowitz (2019) discuss that bootstrap standard errors control for bias in predictive regressions as outlined in (Stambaugh, 1999).

errors reported in parentheses for all the alternative employee sentiment proxies. While we report significant coefficients with the anticipated negative sign for all the ES proxies, we find that an employee sentiment measure that aggregates look-ahead business outlook opinions offer increasing forecasting benefits. In particular, the coefficients of the expected ES are strongly significant, with t-statistics ranging between -2.05 and -2.89. The estimated slope parameters vary between 0.65% (at the 1-month horizon) and 2.43% (at the 9-month horizon). In economic terms, this means that a one standard deviation shock to ES predicts a 2.43 (%) decrease in stock market returns over the subsequent 9 months. The explanatory power (R^2) of the 1-month ahead predictive regression is 5.57%, but becomes substantially higher at longer horizons, such as the 9- and the 12-month with 14.31% and 7.30%, respectively. This result is in line with evidence from the literature that stock returns are more predictable at longer horizons (Cochrane, 2011*a*). The coefficient of the contemporaneous regression in column 1 is not significant.

The preceding analysis clearly shows that the aggregate employee sentiment predicts a reversal in excess stock market returns. A question is whether the ES captures time-varying economic fundamentals or it contains unique information about stock market returns. If the latter is the case, then the documented relationship should be present after controlling for the effect of the various economic factors presented in Section 3.2.2. To do this, we estimate the following regression:

$$ER_{t+1:t+h} = \alpha + \beta \Delta ES_t + \gamma X_{i,t} + \varepsilon_{t+1:t+h}, \qquad (5)$$

where $X_{i,t}$ is the time series of economic variable i. The rest of the notation is as defined previously.

Panel A in Table 3 shows that the ES is negative and strongly significant in bivariate regressions that add the economic variable [*name in row*] to the regression of Equation (4). The column labeled ΔR^2 displays the change in the R^2 by adding the ES to a regression of the h-month ahead excess S&P 500 returns on a constant and the economic variable [*name in row*]. Henceforth we display only the output for the ES_E based on the business outlook given that it outperforms the current employee sentiment proxies (ES from now will refer to ES_E). Moreover, testing a forward-looking employee sentiment measure is more meaningful within a stock return predictability context. The results draw a similar picture to that of Table

2. Specifically, the incremental information content of the *ES* is substantially higher at longer horizons (i.e., at the 9- and 12-month horizons). For instance, focusing at the 9-month forecast horizon, the increase in the explanatory power by adding the expected *ES* is between 11.43% and 24.39%, which is sizeable. In sum, the results show that the information content of our sentiment index augments that contained in economic fundamentals.

The above analysis shows that the *ES* is a robust predictor of excess stock market returns. However, it would be interesting to investigate its predictive performance relative to other sentiment proxies. To this end, we add each one of the sentiment proxies we examined earlier to a bivariate predictive model providing further evidence for the information content of our measure when accounting for the common variation across the alternative sentiment proxies, defined as follows:

$$ER_{t+1:t+h} = \alpha + \beta \Delta ES_t + \gamma \Delta Sentiment_{j,t} + \varepsilon_{t+1:t+h}, \tag{6}$$

where $\Delta Sentiment_{j,t}$ is the first difference in the value of the sentiment proxy j, where j = {UMCI, CBCI, PMI, BWSI} in month t. The rest of the notation remains as defined earlier. Panel B in Table 3 shows that the slope estimate of the *ES* is negative and strongly significant with a similar magnitude to that found in the univariate regressions of the economic variables. Columns that report ΔR^2 further indicate that the *ES* has a substantial contribution to the prediction of future excess stock market returns. Taken together, the results suggest that aggregated employee outlook predictions carry additional information to that contained in other sentiment proxies and often fully subsume their predictive content with few exceptions, i.e., the *UMCI* at the 1-month horizon, the *BWSI* at the 3-month horizon and the *PMI* at the 12-month horizon.

One may reasonably argue that the two-variable regressions involving the *ES* and a control variable (i.e., either another sentiment index or an economic variable) are likely to suffer from known statistical biases (e.g., omitted variables bias). In this context, a regression with multiple factors would be more appropriate. To avoid multicollinearity issues which may arise due to highly correlated economic variables, we obtain the principal components (*PCs*) of the sentiment proxies and economic variables. We then estimate a multivariate regression which employs the estimated PCs as explanatory variables to predict the h-period ahead excess stock

Table 3

Employee Sentiment and Excess Stock Market Return Predictability Controlling for Economic and Sentiment Variables

Horizon	t+1			t+3			t+6			t+9			t+12		
	β	γ	ΔR^2	β	γ	ΔR^2	β	γ	ΔR^2	β	γ	ΔR^2	β	γ	ΔR^2
Panel A: Economic Variables															
DP	−0.0057**	0.0056*	0.0304	−0.0090**	0.0125***	0.0357	−0.0074**	0.0229***	0.0719	−0.0196***	0.0332***	0.1879	−0.0138**	0.0512***	0.1388
	(0.0029)	(0.0038)		(0.0040)	(0.0042)		(0.0044)	(0.0052)		(0.0067)	(0.0054)		(0.0063)	(0.0066)	
DY	−0.0062**	0.0032	0.0514	−0.0101**	0.0065	0.0618	−0.0087**	0.0191***	0.0641	−0.0209***	0.0345***	0.2061	−0.0162***	0.0490***	0.1212
	(0.0030)	(0.0032)		(0.0044)	(0.0051)		(0.0046)	(0.0059)		(0.0065)	(0.0058)		(0.0069)	(0.0077)	
EP	−0.0059**	0.0037*	0.0721	−0.0094**	0.0084**	0.0773	−0.0079*	0.0163***	0.0629	−0.0211***	0.0191***	0.1545	−0.0172**	0.0235***	0.1189
	(0.0031)	(0.0027)		(0.0044)	(0.0038)		(0.0050)	(0.0042)		(0.0082)	(0.0055)		(0.0095)	(0.0074)	
BM	−0.0063**	0.0027	0.0762	−0.0102***	0.0083*	0.0982	−0.0093**	0.0221***	0.1276	−0.0224***	0.0300***	0.2439	−0.0186**	0.0396***	0.2093
	(0.0030)	(0.0037)		(0.0043)	(0.0056)		(0.0049)	(0.0067)		(0.0077)	(0.0071)		(0.0084)	(0.0089)	
SVAR	−0.0065**	−0.0013	0.0713	−0.0108**	0.0022	0.0793	−0.0107**	0.0014	0.0558	−0.0242***	−0.0100	0.1769	−0.0211**	0.0008	0.0875
	(0.0031)	(0.0056)		(0.0047)	(0.0047)		(0.0053)	(0.0072)		(0.0086)	(0.0079)		(0.0100)	(0.0104)	
NTIS	−0.0065**	0.0007	0.0718	−0.0109**	0.0028	0.0918	−0.0107**	0.0010	0.0530	−0.0241***	−0.0046	0.1143	−0.0207**	−0.0118**	0.0035
	(0.0031)	(0.0030)		(0.0047)	(0.0037)		(0.0053)	(0.0043)		(0.0084)	(0.0056)		(0.0099)	(0.0067)	
TBL	−0.0064**	0.0016	0.0710	−0.0108**	−0.0003	0.0871	−0.0114**	−0.0199**	0.1138	−0.0255***	−0.0292***	0.2207	−0.0224***	−0.0344***	0.1309
	(0.0032)	(0.0047)		(0.0047)	(0.0084)		(0.0050)	(0.0114)		(0.0079)	(0.0103)		(0.0090)	(0.0127)	
LTR	−0.0061**	0.0045	0.0724	−0.0103**	0.0056	0.0969	−0.0106**	0.0006	0.0571	−0.0246***	−0.0036	0.1564	−0.0216**	−0.0059	0.0820
	(0.0032)	(0.0037)		(0.0045)	(0.0047)		(0.0054)	(0.0074)		(0.0085)	(0.0087)		(0.0102)	(0.0123)	
LTY	−0.0065**	−0.0002	0.0712	−0.0107**	0.0031	0.0830	−0.0106**	0.0018	0.0472	−0.0243***	−0.0001	0.1547	−0.0211**	−0.0023	0.0905
	(0.0031)	(0.0033)		(0.0047)	(0.0039)		(0.0053)	(0.0052)		(0.0085)	(0.0060)		(0.0100)	(0.0077)	
DFS	−0.0070**	−0.0042*	0.0730	−0.0118***	−0.0093***	0.1191	−0.0111**	−0.0044	0.0607	−0.0236***	0.0067	0.1499	−0.0207**	0.0034	0.0714
	(0.0030)	(0.0031)		(0.0047)	(0.0038)		(0.0052)	(0.0067)		(0.0084)	(0.0071)		(0.0099)	(0.0105)	
DFY	−0.0065**	−0.0018	0.0620	−0.0106**	−0.0044	0.0800	−0.0106**	−0.0003	0.0556	−0.0245***	0.0071*	0.1680	−0.0216**	0.0177***	0.1139
	(0.0032)	(0.0036)		(0.0049)	(0.0049)		(0.0053)	(0.0045)		(0.0083)	(0.0046)		(0.0093)	(0.0064)	
DE	−0.0063**	−0.0016	0.0700	−0.0104**	−0.0037	0.0785	−0.0098**	−0.0078**	0.0481	−0.0235***	−0.0069	0.1473	−0.0205**	−0.0049	0.0895
	(0.0032)	(0.0030)		(0.0047)	(0.0038)		(0.0052)	(0.0044)		(0.0086)	(0.0055)		(0.0100)	(0.0073)	
TMS	−0.0065**	−0.0006	0.0712	−0.0107**	0.0075	0.0853	−0.0106**	0.0070	0.0509	−0.0243***	0.0042	0.1564	−0.0211**	0.0000	0.0897
	(0.0031)	(0.0079)		(0.0047)	(0.0090)		(0.0053)	(0.0115)		(0.0085)	(0.0138)		(0.0100)	(0.0176)	
INFL	−0.0069**	−0.0020	0.0764	−0.0115**	−0.0038	0.0931	−0.0121**	−0.0077	0.0753	−0.0253***	−0.0051	0.1606	−0.0210**	0.0003	0.0758
	(0.0033)	(0.0032)		(0.0050)	(0.0044)		(0.0058)	(0.0065)		(0.0088)	(0.0085)		(0.0108)	(0.0114)	
GIP	−0.0064**	0.0015	0.0399	−0.0096**	0.0114***	0.1227	−0.0102**	0.0044	0.0269	−0.0233***	0.0099*	0.1158	−0.0202**	0.0090	0.0501
	(0.0031)	(0.0033)		(0.0046)	(0.0046)		(0.0054)	(0.0066)		(0.0080)	(0.0074)		(0.0098)	(0.0095)	
GCNG	−0.0064**	0.0008	0.0555	−0.0105**	0.0040	0.0818	−0.0109**	−0.0029	0.0602	−0.0247***	−0.0051	0.1618	−0.0213**	−0.0032	0.0911
	(0.0032)	(0.0026)		(0.0049)	(0.0048)		(0.0053)	(0.0055)		(0.0084)	(0.0060)		(0.0100)	(0.0096)	
GCND	−0.0065**	−0.0005	0.0693	−0.0105**	0.0046	0.0944	−0.0109**	−0.0047	0.0654	−0.0242***	0.0022	0.1544	−0.0210**	0.0023	0.0651
	(0.0031)	(0.0029)		(0.0048)	(0.0051)		(0.0053)	(0.0079)		(0.0085)	(0.0069)		(0.0100)	(0.0116)	
GCS	−0.0067**	0.0020	0.0668	−0.0105**	−0.0028	0.0897	−0.0097**	−0.0098**	0.0657	−0.0229***	−0.0151***	0.1483	−0.0191**	−0.0214***	0.0718
	(0.0033)	(0.0029)		(0.0046)	(0.0048)		(0.0053)	(0.0048)		(0.0081)	(0.0065)		(0.0092)	(0.0073)	
Panel B: Sentiment Variables															
UMCI	−0.0053*	−0.0062**	0.0353	−0.0100**	−0.0038	0.0601	−0.0116**	0.0046	0.0478	−0.0244***	0.0005	0.1385	−0.0207**	−0.0020	0.0677
	(0.0034)	(0.0033)		(0.0052)	(0.0057)		(0.0050)	(0.0058)		(0.0088)	(0.0078)		(0.0101)	(0.0100)	
CBCI	−0.0069**	−0.0026	0.0611	−0.0104**	0.0022	0.0667	−0.0100**	0.0043	0.0364	−0.0240***	0.0019	0.1367	−0.0208**	0.0022	0.0692
	(0.0032)	(0.0041)		(0.0050)	(0.0049)		(0.0055)	(0.0070)		(0.0086)	(0.0078)		(0.0106)	(0.0109)	
PMI	−0.0066**	−0.0009	0.0565	−0.0106**	0.0022	0.0702	−0.0101**	0.0071	0.0381	−0.0237***	0.0083	0.1356	−0.0201**	0.0135*	0.0663
	(0.0031)	(0.0033)		(0.0048)	(0.0048)		(0.0055)	(0.0062)		(0.0086)	(0.0073)		(0.0103)	(0.0092)	
BWSI	−0.0068**	−0.0015	0.0586	−0.0122***	−0.0074**	0.0894	−0.0107**	−0.0005	0.0414	−0.0249***	−0.0033	0.1451	−0.0216**	−0.0026	0.0737
	(0.0032)	(0.0030)		(0.0047)	(0.0044)		(0.0052)	(0.0061)		(0.0087)	(0.0081)		(0.0097)	(0.0090)	

This Table presents the coefficient estimates and their associated bootstrapped standard errors (in parenthesis) from bivariate regressions. The dependent variable is the cumulative h-month excess return of the S&P 500 index, $ER_{t+1:t+h}$ (h=1, 3, 6, 9 and 12). β is the coefficient of the Employee Sentiment, $ES_{E,t}$, and γ is the coefficient of economic variables, $X_{i,t}$, in Panel A and changes in alternative measures of sentiment, $\Delta Sentiment_{j,t}$, in Panel B. We control for the following economic variables: dividend to price ratio (DP), dividend yield (DY), earning to price ratio (EP), book to market ratio (BM), stock market variance ($SVAR$), net equity expansion ($NTIS$), treasury-bill rate (TBL), long-term bond return (LTR), long-term bond yield (LTY), default return spread (DFR), default yield spread (DFY), dividend payout ratio (DE), term spread (TMS) inflation rate ($INFL$), industrial production growth(GIP), consumer durable goods growth ($GCDR$), non-durable goods growth ($GCNDG$) and services growth (GCS). We control for the following sentiment variables: the University of Michigan Consumer Sentiment Index ($UMCI$) , the Conference Board Consumer Confidence Index ($CBCI$), the Purchasing Managers' Index (PMI), and the Baker and Wurgler Investor Sentiment Index ($BWSI$). All variables are standardized. ΔR^2 shows the change in the R^2 by adding the ES to a regression of excess stock market return on a constant and the control variable [*name in row*]. The sample period is from June 2012 to July 2018. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively, using a one-side test.

market returns. The use of common factors instead of the variables themselves does not only alleviate potential multicollinearity issues, but it also reduces the dimension of the model which is important in finite samples. The first principal component of sentiment explains 61.40% of the common variation across sentiment proxies. Likewise, the five principal components of the 18 economic variables explain roughly 73.29% of their common variation.[18] The results are tabulated in the Online Appendix.

Finally, it could be argued that firms with a low number of employee outlook reviews add noise to the index. To accommodate potential concerns related to the this point, we construct three different indexes. First, we use a different construction methodology that allocates more weight to firms with more reviews (number of reviews-weighted index). The results (untabulated) are qualitatively similar. Second, we construct the employee sentiment index using only firms with at least q reviews per month (q is either 2 or 5). This filter excludes from the index companies with low participation as a low number of reviews may represent a biased view of the outlook of a particular firm. This alternative proxy leads to weaker but significant results with coefficients ranging from -0.50% to -1.64% (untabulated). Finally, we test an index using only reviews from public firms. The coefficients range from -0.48% to -1.50% (untabulated). The slight deterioration in the results may be due to the fact that firms with low participation are likely to be small firms, which have been shown to be more sensitive to sentiment (Bachmann and Elstner, 2015; Massenot and Pettinicchi, 2018). Moreover, by eliminating a large part of the labor market, the proposed index is no longer a market-wide employee sentiment proxy, but a large-firm or public-firm sentiment measure, instead.

4.2. Employee Sentiment and Portfolio Returns

Baker and Wurgler (2006) show that sentiment has a strong effect on the cross-section of stock returns. Mispricing is the main explanation for the documented cross-sectional variation in stock returns, arising from sentiment-driven variation in relative speculative demand and limits to arbitrage. Sentiment is expected to impact the relative value of different securities as: (a) certain types of stocks (e.g., those of young or smaller firms) are more prone to speculative trading and misvaluation than others, and (b) certain stocks that face arbitrage limitations are

[18] We retain the number of principal components for the economic and sentiment variables, respectively, based on the Kaiser Criterion (eigenvalues>1).

more costly to trade and more difficult to short-sell (Shleifer and Vishny, 1997; D'Avolio, 2002). Therefore, these stocks are subject to speculation and are sensitive to shifts in sentiment.[19]

Motivated by these theoretical predictions, we analyze the impact of shifts in employee sentiment on stock portfolios sorted by different characteristics related to firm size and age, profitability, dividend payout, riskiness, firm valuation, growth, and financial distress. The results of this analysis provide additional implications for the cross-section of stock returns. More precisely, all stocks are sorted on a certain characteristic to form equal-weighted[20] decile portfolios.[21] We then regress the cumulative 1-, 3-, 6-, 9-, and 12-month portfolio returns of the 1, 5 and 10 deciles on the lagged value of the change in the ES index.

The characteristics used to form portfolios are the following. Single sorts of stocks on size, measured by the market value of equity, ME (stock price times the number of shares outstanding), and age are considered. Age corresponds to the number of years a firm is listed in *Compustat*. Baker and Wurgler (2006) find a significant cross-sectional relationship between sentiment and the returns of different size and age stocks. Several papers link stock beta to sentiment based on speculation and constraints to arbitrage arguments (e.g., see Stambaugh et al., 2012; Baker et al., 2011; Antoniou et al., 2015), implying that high-beta stocks are more sensitive to sentiment than low-beta stocks. To this end, we examine beta-sorted decile portfolios. We also employ portfolios formed sorting stocks based on their variance, V (variance of 60-day lagged daily returns with at least 20 observations).

We also examine the relationship between employee sentiment and profitability, accruals and operating performance (Stambaugh et al., 2012; Hirshleifer et al., 2004; Sloan, 1996). Decile portfolios are formed by sorting the stocks on E/P (total earnings before extraordinary items for the last fiscal year end in $t-1$ divided by price times shares outstanding at the end of December of $t-1$), accruals, AC (change in operating working capital per split-adjusted share from the fiscal

[19]As highlighted in Baker and Wurgler (2006) it is not easy to distinguish between these two channels since high speculative assets also tend to be riskier and more expensive to arbitrage. Thus, they have similar predictions.

[20]In the online Appendix we also perform analyses with value-weighted portfolios investigating whether the effects of market-wide employee sentiment are economically significant.

[21]The majority of the characteristic-sorted portfolios are obtained from the web page of Ken French (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). When these portfolios are not available, we estimate them using the *CRSP* and *Compustat* databases to access return and financial information on common stocks (with share codes 10 and 11) traded in *NYSE*, *AMEX* and *NASDAQ*. Following prior literature, financial firms are excluded from the sample (sic codes between 6000–6999). Following the standard practice (e.g., Fama and French, 1992, 2015), accounting data for fiscal year-ends in year $t-1$ are matched with the monthly excess returns from July of year t to June of year $t+1$. This is done to ensure that the accounting information is known prior to the period used to compute the stock returns. In all cases, the *NYSE* breakpoints are used to allocate the firms to each decile portfolio according to a certain characteristic.

year end $t-2$ to $t-1$ divided by book equity per share in $t-1$), and operating profitability, OP (annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses divided by book equity for the last fiscal year end in $t-1$). High dividend-paying stocks are considered less risky, easier to value and more liquid (Litzenberger and Ramaswamy, 1979; Amihud, 2002). To this end, dividend-sorted portfolios (DY) are formed (ex-date dividends times the number of shares outstanding, divided by the book equity value).

It is often argued in the literature that firms with more intangible assets, measured by the ratio of fixed assets (i.e., property, plant and equipment) over the total assets (PPE), and firms with higher research and development spending as a percentage of their total assets (RD), are more difficult to value (Chan et al., 2001; Baker and Wurgler, 2006; Lin, 2012; Hou et al., 2015). Therefore, the returns of these firms are expected to be more sensitive to fluctuations in sentiment. We thus form portfolios by sorting the stocks on these characteristics.

Finally, portfolios that sort stocks by growth opportunities and financial distress are used with the following characteristics: the ratio of book value to market value of equity, BM; investments, I (change in total assets from the fiscal year ending in year $t-2$ to the fiscal year ending in $t-1$, divided by $t-2$ total assets) (e.g., Titman et al., 2004; Stambaugh et al., 2012); net stock issues, NSI (change in the natural log of split-adjusted shares from the $t-2$ fiscal year end to the $t-1$ fiscal year end) (e.g., Stambaugh et al., 2012; Loughran and Ritter, 1995); sales growth, SG (annual percentage change in net sales) (Lakonishok et al., 1994); external finance, EF (change in total assets minus the change in retained earnings over the value of total assets) (e.g., see Bradshaw et al., 2006); and leverage, L (long-term plus current debt over the total value of shareholders' equity) (e.g., see George and Hwang, 2010). High book-to-market, high leverage, low sales growth, and low external finance often signal distress. Firms on the other side of the spectrum have high growth opportunities. Firms in the middle are usually stable. Furthermore, both high growth, and distressed firms are more costly to arbitrage.

The results in Table 4 show that changes in the aggregate employee sentiment predict characteristic-sorted portfolios with a negative sign for the longer-term horizons. In the short-term horizons, market-wide sentiment effects manifest consistently for the size, risk and profitability-sorted portfolios including dividend yield. Altogether, sentiment betas differ across different characteristics and horizons. A closer look at the results reveals interesting monotonic cross-sectional patterns at the top and bottom deciles with firms exhibiting different sensitivity to sentiment shocks. For example, harder to value and more costly to arbitrage stocks are more sensitive to sentiment changes. The results for the portfolios formed on tangibility, in line with the evidence of Baker and Wurgler (2006), exhibit systematically a negative association with sentiment, which is similar in magnitude across deciles. A similar picture emerges from the RD-sorted portfolios.

The slope estimates of the portfolios sorted on SG, BM, and EF, reported in Panel E, show the U–shaped pattern documented by Baker and Wurgler (2006). More specifically, employee sentiment has a stronger impact on the top and bottom deciles compared to the middle deciles. Finally, the results from the sorts on leverage, point towards a much stronger effect for low leverage than for high leverage stocks. This rather counterintuitive finding is well documented with empirical studies on the "distress risk puzzle" (e.g., Griffin and Lemmon, 2002; Garlappi et al., 2008; Campbell et al., 2008; George and Hwang, 2010).

We also investigate whether aggregated employee sentiment can predict the returns on Fama-French industry portfolios. Our results (presented in the Online Appendix) show that the slope estimate of the *ES* is consistently negative and significant for several of the tested industries.

4.3. Firm-specific and Market-Wide Employee Sentiment: Predicting Firm Stock Returns

In this section, we examine both changes in firm-specific and market-wide employee sentiment and their ability in predicting contemporaneous and future returns. To this end, we hand match *Glassdoor* data with assets on *CRSP* database. From this process, we have data for 1,036 stocks yielding a total sample of 43,079 firm-month contemporaneous observations.[22] The empirical model we test is described as follows:

$$R_{i,t+1:t+h} = \alpha + \beta \Delta ES_t + \gamma \Delta FES_{i,t} + \delta EMR_t + \varepsilon_{i,t+1:t+h}, \tag{7}$$

where the $R_{i,t+h}$ is the return on the asset i (h = 0, 1, 3, 6, 9, 12), $FES_{i,t}$ is the firm-level employee business outlook expectations, EMR controls for market returns, and $\varepsilon_{i,t+1:t+h}$ represents robust innovations clustered at the firm level.

[22]For each firm, we estimate the correlation between the firm-level employee sentiment and market employee sentiment. We find that the average correlation is 0.0637 for the levels and 0.0128 for differences that are used in the analysis, suggesting that the information content of these two variables is different.

Table 4

Employee Sentiment and Characteristic-sorted Portfolio Return Predictability (Equal–Weighted)

	Dependent Variable: Characteristic Portfolio Returns (Equal–Weighted)														
Horizon	t+1			t+3			t+6			t+9			t+12		
Decile	(1)	(5)	(10)	(1)	(5)	(10)	(1)	(5)	(10)	(1)	(5)	(10)	(1)	(5)	(10)
							Panel A: Size and Age								
ME	−0.0046	−0.0087**	−0.0069**	−0.0110*	−0.0128**	−0.0127***	−0.0123	−0.0101	−0.0143***	−0.0398***	−0.0374***	−0.0297***	−0.0353**	−0.0269**	−0.0253***
	(0.0041)	(0.0041)	(0.0031)	(0.0068)	(0.0058)	(0.0044)	(0.0123)	(0.0097)	(0.0057)	(0.0171)	(0.0135)	(0.0092)	(0.0188)	(0.0151)	(0.0107)
Age	0.0001	−0.0061*	−0.0056*	−0.0071	−0.0090*	−0.0092**	−0.0079	−0.0058	−0.0111*	−0.0383**	−0.0310**	−0.0274***	−0.0341**	−0.0204	−0.0171*
	(0.0051)	(0.0039)	(0.0034)	(0.0080)	(0.0056)	(0.0053)	(0.0138)	(0.0096)	(0.0068)	(0.0181)	(0.0138)	(0.0099)	(0.0175)	(0.0163)	(0.0118)
							Panel B: Risk								
Beta	−0.0044*	−0.0059*	−0.0114**	−0.0083**	−0.0132**	−0.0224**	−0.0073	−0.0125	−0.0216	−0.0264***	−0.0371***	−0.0603***	−0.0221**	−0.0297**	−0.0529**
	(0.0030)	(0.0040)	(0.0057)	(0.0041)	(0.0063)	(0.0104)	(0.0069)	(0.0100)	(0.0178)	(0.0091)	(0.0134)	(0.0254)	(0.0107)	(0.0157)	(0.0276)
V	−0.0028	−0.0056*	−0.0064	−0.0059*	−0.0118**	−0.0143*	−0.0054	−0.0122*	−0.0155	−0.0184***	−0.0331***	−0.0539**	−0.0113	−0.0260**	−0.0443**
	(0.0033)	(0.0038)	(0.0054)	(0.0038)	(0.0052)	(0.0095)	(0.0050)	(0.0089)	(0.0163)	(0.0077)	(0.0122)	(0.0233)	(0.0096)	(0.0147)	(0.0231)
							Panel C: Profitability and Dividend Payout								
E/P	−0.0059*	−0.0051*	−0.0072**	−0.0137***	−0.0124***	−0.0137**	−0.0122*	−0.0113*	−0.0145	−0.0368***	−0.0306***	−0.0436**	−0.0311**	−0.0225*	−0.0358**
	(0.0037)	(0.0035)	(0.0040)	(0.0054)	(0.0048)	(0.0072)	(0.0093)	(0.0079)	(0.0130)	(0.0132)	(0.0111)	(0.0189)	(0.0146)	(0.0141)	(0.0206)
AC	−0.0065*	−0.0061**	−0.0048	−0.0153***	−0.0121**	−0.0114*	−0.0135	−0.0125*	−0.0134	−0.0402***	−0.0390***	−0.0450**	−0.0363**	−0.0327**	−0.0405**
	(0.0046)	(0.0035)	(0.0047)	(0.0063)	(0.0058)	(0.0077)	(0.0111)	(0.0094)	(0.0139)	(0.0157)	(0.0133)	(0.0194)	(0.0172)	(0.0149)	(0.0210)
OP	−0.0041	−0.0064*	−0.0084**	−0.0123*	−0.0125***	−0.0141**	−0.0127	−0.0106*	−0.0173*	−0.0479***	−0.0283***	−0.0432***	−0.0423**	−0.0188*	−0.0386**
	(0.0052)	(0.0039)	(0.0037)	(0.0082)	(0.0044)	(0.0061)	(0.0141)	(0.0078)	(0.0105)	(0.0202)	(0.0107)	(0.0150)	(0.0212)	(0.0138)	(0.0167)
DY	−0.0089**	−0.0066*	−0.0057*	−0.0144***	−0.0130***	−0.0107**	−0.0121*	−0.0126**	−0.0134*	−0.0351***	−0.0299***	−0.0368***	−0.0250**	−0.0203*	−0.0295**
	(0.0040)	(0.0041)	(0.0036)	(0.0057)	(0.0046)	(0.0056)	(0.0091)	(0.0074)	(0.0101)	(0.0122)	(0.0109)	(0.0141)	(0.0146)	(0.0137)	(0.0163)
							Panel D: Tangibility								
PPE	0.0040	0.0034	0.0021	0.0024	−0.0012	−0.0088	−0.0072	−0.0117	−0.0074	−0.0269*	−0.0242**	−0.0247	−0.0434**	−0.0373***	−0.0419**
	(0.0052)	(0.0041)	(0.0062)	(0.0070)	(0.0060)	(0.0094)	(0.0134)	(0.0105)	(0.0161)	(0.0169)	(0.0129)	(0.0208)	(0.0204)	(0.0138)	(0.0232)
RD	0.0003	0.0015	0.0036	−0.0070*	−0.0039	0.0083	−0.0132	−0.0125	0.0014	−0.0253**	−0.0272**	−0.0237	−0.0431***	−0.0405***	−0.0508**
	(0.0034)	(0.0039)	(0.0074)	(0.0051)	(0.0050)	(0.0112)	(0.0113)	(0.0109)	(0.0172)	(0.0141)	(0.0137)	(0.0203)	(0.0157)	(0.0160)	(0.0234)
							Panel E: Financial Distress/Growth Opportunities								
BM	−0.0041	−0.0065*	−0.0079**	−0.0114**	−0.0142***	−0.0149**	−0.0106	−0.0128*	−0.0154	−0.0383***	−0.0385***	−0.0421**	−0.0334**	−0.0315**	−0.0343*
	(0.0044)	(0.0040)	(0.0043)	(0.0067)	(0.0054)	(0.0077)	(0.0108)	(0.0096)	(0.0134)	(0.0141)	(0.0136)	(0.0197)	(0.0145)	(0.0157)	(0.0223)
I	−0.0046	−0.0058*	−0.0048	−0.0145**	−0.0130***	−0.0090	−0.0176	−0.0138*	−0.0086	−0.0509***	−0.0339***	−0.0421***	−0.0443**	−0.0256**	−0.0361**
	(0.0051)	(0.0038)	(0.0048)	(0.0087)	(0.0046)	(0.0073)	(0.0149)	(0.0086)	(0.0124)	(0.0208)	(0.0119)	(0.0180)	(0.0223)	(0.0151)	(0.0183)
NSI	−0.0055*	−0.0073**	−0.0039	−0.0150***	−0.0167***	−0.0067	−0.0118	−0.0159*	−0.0098	−0.0325**	−0.0436***	−0.0463**	−0.0208	−0.0384**	−0.0429**
	(0.0042)	(0.0041)	(0.0060)	(0.0058)	(0.0059)	(0.0089)	(0.0101)	(0.0108)	(0.0156)	(0.0141)	(0.0145)	(0.0216)	(0.0171)	(0.0172)	(0.0214)
SG	0.0039	0.0011	0.0019	0.0008	−0.0025	−0.0016	−0.0065	−0.0125*	−0.0067	−0.0269*	−0.0235**	−0.0235	−0.0438***	−0.0299***	−0.0470***
	(0.0050)	(0.0036)	(0.0050)	(0.0081)	(0.0051)	(0.0070)	(0.0132)	(0.0090)	(0.0141)	(0.0167)	(0.0110)	(0.0186)	(0.0181)	(0.0126)	(0.0200)
EF	0.0040	0.0023	0.0032	−0.0025	−0.0023	0.0052	−0.0110	−0.0100	−0.0024	−0.0248**	−0.0194**	−0.0219	−0.0368**	−0.0279**	−0.0471**
	(0.0044)	(0.0035)	(0.0058)	(0.0062)	(0.0048)	(0.0084)	(0.0111)	(0.0091)	(0.0146)	(0.0144)	(0.0118)	(0.0187)	(0.0162)	(0.0140)	(0.0212)
L	−0.0001	0.0028	0.0005	−0.0012	−0.0029	−0.0057	−0.0106	−0.0130	−0.0133	−0.0359*	−0.0257**	−0.0313**	−0.0700***	−0.0385***	−0.0475***
	(0.0063)	(0.0039)	(0.0040)	(0.0106)	(0.0050)	(0.0061)	(0.0203)	(0.0102)	(0.0125)	(0.0263)	(0.0138)	(0.0165)	(0.0245)	(0.0158)	(0.0180)

Note: This table presents the slope estimates from OLS regressions of 1, 3, 6, 9 and 12–month ahead equal-weighted returns on characteristic–sorted decile portfolios on the lagged value of the *ES*. Bootstrap robust standard errors are reported in parentheses. 1, 5 and 10 decile portfolios are formed from single sorts on market equity (ME), age, beta, variance (V), residual variance (RV), earnings (E/P), accruals (AC), operating profit (OP), dividend yield (DY), fixed over total assets (PPE), research and development over total assets (RD), book-to-market equity ratio (BM), investment (I), net share issues (NSI), sales growth (SG), external finance over total assets (EF) and market leverage (L). ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table 5 reports a strong negative effect for the *ES* on 1, 3, 6, 9, and 12-horizon ahead asset returns, while there is no significant effect on contemporaneous returns. This reveals a similar pattern to how the *ES* performs in predicting the aggregate stock market. In line with the literature that investigates firm-level employee sentiment (Green et al., 2019; Sheng, 2019; Hales et al., 2018), we find a positive association between changes in *FES* and contemporaneous and future asset returns. Our findings accord with a distinct effect of firm-level and market-level employee sentiment. While the former appears to reveal information related to fundamentals, the latter appears to capture correlated judgement errors. Such empirical findings complement theoretical studies that discuss how overconfidence to private signals may distort perceptions across several assets leading to correlated trading (e.g., Daniel et al., 2001; Peng and Xiong, 2006), and studies that call for asset pricing models extended with both market-wide and firm-related sentiment (Mahmoudi et al., 2020).

Table 5

Aggregate Employee Sentiment vs. Firm Employee Sentiment

	Dependent: Excess Stock Returns (Firm level)					
	(1)	(2)	(3)	(4)	(5)	(6)
Horizon	t+0	t+1	t+3	t+6	t+9	t+12
ES	−0.0043	−0.0597***	−0.0546***	−0.0458***	−0.1123***	−0.0725***
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
FES	0.0165***	0.0031	0.0060*	0.0092***	0.0074**	0.0055**
	(0.0008)	(0.0008)	(0.0010)	(0.0012)	(0.0014)	(0.0014)
EMR	0.2819***	−0.0377***	−0.0751***	−0.0094**	0.0337***	−0.0067*
	(0.0200)	(0.0177)	(0.0281)	(0.0337)	(0.0405)	(0.0451)
Observations	43,079	42,959	42,724	42,315	41,888	41,458

This table presents the results from the predictive regressions of h–period ahead firm level returns using both firm-level employee sentiment and aggregate employee sentiment and controlling for excess aggregate stock market returns. Robust standard errors clustered at firm level are reported in parentheses. We consider a panel data set of contemporaneous stock returns and stock returns in forecast horizons of 1, 3, 6, 9 and 12 months, respectively. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

4.4. Employee Sentiment and Economic Channels

We next explore the channel that explain the predictability of stock market returns by employee sentiment. According to traditional asset pricing theory, prices vary due to changing expectations of future cash flows, discount rates or both (Campbell and Shiller, 1988*b*). There are thus two possible channels that can explain the predictability of stock market returns by employee sentiment, the cash-flow channel and the discount rate channel.

The explanation of biased expectations about future cash flows unrelated to fundamentals is well-supported in the literature (e.g., Baker and Wurgler, 2007; Greenwood and Shleifer, 2014; Huang, Li, Meschke and Guthrie, 2015). The rationale of this channel is that extrapolation of past good cash-flow news, due to overoptimism driven by positive sentiment shocks, leads to mispricing (overvaluation). The subsequent correction to this overvaluation yields lower returns. Therefore, if this channel explains the negative return predictability, positive current sentiment shocks should predict subsequent negative cash-flow changes. To test this hypothesis, we perform the following regression:

$$CF_{t+1:t+h} = \kappa + \lambda \Delta ES_t + u_{t+1:t+h}, \tag{8}$$

where $CF_{t+1:t+h}$ is the h-month growth rate of the cash flows ($h = 1, 3, 6, 9, 12$). We first employ the aggregate dividend growth as a cash flow proxy (e.g., Campbell and Shiller, 1988*a*; Fama and French, 2000; Lettau and Ludvigson, 2005; Cochrane, 2008, 2011*b*). Given the evidence on dividend smoothing policies (Leary and Michaely, 2011) and their effect on return predictability (Chen et al., 2012), we also use aggregate earnings growth (e.g., Ang and Bekaert, 2006; Chen et al., 2012) and industrial production growth (e.g., Fama, 1990; Greenwood and Shleifer, 2014) as additional cash flow measures. Thus, $CF = \{$*Dividend Growth, Earnings Growth, Industrial Production Growth*$\}$, where growth rates are computed as the h-period logarithmic differences of the corresponding measure.

Panel A of Table 6 presents that the coefficients have the anticipated negative sign and are significant, especially at longer horizons. When cash flows are measured based on earnings rather than dividends, the predictability becomes stronger. This result implies that dividend smoothing may be a reason behind the lower information content of dividend-based cash flow measures (Chen et al., 2014). In sum, the empirical evidence provides support for the cash flow channel indicating that the extrapolation of positive sentiment shocks is expected to lead to negative future cash flows. Not surprisingly, the effect becomes more prominent at longer horizons, given that fundamentals are gradually revealed. Our results are consistent with the empirical implications of extrapolative expectations models (e.g., Greenwood and Shleifer, 2014).

If discount rate is the main explanation for the negative predictive ability of employee sentiment for future stock returns, then the employee sentiment should be significantly related to discount rates. Following the literature, we employ the dividend-to-price (D/P) ratio as a proxy

Table 6

Employee Sentiment Index and Economic Channels

Horizon	t+1	t+3	t+6	t+9	t+12
			Panel A: Cash Flows		
Earning Growth					
L.ES	−0.0018*	−0.0058*	−0.0130*	−0.0176*	−0.0235*
	(0.0014)	(0.0042)	(0.0085)	(0.0123)	(0.0150)
Dividend Growth					
L.ES	−0.0005	−0.0011	−0.0019	−0.0032*	−0.0045*
	(0.0005)	(0.0012)	(0.0017)	(0.0023)	(0.0029)
Industrial Production Growth					
L.ES	−0.0007*	−0.0019**	−0.0026*	−0.0042*	−0.0048*
	(0.0006)	(0.0010)	(0.0019)	(0.0026)	(0.0029)
			Panel B: Discount Rate		
Dividend Price					
L.ES	−0.0016	−0.0023	0.0064	0.0264	0.0592
	(0.0063)	(0.0170)	(0.0327)	(0.0472)	(0.0564)
			Panel C: Volatility		
LVol					
L.ES	0.0415*	0.0386**	0.0284**	0.0201	0.0225*
	(0.0253)	(0.0183)	(0.0157)	(0.0158)	(0.0158)
L.LVol	0.6182***	0.3911***	0.2571***	0.2140***	0.1498***
	(0.1003)	(0.0948)	(0.0886)	(0.0638)	(0.0626)

This table presents the results from the predictive regressions of the h–period ahead cash flows (Panel A), discount rates (Panel B) and volatility (Panel C) on the lagged value of the *ES*. Bootstrap robust standard errors are reported in parentheses. We consider forecast horizons of 1, 3, 6, 9 and 12 months, respectively. The cash flow proxies include the growth rate in the S&P 500 dividends, the growth rate in the S&P 500 earnings and the growth rate in the US industrial production. For the discount rate channel, we employ the dividend-price ratio of the S&P 500 index. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

for discount rates (e.g., Cochrane, 2008; Huang, Jiang, Tu and Zhou, 2015). Several rational expectations models (e.g., Campbell and Cochrane, 1999; Bansal and Yaron, 2004; Cochrane, 2011*b*; Wachter, 2013) use the D/P ratio to capture discount rates variation. Therefore, we test whether changes in employee sentiment significantly predict lower future D/P ratios. Our results presented in Panel B of Table 6 show that the slope estimates of *ES* are not statistically significant. This result is in line with the findings of Huang, Jiang, Tu and Zhou (2015). Overall, the cash flow channel provides the main explanation for the stock market return predictability of the aggregate employee sentiment.

In addition to stock market returns, a large stream of the financial literature studies the impact of sentiment on volatility. This relationship is established through the role of noise traders (De Long et al., 1990). The main idea is that the presence of sentiment-driven noise traders causes prices to deviate from their fundamental values leading to mispricing and excess volatility. Moreover, De Long et al. (1990) argue that changes in sentiment reflect a risk to arbitrageurs. An implication of these is that market-wide sentiment should positively predict future stock market volatility. Another possibility is that higher stock market volatility will lead to higher risk premia (discount rates) (French et al., 1987) which imply higher future returns. Given that bearish sentiment is associated with higher stock market volatility, this argument predicts a negative relationship between sentiment and stock market volatility. We, therefore, test the above conjectures through the following regression:

$$\mathrm{LVol}_{t:t+h} = \zeta + \theta \Delta ES_t + \xi \mathrm{LVol}_t + \eta_{t:t+h}, \tag{9}$$

where $LVol_{t:t+h} = log(\sqrt{\mathrm{Var}_{t:t+h}})$ is the h-month logarithmic volatility ($\sqrt{\mathrm{Var}}$) of the returns on the S&P 500 index. The $\mathrm{Var}_{t:t+h}$ is computed from the sum of squared daily returns over the period from month t to month $t+h$. We consider the natural logarithm of the volatility as its empirical distribution is closer to Gaussian, making it more suitable as a dependent variable (Paye, 2012).

Panel C in Table 6 shows the results for predicting the h–month logarithmic volatility. Our findings indicate that aggregate employee sentiment positively predicts stock market volatility. In particular, the estimated slopes are statistically significant. This significant positive association between employee sentiment and stock market volatility is in contradiction to the discount rate argument lending support to the argument that mispricing caused by sentiment shocks leads to excess volatility. The findings of Lee et al. (1991) also display the pricing of noise traders' sentiment risk through the discounts of closed-end funds that are not related to the riskiness of their underlying assets.

5. Discussion

This paper estimates the market-wide employee sentiment using online information from *Glassdoor* extending the literature that investigates the value of employee opinions for stock markets (e.g., Green et al., 2019; Symitsi et al., 2018; Sheng, 2019; Huang et al., 2020; Symitsi et al., 2021). We differ though from this stream of scholarly thought as we measure the power of aggregate employee sentiment rather than firm-specific employee sentiment, providing insights on the potential of employee data embedded in asset pricing models. In particular, we present a comprehensive empirical investigation of the ability of this factor to complement existing sentiment measures and other economic variables commonly used in financial models extending theoretical and empirical research for the importance of market-wide sentiment factors in

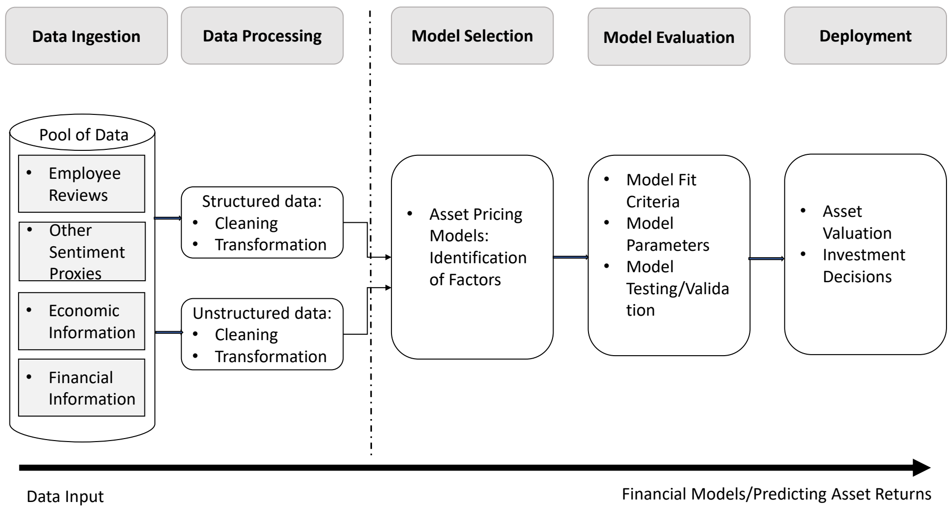

Figure 2. A Machine Learning (ML) Scoring Pipeline for Expert Financial Models in Asset Pricing

asset pricing (Lemmon and Portniaguina, 2006; Baker and Wurgler, 2007). We also test the theoretical underpinnings for the economic channels that drive this predictive force.

Our findings corroborate the idea that sentiment is not a unidimensional factor demonstrating that sentiment from different stakeholders can offer incremental information value in predicting stock returns. Employee opinions altogether seem to convey valuable information and knowledge. These results are consistent even if we control for firm-level employee sentiment, indicating that market-wide sentiment proxies could complement firm-level sentiment proxies (Mahmoudi et al., 2020). More specifically, our results support that non-financial information sources can upgrade financial models. According to that, asset pricing and investment decisions can be assisted when models are enriched and tested with new sources of information. Figure 2 displays an example of a machine learning pipeline that augments traditional data sources and asset pricing factors with online data to facilitate asset pricing and investment decisions.

Estimating market-wide sentiment from employee expectations posted online has several advantages. The anonymity and the voluntary basis for online reviews reduces well-documented biases compared to eponymous interviews and surveys which lead to employee silence under the fear of retribution and senior manager intolerance to negative employee feedback (Milliken et al., 2003; Holland et al., 2016). As discussed earlier, managers may provide biased narratives due to their agency relationship with their employer. Importantly, building a sentiment measure based on an online big dataset of readily available reviews offers a greater coverage without publication delay compared to survey-based sentiment measures, such as the University of Michigan Consumer Sentiment, the Conference Board Consumer Confidence and the Purchasing

Manager's Index which compile responses of up to 500, 5,000 and 400 respondents, respectively.

Despite the appealing properties of our market-wide sentiment measure, online reviews are also known to be subject to biases (Li and Hitt, 2008; Askalidis et al., 2017; Hu et al., 2009). For example, polarization is typically observed in online reviews (U-shape distribution) reflecting a selection bias with extreme views being more common than moderate views, which could convey misleading information. However, non-economic stimuli have been found to reduce such biases (Marinescu et al., 2018). Marinescu et al. (2018) show that the incentive policy of *Glassdoor* encourages also reviews with neutral opinions reducing the polarization bias and leading to more balanced ratings per employer.[23] While fake reviews and manipulation could also be a concern in online anonymous posts (Mayzlin et al., 2014), *Glassdoor* mitigates such risks by employing algorithm and manual controls. Moreover, an index averaging thousands opinions per month across many employers will eliminate idiosyncratic errors rendering a good proxy of the overall employee sentiment across the labor market.

We perform a comprehensive empirical analysis to test whether our sentiment index predicts future market, portfolio and asset returns. However, our empirical approach does not come without limitations. First, using predictive regressions, we can only study the in-sample performance of our sentiment proxy. Future studies, by employing extended samples, could explore the predictive power of market-wide employee sentiment in forecasting returns out-of-sample. Second, given that business outlook ratings are present in *Glassdoor* since 2012, our empirical results have not been tested across larger periods or business cycles. Market sentiment has been found to have asymmetric effects in asset pricing that vary with market conditions (Stambaugh et al., 2012). To this end, further analysis is needed to identify how the employee-sentiment changes across turmoil and tranquil periods and whether its predictive power alters. Third, in line with prior studies, we tested our index employing basic asset pricing models and controlling for economic variables and other sentiment proxies. However, in the literature there is a plethora of studies that do not only model asset returns but also higher moments (e.g., GARCH models). A horse race across advanced and parsimonious models was beyond the purpose of this study, though, future studies could potentially provide evidence on how market-wide employee sentiment performs in augmented financial models.

[23]This "give-to-get" model allows access to the content of the platform after a user submit at least one kind of review, such as company, salary, interview, or benefits review (only one review can be submitted per type per employer per year).

6. Conclusion

We propose a collective employee sentiment measure, aggregating millions of employee opinions voluntarily and anonymously disclosed on *Glassdoor* platform. We find that our index is a significant negative predictor of aggregate stock market returns. The results remain qualitatively unchanged when we control for other measures of investor sentiment and various commonly used macroeconomic variables. We also find that the proposed *ES* forecasts stock returns at the cross-section, particularly for stocks that are difficult to value and costly to arbitrage, though sentiment betas vary significantly with portfolio characteristics, industries and horizons. We also document a distinct power of market-wide employee sentiment on predicting asset price returns after controlling for stock market returns and changes in firm-level employee sentiment. The driving force behind the predictive ability of our index seems to stem from biased beliefs about future cash flows. We also find that market-wide employee sentiment predicts positively future volatility. Overall, our empirical results are consistent with the theoretical predictions of models based on noise-trader sentiment. With investor sentiment used to investigate various issues in finance, employee sentiment, as a new measure of investor sentiment that contains complementary information to existing measures, may yield a variety of future applications in finance, accounting, and economics leading to enhanced forecasting models and expert systems.

References

Aboody, D., Even-Tov, O., Lehavy, R. and Trueman, B. (2018), 'Overnight Returns and Firm-Specific Investor Sentiment', *Journal of Financial and Quantitative Analysis* **53**(2), 485–505.

Ahmad, K., Han, J., Hutson, E., Kearney, C. and Liu, S. (2016), 'Media-expressed Negative Tone and Firm-level Stock Returns', *Journal of Corporate Finance* **37**, 152–172.

Amihud, Y. (2002), 'Illiquidity and stock returns: cross-section and time-series effects', *Journal of Financial Markets* **5**(1), 31–56.

Ang, A. and Bekaert, G. (2006), 'Stock Return Predictability: Is It There?', *Review of Financial Studies* **20**(3), 651–707.

Antoniou, C., Doukas, J. A. and Subrahmanyam, A. (2015), 'Investor sentiment, beta, and the cost of equity capital', *Management Science* **62**(2), 347–367.

Antweiler, W. and Frank, M. Z. (2004), 'Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards', *Journal of Finance* **59**(3), 1259–1294.

Arif, S. and Lee, C. M. (2014), 'Aggregate Investment and Investor Sentiment', *Review of Financial Studies* **27**(11), 3241–3279.

Askalidis, G., Kim, S. J. and Malthouse, E. C. (2017), 'Understanding and overcoming biases in online review systems', *Decision Support Systems* **97**, 23–30.

Babenko, I. and Sen, R. (2015), 'Do Nonexecutive Employees Have Valuable Information? Evidence from Employee Stock Purchase Plans', *Management Science* **62**(7), 1878–1898.

Bachmann, R. and Elstner, S. (2015), 'Firm optimism and pessimism', *European Economic Review* **79**, 297–325.

Baker, M., Bradley, B. and Wurgler, J. (2011), 'Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly', *Financial Analysts Journal* **67**(1), 40–54.

Baker, M. and Stein, J. C. (2004), 'Market liquidity as a sentiment indicator', *Journal of Financial Markets* **7**(3), 271–299.

Baker, M. and Wurgler, J. (2006), 'Investor Sentiment and the Cross-Section of Stock Returns', *Journal of Finance* **61**(4), 1645–1680.

Baker, M. and Wurgler, J. (2007), 'Investor sentiment in the stock market', *Journal of Economic Perspectives* **21**(2), 129–152.

Baker, M., Wurgler, J. and Yuan, Y. (2012), 'Global, local, and contagious investor sentiment', *Journal of Financial Economics* **104**(2), 272–287.

Bansal, R. and Yaron, A. (2004), 'Risks or the Long Run: A Potential Resolution of Asset Pricing Puzzles', *Journal of Finance* **59**(4), 1481–1509.

Barber, B. M., Odean, T. and Zhu, N. (2008), 'Do retail trades move markets?', *Review of Financial Studies* **22**(1), 151–186.

Barberis, N., Shleifer, A. and Vishny, R. (1998), 'A model of investor sentiment', *Journal of Financial Economics* **49**(3), 307–343.

Bartov, E., Faurel, L. and Mohanram, P. S. (2017), 'Can Twitter Help Predict Firm-Level Earnings and Stock Returns?', *Accounting Review* **93**(3), 25–57.

Ben-Rephael, A., Kandel, S. and Wohl, A. (2012), 'Measuring investor sentiment with mutual fund flows', *Journal of Financial Economics* **104**(2), 363–382.

Benartzi, S. (2001), 'Excessive Extrapolation and the Allocation of 401 (k) Accounts to Company Stock', *Journal of Finance* **56**(5), 1747–1764.

Bradshaw, M. T., Richardson, S. A. and Sloan, R. G. (2006), 'The relation between corporate financing activities, analysts' forecasts and stock returns', *Journal of Accounting and Economics* **42**(1-2), 53–85.

Bram, J., Ludvigson, S. et al. (1998), 'Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race', *Economic Policy Review* pp. 59–78.

Campbell, J. L., DeAngelis, M. D. and Moon, J. (2017), 'The Credibility of the Crowds: Personal Stock Holdings and Investors' Response to Stock Opinions on Social Media'.

Campbell, J. Y. and Cochrane, J. H. (1999), 'By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior', *Journal of Political Economy* **107**(2), 205–251.

Campbell, J. Y., Hilscher, J. and Szilagyi, J. (2008), 'In Search of Distress Risk', *Journal of Finance* **63**(6), 2899–2939.

Campbell, J. Y. and Kyle, A. S. (1993), 'Smart Money, Noise Trading and Stock Price Behaviour', *Review of Economic Studies* **60**(1), 1–34.

Campbell, J. Y. and Shiller, R. J. (1988*a*), 'The dividend-price ratio and expectations of future dividends and discount factors', *Review of Financial Studies* **1**(3), 195–228.

Campbell, J. Y. and Shiller, R. J. (1988*b*), 'Stock prices, earnings, and expected dividends', *Journal of Finance* **43**(3), 661–676.

Cao, M. and Wei, J. (2005), 'Stock market returns: A note on temperature anomaly', *Journal of Banking & Finance* **29**(6), 1559–1573.

Chan, L. K., Lakonishok, J. and Sougiannis, T. (2001), 'The Stock Market Valuation of Research and Development Expenditures', *Journal of Finance* **56**(6), 2431–2456.

Chen, H., De, P., Hu, Y. J. and Hwang, B.-H. (2014), 'Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media', *Review of Financial Studies* **27**(5), 1367–1403.

Chen, L., Da, Z. and Priestley, R. (2012), 'Dividend Smoothing and Predictability', *Management science* **58**(10), 1834–1853.

Cochrane, J. H. (2008), 'The Dog That Did Not Bark: A Defense of Return Predictability', *Review of Financial Studies* **21**(4), 1533–1575.

Cochrane, J. H. (2011*a*), 'Presidential address: Discount rates', *Journal of Finance* **66**(4), 1047–1108.

Cochrane, J. H. (2011*b*), 'Presidential Address: Discount Rates', *Journal of Finance* **66**(4), 1047–1108.

Cohen, L. (2008), 'Loyalty-Based Portfolio Choice', *Review of Financial Studies* **22**(3), 1213–1245.

Da, Z., Engelberg, J. and Gao, P. (2015), 'The Sum of All FEARS Investor Sentiment and Asset Prices', *Review of Financial Studies* **28**(1), 1–32.

Daniel, K. D., Hirshleifer, D. and Subrahmanyam, A. (2001), 'Overconfidence, Arbitrage, and Equilibrium Asset Pricing', *Journal of Finance* **56**(3), 921–965.

Daniel, K. and Hirshleifer, D. (2015), 'Overconfident Investors, Predictable Returns, and Excessive Trading', *Journal of Economic Perspectives* **29**(4), 61–88.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (1998), 'Investor Psychology and Security Market Under- and Overreactions', *Journal of Finance* **53**(6), 1839–1885.

D'Avolio, G. (2002), 'The market for borrowing stock', *Journal of Financial Economics* **66**(2-3), 271–306.

De Long, J. B., Shleifer, A., Summers, L. H. and Waldmann, R. J. (1990), 'Noise Trader Risk in Financial Markets', *Journal of Political Economy* **98**(4), 703–738.

DeVault, L., Sias, R., Starks, L. et al. (2019), 'Sentiment Metrics and Investor Demand', *Journal of Finance* **74**(2), 985–1024.

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

Edmans, A., Garcia, D. and Norli, Ø. (2007), 'Sports Sentiment and Stock Returns', *Journal of Finance* **62**(4), 1967–1998.

Edmans, A., Li, L. and Zhang, C. (2014), Employee satisfaction, labor market flexibility, and stock returns around the world, Technical report, National Bureau of Economic Research.

Engelberg, J., McLean, R. D., Pontiff, J. and Ringgenberg, M. (2019), 'Are Cross-Sectional Predictors Good Market-Level Predictors?', *Working Paper* .

Fama, E. F. (1990), 'Stock returns, expected returns, and real activity', *Journal of Finance* **45**(4), 1089–1108.

Fama, E. F. and French, K. R. (1992), 'The cross-section of expected stock returns', *Journal of Finance* **47**(2), 427–465.

Fama, E. F. and French, K. R. (2000), 'Forecasting profitability and earnings', *Journal of Business* **73**(2), 161–175.

Fama, E. F. and French, K. R. (2015), 'A five-factor asset pricing model', *Journal of Financial Economics* **116**(1), 1 – 22.

French, K. R., Schwert, G. W. and Stambaugh, R. F. (1987), 'Expected stock returns and volatility', *Journal of Financial Economics* **19**(1), 3–29.

Gao, Z., Ren, H. and Zhang, B. (2019), 'Googling Investor Sentiment around the World', *Journal of Financial and Quantitative Analysis* **00**(0), 1–32.

Garcia, D. (2013), 'Sentiment during recessions', *Journal of Finance* **68**(3), 1267–1300.

Garlappi, L., Shu, T. and Yan, H. (2008), 'Default Risk, Shareholder Advantage, and Stock Returns', *Review of Financial Studies* **21**(6), 2743–2778.

George, T. J. and Hwang, C.-Y. (2010), 'A resolution of the distress risk and leverage puzzles in the cross section of stock returns', *Journal of Financial Economics* **96**(1), 56–79.

Giannini, R., Irvine, P. and Shu, T. (2018), 'Nonlocal Disadvantage: An Examination of Social Media Sentiment', *The Review of Asset Pricing Studies* **8**(2), 293–336.

Gomes, J., Kogan, L. and Zhang, L. (2003), 'Equilibrium cross section of returns', *Journal of Political Economy* **111**(4), 693–732.

Green, T. C., Huang, R., Wen, Q. and Zhou, D. (2019), 'Crowdsourced Employer Reviews and Stock Returns', *Journal of Financial Economics (forthcoming)* .

Greenwood, R. and Shleifer, A. (2014), 'Expectations of returns and expected returns', *Review of Financial Studies* **27**(3), 714–746.

Griffin, J. M. and Lemmon, M. L. (2002), 'Book-to-Market Equity, Distress Risk, and Stock Returns', *Journal of Finance* **57**(5), 2317–2336.

Grossman, S. J. and Stiglitz, J. E. (1980), 'On the Impossibility of Informationally Efficient Markets', *American Economic Review* **70**(3), 393–408.

Hales, J., Moon, J. and Swenson, L. (2018), 'A New Era of Voluntary Disclosure? Empirical Evidence on the Informativeness of Rank and-File Employees' Business Outlook', *Accounting, Organizations and Society* **68**, 88–108.

Hirshleifer, D., Hou, K. and Teoh, S. H. (2009), 'Accruals, cash flows, and aggregate stock returns', *Journal of Financial Economics* **91**(3), 389–406.

Hirshleifer, D., Hou, K., Teoh, S. H. and Zhang, Y. (2004), 'Do Investors Overvalue Firms with Bloated Balance Sheets?', *Journal of Accounting and Economics* **38**, 297–331.

Hirshleifer, D. and Shumway, T. (2003), 'Good Day Sunshine: Stock Returns and the Weather', *Journal of Finance* **58**(3), 1009–1032.

Holland, P., Cooper, B. K. and Hecker, R. (2016), 'Use of social media at work: a new form of employee voice?', *International Journal of Human Resource Management* **27**(21), 2621–2634.

Hong, H. and Stein, J. C. (1999), 'A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets', *Journal of Finance* **54**(6), 2143–2184.

Horowitz, J. L. (2019), 'Bootstrap methods in econometrics', *Annual Review of Economics* **11**, 193–224.

Hou, K., Xue, C. and Zhang, L. (2015), 'Digesting Anomalies: An Investment Approach', *Review of Financial Studies* **28**(3), 650–705.

Hribar, P. and McInnis, J. (2012), 'Investor sentiment and analysts' earnings forecast errors', *Management Science* **58**(2), 293–307.

Hu, N., Zhang, J. and Pavlou, P. A. (2009), 'Overcoming the j-shaped distribution of product reviews', *Communications of the ACM* **52**(10), 144–147.

Huang, D., Jiang, F., Tu, J. and Zhou, G. (2015), 'Investor Sentiment Aligned: A Powerful Predictor of Stock Returns', *Review of Financial Studies* **28**(3), 791–837.

Huang, J. (2018), 'The customer knows best: The investment value of consumer opinions', *Journal of Financial Economics* **128**(1), 164–182.

Huang, K., Li, M. and Markov, S. (2020), 'What Do Employees Know? Evidence from a Social Media Platform', *The Accounting Review* **95**(2), 199–226.

Huang, M., Li, P., Meschke, F. and Guthrie, J. P. (2015), 'Family firms, employee satisfaction, and corporate performance', *Journal of Corporate Finance* **34**, 108–127.

Huddart, S. and Lang, M. (2003), 'Information distribution within firms: evidence from stock option exercises', *Journal of Accounting and Economics* **34**(1-3), 3–31.

Hvide, H. K. and Östberg, P. (2015), 'Social interaction at work', *Journal of Financial Economics* **117**(3), 628–652.

Inoue, A. and Kilian, L. (2005), 'In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?', *Econometric Reviews* **23**(4), 371–402.

Jame, R., Johnston, R., Markov, S. and Wolfe, M. C. (2016), 'The Value of Crowdsourced Earnings Forecasts', *Journal of Accounting Research* **54**(4), 1077–1110.

Jiang, F., Lee, J. A., Martin, X. and Zhou, G. (2018), 'Manager sentiment and stock returns', *Journal of Financial Economics (forthcoming)* .

Johnson, E. J. and Tversky, A. (1983), 'Affect, generalization, and the perception of risk', *Journal of Personality and Social Psychology* **45**(1), 20–31.

Kamstra, M. J., Kramer, L. A. and Levi, M. D. (2000), 'Losing Sleep at the Market: The Daylight Saving Anomaly', *American Economic Review* **90**(4), 1005–1011.

Kamstra, M. J., Kramer, L. A. and Levi, M. D. (2003), 'Winter Blues: A SAD Stock Market Cycle', *American Economic Review* **93**(1), 324–343.

Kaniel, R., Saar, G. and Titman, S. (2008), 'Individual investor trading and stock returns', *Journal of Finance* **63**(1), 273–310.

Kaplanski, G. and Levy, H. (2010), 'Sentiment and stock prices: The case of aviation disasters', *Journal of Financial Economics* **95**(2), 174–201.

Keynes, J. M. (1936), *The general theory of employment, interest, and money*, London: Palgrave Macmillan.

Kogan, L., Ross, S. A., Wang, J. and Westerfield, M. M. (2006), 'The Price Impact and Survival of Irrational Traders', *Journal of Finance* **61**(1), 195–229.

Kothari, S., Lewellen, J. and Warner, J. B. (2006), 'Stock returns, aggregate earnings surprises, and behavioral finance', *Journal of Financial Economics* **79**(3), 537–568.

Kothari, S. P. and Shanken, J. (1997), 'Book-to-market, dividend yield, and expected market returns: A time-series analysis', *Journal of Financial Economics* **44**(2), 169–203.

Kumar, A. and Lee, C. M. (2006), 'Retail Investor Sentiment and Return Comovements', *Journal of Finance* **61**(5), 2451–2486.

Lakonishok, J., Shleifer, A. and Vishny, R. W. (1994), 'Contrarian Investment, Extrapolation, and Risk', *Journal of Finance* **49**(5), 1541–1578.

Larcker, D. F. and Zakolyukina, A. A. (2012), 'Detecting deceptive discussions in conference calls', *Journal of Accounting Research* **50**(2), 495–540.

Leary, M. T. and Michaely, R. (2011), 'Determinants of Dividend Smoothing: Empirical Evidence', *Review of Financial Studies* **24**(10), 3197–3249.

Lee, C. M., Shleifer, A. and Thaler, R. H. (1991), 'Investor Sentiment and the Closed-End Fund Puzzle', *Journal of Finance* **46**(1), 75–109.

Lemmon, M. and Portniaguina, E. (2006), 'Consumer Confidence and Asset Prices: Some Empirical Evidence', *Review of Financial Studies* **19**(4), 1499–1529.

Lettau, M. and Ludvigson, S. C. (2005), 'Expected returns and expected dividend growth', *Journal of Financial Economics* **76**(3), 583–626.

Li, X. and Hitt, L. M. (2008), 'Self-selection and information role of online product reviews', *Information Systems Research* **19**(4), 456–474.

Lin, X. (2012), 'Endogenous technological progress and the cross-section of stock returns', *Journal of Financial Economics* **103**(2), 411–427.

Litzenberger, R. H. and Ramaswamy, K. (1979), 'The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence', *Journal of Financial Economics* **7**(2), 163–195.

Loughran, T. and Ritter, J. R. (1995), 'The New Issues Puzzle', *Journal of Finance* **50**(1), 23–51.

Ludvigson, S. C. (2004), 'Consumer Confidence and Consumer Spending', *Journal of Economic perspectives* **18**(2), 29–50.

Mahmoudi, N., Docherty, P. and Melia, A. (2020), 'Firm-Level Versus Aggregate Market Investor Sentiment: What Matters Most for Corporate Announcement Returns', *Working paper. Available at SSRN 3391698* .

Marinescu, I., Klein, N., Chamberlain, A. and Smart, M. (2018), Incentives can reduce bias in online reviews, Technical report, National Bureau of Economic Research.

Massenot, B. and Pettinicchi, Y. (2018), 'Can firms see into the future? Survey evidence from Germany', *Journal of Economic Behavior & Organization* **145**, 66–79.

Mayzlin, D., Dover, Y. and Chevalier, J. (2014), 'Promotional reviews: An empirical investigation of online review manipulation', *American Economic Review* **104**(8), 2421–55.

Miller, G. S. and Skinner, D. J. (2015), 'The Evolving Disclosure Landscape: How Changes in Technology, the Media, and Capital Markets Are Affecting Disclosure', *Journal of Accounting Research* **53**(2), 221–239.

Milliken, F. J., Morrison, E. W. and Hewlin, P. F. (2003), 'An Exploratory Study of Employee Silence: Issues that Employees Don't Communicate Upward and Why', *Journal of Management Studies* **40**(6), 1453–1476.

Nelson, C. R. and Kim, M. J. (1993), 'Predictable Stock Returns: The Role of Small Sample Bias', *Journal of Finance* **48**(2), 641–661.

Nelson, S. A. (1902), *The ABC of Stock Speculation*, New York: S.A. Nelson.

Nguyen, T. H., Shirai, K. and Velcin, J. (2015), 'Sentiment analysis on social media for stock movement prediction', *Expert Systems with Applications* **42**(24), 9603–9611.

Oliveira, N., Cortez, P. and Areal, N. (2017), 'The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices', *Expert Systems with Applications* **73**, 125–144.

Paye, B. S. (2012), ''Déjà vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables', *Journal of Financial Economics* **106**(3), 527–546.

Peng, L. and Xiong, W. (2006), 'Investor attention, overconfidence and category learning', *Journal of Financial Economics* **80**(3), 563–602.

Saunders, E. M. (1993), 'Stock Prices and Wall Street Weather', *American Economic Review* **83**(5), 1337–1345.

Sheng, J. (2019), 'Asset pricing in the information age: Employee expectations and stock returns', *Working Paper. University of California, Irvine. Available at SSRN 3321275* .

Shiller, R. J. and Pound, J. (1989), 'Survey evidence on diffusion of interest and information among investors', *Journal of Economic Behavior & Organization* **12**(1), 47–66.

Shleifer, A. and Summers, L. H. (1990), 'The noise trader approach to finance', *Journal of Economic perspectives* **4**(2), 19–33.

Shleifer, A. and Vishny, R. W. (1997), 'The limits of arbitrage', *Journal of Finance* **52**(1), 35–55.

Siganos, A., Vagenas-Nanos, E. and Verwijmeren, P. (2014), 'Facebook's daily sentiment and international stock markets', *Journal of Economic Behavior & Organization* **107**, 730–743.

Sloan, R. G. (1996), 'Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings?', *Accounting Review* pp. 289–315.

Stambaugh, R. F. (1999), 'Predictive regressions', *Journal of Financial Economics* **54**(3), 375–421.

Stambaugh, R. F., Yu, J. and Yuan, Y. (2012), 'The short of it: Investor sentiment and anomalies', *Journal of Financial Economics* **104**(2), 288–302.

Stamolampros, P., Korfiatis, N., Chalvatzis, K. and Buhalis, D. (2019), 'Job satisfaction and employee turnover determinants in high contact services: Insights from Employees' Online reviews', *Tourism Management* **75**, 130–147.

Symitsi, E., Stamolampros, P. and Daskalakis, G. (2018), 'Employees' online reviews and equity prices', *Economics Letters* **162**, 53–55.

Symitsi, E., Stamolampros, P., Daskalakis, G. and Korfiatis, N. (2021), 'The informational value of employee online reviews', *European Journal of Operational Research* **288**(2), 605 – 619.

Tang, V. W. (2018), 'Wisdom of Crowds: Cross-Sectional Variation in the Informativeness of Third-Party-Generated Product Information on Twitter', *Journal of Accounting Research* **56**(3), 989–1034.

Tetlock, P. C. (2007), 'Giving Content to Investor Sentiment: The Role of Media in the Stock Market', *Journal of Finance* **62**(3), 1139–1168.

Tetlock, P. C., Saar-Tsechansky, M. and Macskassy, S. (2008), 'More Than Words: Quantifying Language to Measure Firms' Fundamentals', *Journal of Finance* **63**(3), 1437–1467.

Titman, S., Wei, K. J. and Xie, F. (2004), 'Capital Investments and Stock Returns', *Journal of Financial and Quantitative Analysis* **39**(4), 677–700.

Wachter, J. A. (2013), 'Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?', *Journal of Finance* **68**(3), 987–1035.

Welch, I. and Goyal, A. (2008), 'A Comprehensive Look at the Empirical Performance of Equity Premium Prediction', *Review of Financial Studies* **21**(4), 1455–1508.

Weng, B., Lu, L., Wang, X., Megahed, F. M. and Martinez, W. (2018), 'Predicting short-term stock prices using ensemble methods and online data sources', *Expert Systems with Applications* **112**, 258–273.

Wright, W. F. and Bower, G. H. (1992), 'Mood effects on subjective probability assessment', *Organizational Behavior and Human Decision Processes* **52**(2), 276–291.

Yuan, K., Zheng, L. and Zhu, Q. (2006), 'Are investors moonstruck? Lunar phases and stock returns', *Journal of Empirical Finance* **13**(1), 1–23.

Online Appendix to "Employee Sentiment Index: Predicting Stock Returns with Online Employee Data"

A. Definition of economic variables

- *Dividend-to-price ratio (DP)*: defined as the difference between the logarithm of the 12-month moving sum of the dividends on the S&P 500 index and the logarithm of the S&P 500 index level.
- *Dividend yield (DY):* given by the difference between the logarithm of the dividends on the S&P 500 index and the logarithm of the lagged S&P 500 prices.
- *Earnings-to-price ratio (EP):* defined as the difference between the logarithm of the 12-month moving sum of the earnings on the S&P 500 index and the logarithm of the S&P 500 index level.
- *Book-to-market ratio (BM):* corresponds to the ratio of the book to market values of the Dow Jones Industrial Average.
- *Stock market variance (SVAR):* is the monthly stock market variance computed as the sum of intra–month squared daily returns on the S&P 500 index.
- *Net equity expansion (NTIS):* is the ratio of the 12-month moving sum of the net issues to the total end-of-year market capitalization of NYSE stocks.
- *Treasury bill rate (TBL):* is the 3–month treasury bill rate.
- *Long-term return (LTR):* is the return on long-term US government bonds.
- *Long-term yield (LTY):* is the yield on long-term US government bonds.
- *Default return spread (DFR):* is the spread between the long-term corporate and government bond returns.
- *Default yield spread (DFY):* is the spread between the Moody's Baa and Aaa bond yields.
- *Dividend payout ratio (DE):* is the difference between the log of dividends and the log of earnings.
- *Term spread (TMS):* is the difference between the yield on the long-term government bond and the treasury bill rate.
- *Inflation (INFL):* is the inflation rate computed as the growth rate of the Consumer Price Index.
- *Industrial production (GIP):* corresponds to the growth rate in the US industrial production.
- *Consumption of durables (GCDG):* is the growth rate in the consumption of durable goods.
- *Consumption of non-durables (GCNDG):* is the growth rate in the consumption of non-durable goods.
- *Consumption of services (GCS):* is the growth rate in consumer services.

B. Controlling for economic and sentiment factors

The multivariate predictive regression that controls for factors of economic and sentiment variables is specified as follows:

$$ER_{t+1:t+h} = \alpha + \beta\, \Delta ES_t + \sum_{j=1}^{n} \gamma_j\, F\_SENT_{jt} + \sum_{m=1}^{\omega} \delta_m\, F\_ECON_{mt} + \varepsilon_{t+1:t+h}, \tag{a}$$

where F_SENT_{jt} is the j^{th} principal component of the correlation matrix of the alternative sentiment proxies (i.e., UMCI, CBCI, PMI, and $BWSI$) and F_ECON_{mt} is the m^{th} principal

Table I. Correlation Matrix of Employee Sentiment, Sentiment Indexes and Economic Variables

Panel A: Sentiment Variables

	ES_O	ES_{SU}	ES_E	UMCI	CBCI	PMI
ES_{SU}	0.9983*					
ES_E	0.9738*	0.9729*				
UMCI	0.7987*	0.8155*	0.8508*			
CBCI	0.8692*	0.8839*	0.8554*	0.9084*		
PMI	0.3833*	0.4069*	0.2619	0.3871*	0.5962*	
BWSI	−0.5069*	−0.4812*	−0.4811*	−0.2024	−0.2697	0.1191

Panel B: Economic Variables

	ES_O	ES_{SU}	ES_E	DP	DY	EP	BM	SVAR	NTIS	TBL
DP	−0.2420	−0.2696	−0.1507							
DY	−0.2637	−0.2932	−0.1918	0.8616*						
EP	−0.8620*	−0.8622*	−0.9123*	0.1915	0.1864					
BM	−0.8009*	−0.8177*	−0.7465*	0.6495*	0.6045*	0.6707*				
SVAR	0.0020	0.0096	0.0779	0.3079*	0.0312	−0.0216	0.0972			
NTIS	−0.7176*	−0.7009*	−0.7494*	−0.3235*	−0.3103*	0.5975*	0.3461*	−0.0987		
TBL	0.7851*	0.8016*	0.6917*	−0.5451*	−0.5408*	−0.5226*	−0.8955*	−0.0364	−0.4001*	
LTR	−0.0380	−0.0405	−0.0408	0.0640	−0.0550	0.0317	0.1356	0.2437	0.1309	−0.0576
LTY	−0.3855*	−0.3685*	−0.4614*	−0.4063*	−0.3639*	0.3949*	−0.0338	−0.0939	0.7327*	−0.0078
DFS	0.0471	0.0367	0.0186	−0.0840	0.1908	−0.0331	−0.0717	−0.4284*	−0.1260	0.0343
DFY	−0.0876	−0.1086	0.0311	0.6901*	0.6477*	0.0653	0.3524*	0.3792*	−0.4360*	−0.2863
DE	0.7681*	0.7575*	0.8535*	0.1947	0.1464	−0.9254*	−0.4195*	0.1404	−0.7221*	0.3118*
TMS	−0.4655*	−0.4503*	−0.5314*	−0.3476*	−0.3053*	0.4482*	0.0598	−0.0907	0.7706*	−0.1105
INFL	0.1622	0.1567	0.1319	−0.1426	−0.1316	−0.1254	−0.1591	−0.0930	−0.0804	0.2070
GIP	0.0869	0.0900	−0.0084	−0.1948	−0.2489	0.0558	−0.1508	−0.0904	0.1236	0.2166
GCDG	−0.0179	−0.0246	−0.0303	−0.0372	−0.0491	0.0069	0.0201	−0.131	0.0276	0.0199
GCNDG	0.1102	0.1031	0.0784	−0.1116	−0.1715	−0.0714	−0.1464	−0.1427	−0.0311	0.1329
GCS	0.1412	0.1478	0.1351	−0.2180	−0.2089	−0.2079	−0.148	−0.0641	0.0526	0.1311

	LTR	LTY	DFR	DFY	DE	TMS	INFL	GIP	GCDG	GCNDG
DP										
DY										
EP										
BM										
SVAR										
NTIS										
TBL										
LTR										
LTY	−0.1963									
DFS	−0.4725*	−0.0129								
DFY	0.0561	−0.4635*	0.0611							
DE	−0.0070	−0.5515*	0.0006	0.2012						
TMS	−0.1882	0.9947*	−0.0162	−0.4315*	−0.5821*					
INFL	−0.0161	0.0871	−0.0050	−0.2160	0.0703	0.0654				
GIP	0.0454	0.2406	−0.1568	−0.3663*	−0.1310	0.2164	0.1706			
GCDG	−0.1547	−0.0057	0.0882	−0.0844	−0.0212	−0.0072	0.0567	0.0649		
GCNDG	0.0429	0.1247	−0.1759	−0.1882	0.0282	0.1103	0.5162*	0.2940	0.3232*	
GCS	0.1166	0.0406	−0.0755	−0.2053	0.1236	0.0280	−0.1857	0.0424	0.0483	−0.0273

Note: * $p < 0.05$ denotes the level of significance.

Table II

Employee Sentiment and Excess Stock Market Return Predictability - Controlling for Investor Sentiment and Economic Variables

	(1)	(2)	(3)	(4)	(5)
Horizon	t+1	t+3	t+6	t+9	t+12
ES	−0.0061**	−0.0089**	−0.0076*	−0.0189**	−0.0120*
	(0.0035)	(0.0051)	(0.0058)	(0.0086)	(0.0088)
F_$SENT1$	−0.0053	−0.0018	0.0050	0.0012	−0.0016
	(0.0044)	(0.0056)	(0.0063)	(0.0069)	(0.0083)
F_$SENT2$	0.0002	0.0005	0.0062	0.0019	0.0068
	(0.0030)	(0.0049)	(0.0059)	(0.0072)	(0.0078)
F_$ECON1$	0.0010	0.0043	0.0062*	0.0034	−0.0002
	(0.0030)	(0.0041)	(0.0045)	(0.0050)	(0.0057)
F_$ECON2$	0.0026	0.0070*	0.0188***	0.0291***	0.0409***
	(0.0030)	(0.0044)	(0.0057)	(0.0058)	(0.0068)
F_$ECON3$	0.0037	0.0073**	0.0030	−0.0080	−0.0045
	(0.0039)	(0.0044)	(0.0066)	(0.0065)	(0.0087)
F_$ECON4$	0.0012	0.0075*	0.0000	0.0080	0.0137*
	(0.0034)	(0.0053)	(0.0077)	(0.0074)	(0.0103)
F_$ECON5$	0.0024	0.0013	−0.0034	−0.0078	−0.0182***
	(0.0033)	(0.0048)	(0.0048)	(0.0065)	(0.0074)
Constant	0.0098***	0.0277***	0.0521***	0.0762***	0.0983***
	(0.0032)	(0.0046)	(0.0057)	(0.0064)	(0.0071)
Obs.	73	73	73	73	73
ΔR^2	0.0446	0.0452	0.0194	0.0793	0.0218

This table presents the results from the regression of the h-month ahead cumulative excess return of the S&P 500 index on a constant, the standardized changes in *ES* of month t and the month t value of common factors (principal components) of other sentiment proxies (F_SENT_j, j=1,2) and of economic variables (F_ECON_m, m=1,2,...5). We consider forecast horizons h = 1, 3, 6, 9, and 12 months, respectively. The sample period is from June 2012 to July 2018. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively, using a one-side test.

component of the correlation matrix of the economic variables. Changes in the *ES* maintain a negative and significant loading across all forecast horizons. The incremental explanatory power of the *ES* peaks at the 9-month horizon. With the exception of the second principal component of the economic variables, most of the slope estimates of the common factors are insignificant.

C. Alternative market portfolios

We next examine whether the evidence of predictability of stock market returns by aggregated employee outlook predictions is robust to alternative proxies of the market portfolio. We use the returns on the following assets: (i) the *Russell 3000* index, (ii) the *CRSP* value–weighted index (FF Mkt) consisting of all US firms listed on *NYSE*, *AMEX*, and *Nasdaq*, (iii) the *S&P 500 E-Mini* futures, and (iv) the *SPDR S&P 500 ETF*. The *CRSP* value–weighted index is taken from *CRSP*, while data on the other three series are collected from *Thomson Reuters Datastream*.

Panels A–C of Table III show that the *ES* strongly predicts a subsequent reversal in excess stock market returns, in line with our previous evidence. All coefficients are negative and statistically significant at the 5% level and their impact is much stronger at longer horizons (e.g., 9 or 12 months). Furthermore, the slope estimates have a very similar magnitude across the different portfolios. These results remain qualitatively unchanged if we control for economic and sentiment variables (we do not report these results to conserve space). Overall, the above analysis clearly suggests that the strong predictability of the *ES* is robust to alternative market portfolio proxies.

D. Value-weighted characteristic portfolios

Table IV examines value-weighted portfolios investigating whether the effects of market-wide employee sentiment are economically significant. We find that the power of employee sentiment to predict longer-term returns is stronger. Significant evidence in predicting returns is found for portfolios sorted on size, risk and profitability across all the tested horizons.

Table III

Employee Sentiment and Excess Stock Market Returns - Alternative Market Portfolio Proxies (dependent variables)

	(1)	(2)	(3)	(4)	(5)
Horizon	t+1	t+3	t+6	t+9	t+12
Panel A: Alternative Stock Market Indexes					
RUSSELL3000	−0.0058**	−0.0108***	−0.0110**	−0.0260***	−0.0219**
	(0.0029)	(0.0044)	(0.0053)	(0.0086)	(0.0101)
CRSP	−0.0058**	−0.0108***	−0.0109**	−0.0245***	−0.0203**
	(0.0029)	(0.0044)	(0.0049)	(0.0078)	(0.0092)
Panel B: Futures					
EMINI	−0.0061**	−0.0110***	−0.0112**	−0.0249***	−0.0207**
	(0.0031)	(0.0047)	(0.0052)	(0.0085)	(0.0102)
Panel C: Exchange Traded Funds					
SPY	−0.0059**	−0.0109***	−0.0110**	−0.0246***	−0.0207**
	(0.0030)	(0.0046)	(0.0051)	(0.0084)	(0.0101)
Observations	73	73	73	73	73

This table presents coefficient estimates along with their associated bootstrapped robust standard errors for regressions of an asset's h-month cumulative excess return (asset names in rows) on the lagged value of the employee sentiment index (ES). Excess returns are predicted over horizons of 1, 3, 6, 9 and 12 months, respectively. The following market proxies are considered: the RUSSELL 3000 index, the value-weighted index of all US stocks as in Fama-French (FF Mkt), the S&P 500 E-Mini futures (E–MINI) and the SPDR S&P 500 ETF (SPY ETF). ***, **, and * denote statistical significance at the 1%, 5% and 10% level using a one-sided test.

Predicting industry portfolios

In this analyses we employ industry returns obtained from the webpage of Ken French. The industries covered are consumer non-durables, consumer durables, manufacturing, energy, high technology, telecommunications, wholesale and retail shops, healthcare and utilities. We exclude the 10th portfolio that classifies firms to "Other" as is not interpretable in a way consistent with the dominant business activity. We estimate the following regression:

$$R_{t+1:t+h} = \alpha + \beta \Delta ES_t + \varepsilon_{t+1:t+h}, \tag{b}$$

where $R_{t+1:t+h}$ are the cumulative h-month ahead returns (h=1,3,6,9,12) of each industry portfolio formed both equally- and value-weighted.

In line with the results for stock market returns, we see that the predictability is stronger for the longer horizons of 9 and 12 months. Looking across the different industries, we find that consumer durables and manufacture exhibit the highest slope estimates that are also economically meaningful for both equal-weighted and value-weighted portfolio returns. The portfolios of high technology, energy, telecommunications, and shops are generally predicted by employee sentiment. The effect in value-weighted portfolios is somewhat reduced, but remains significant. The returns on utilities portfolios are mostly insensitive to changes in aggregate market sentiment. The effect of aggregate employee sentiment on the non-durables portfolio returns is highly significant, but becomes insignificant once portfolios are constructed by weighing the stocks based on their market capitalization.

Table IV

Employee Sentiment and Characteristic-sorted Portfolio Return Predictability (Value–Weighted)

	Dependent Variable: Characteristic Portfolio Returns (Equal–Weighted)														
Horizon	t+1			t+3			t+6			t+9			t+12		
Decile	(1)	(5)	(10)	(1)	(5)	(10)	(1)	(5)	(10)	(1)	(5)	(10)	(1)	(5)	(10)
Panel A: Size and Age															
ME	−0.0036	−0.0075**	−0.0060**	−0.0110*	−0.0128**	−0.0127***	−0.0124	−0.0098	−0.0114***	−0.0392***	−0.0352***	−0.0244***	−0.0346**	−0.0276**	−0.0205**
	(0.0045)	(0.0040)	(0.0031)	(0.0068)	(0.0058)	(0.0044)	(0.0116)	(0.0095)	(0.0047)	(0.0160)	(0.0130)	(0.0077)	(0.0186)	(0.0148)	(0.0093)
Age	−0.0055*	−0.0067*	−0.0041	−0.0071	−0.0090*	−0.0092**	−0.0152*	−0.0067	−0.0113**	−0.0356***	−0.0216***	−0.0233***	−0.0361**	−0.0161*	−0.0166**
	(0.0040)	(0.0041)	(0.0033)	(0.0080)	(0.0056)	(0.0053)	(0.0114)	(0.0066)	(0.0052)	(0.0150)	(0.0092)	(0.0075)	(0.0171)	(0.0114)	(0.0085)
Panel B: Risk															
Beta	−0.0027	−0.0044	−0.0148***	−0.0083**	−0.0132**	−0.0224**	−0.0068*	−0.0090*	−0.0312**	−0.0125***	−0.0272***	−0.0654***	−0.0060	−0.0211**	−0.0636***
	(0.0030)	(0.0041)	(0.0054)	(0.0041)	(0.0063)	(0.0104)	(0.0048)	(0.0057)	(0.0162)	(0.0051)	(0.0094)	(0.0242)	(0.0053)	(0.0109)	(0.0267)
V	−0.0045*	−0.0068*	−0.0074*	−0.0036	−0.0115**	−0.0149*	−0.0101**	−0.0138**	−0.0271**	−0.0180***	−0.0322***	−0.0637***	−0.0145**	−0.0251**	−0.0489**
	(0.0033)	(0.0044)	(0.0055)	(0.0032)	(0.0052)	(0.0095)	(0.0054)	(0.0066)	(0.0154)	(0.0064)	(0.0101)	(0.0234)	(0.0069)	(0.0116)	(0.0224)
Panel C: Profitability and Dividend Payout															
E/P	−0.0035	−0.0054**	−0.0075**	−0.0137***	−0.0124***	−0.0137**	−0.0068	−0.0139**	−0.0199**	−0.0265***	−0.0267***	−0.0387***	−0.0292***	−0.0216**	−0.0324**
	(0.0031)	(0.0031)	(0.0043)	(0.0054)	(0.0048)	(0.0072)	(0.0078)	(0.0061)	(0.0090)	(0.0101)	(0.0084)	(0.0140)	(0.0116)	(0.0099)	(0.0150)
AC	−0.0061*	−0.0068**	−0.0091***	−0.0153***	−0.0121**	−0.0114*	−0.0110*	−0.0132**	−0.0230***	−0.0300***	−0.0290***	−0.0387***	−0.0253**	−0.0211**	−0.0404***
	(0.0038)	(0.0034)	(0.0033)	(0.0063)	(0.0058)	(0.0077)	(0.0073)	(0.0066)	(0.0082)	(0.0105)	(0.0089)	(0.0127)	(0.0116)	(0.0104)	(0.0149)
OP	−0.0077*	−0.0061*	−0.0051*	−0.0123*	−0.0125***	−0.0141**	−0.0188*	−0.0127**	−0.0094**	−0.0501***	−0.0298***	−0.0221***	−0.0477***	−0.0258**	−0.0179**
	(0.0051)	(0.0038)	(0.0031)	(0.0082)	(0.0044)	(0.0061)	(0.0126)	(0.0066)	(0.0055)	(0.0189)	(0.0108)	(0.0075)	(0.0197)	(0.0123)	(0.0086)
DY	−0.0107***	−0.0044*	−0.0040	−0.0144***	−0.0130***	−0.0107**	−0.0191**	−0.0123**	−0.0104*	−0.0386***	−0.0289***	−0.0170*	−0.0368***	−0.0222***	−0.0171*
	(0.0038)	(0.0033)	(0.0034)	(0.0057)	(0.0046)	(0.0056)	(0.0087)	(0.0057)	(0.0070)	(0.0127)	(0.0089)	(0.0104)	(0.0144)	(0.0096)	(0.0115)
Panel D: Tangibility															
PPE	0.0044	−0.0001	0.0023	0.0024	−0.0012	−0.0088	−0.0132*	−0.0133**	−0.0115	−0.0222**	−0.0180**	−0.0237*	−0.0309**	−0.0229***	−0.0312**
	(0.0046)	(0.0033)	(0.0043)	(0.0070)	(0.0060)	(0.0094)	(0.0096)	(0.0061)	(0.0116)	(0.0128)	(0.0078)	(0.0147)	(0.0152)	(0.0087)	(0.0147)
RD	0.0012	0.0026	0.0093	−0.0070*	−0.0039	0.0083	−0.0086	−0.0009	0.0056	−0.0158**	−0.0128	−0.0072	−0.0209**	−0.0216*	−0.0270*
	(0.0030)	(0.0059)	(0.0096)	(0.0051)	(0.0050)	(0.0112)	(0.0075)	(0.0134)	(0.0135)	(0.0091)	(0.0161)	(0.0157)	(0.0113)	(0.0168)	(0.0182)
Panel E: Financial Distress/Growth Opportunities															
BM	−0.0032	−0.0070**	−0.0110**	−0.0114**	−0.0142***	−0.0149**	−0.0042	−0.0131**	−0.0254**	−0.0191***	−0.0317***	−0.0477***	−0.0167**	−0.0238**	−0.0330*
	(0.0030)	(0.0035)	(0.0061)	(0.0067)	(0.0054)	(0.0077)	(0.0056)	(0.0065)	(0.0119)	(0.0072)	(0.0102)	(0.0189)	(0.0081)	(0.0119)	(0.0222)
I	−0.0074**	−0.0042	−0.0068**	−0.0145**	−0.0130***	−0.0090	−0.0126**	−0.0177***	−0.0034	−0.0263***	−0.0343***	−0.0296**	−0.0230**	−0.0277**	−0.0296**
	(0.0039)	(0.0035)	(0.0037)	(0.0087)	(0.0046)	(0.0073)	(0.0074)	(0.0071)	(0.0100)	(0.0102)	(0.0107)	(0.0132)	(0.0120)	(0.0123)	(0.0131)
NSI	−0.0055*	−0.0060**	−0.0033	−0.0150***	−0.0167***	−0.0067	−0.0187***	−0.0083	−0.0020	−0.0344***	−0.0279***	−0.0162	−0.0282**	−0.0229**	−0.0155
	(0.0036)	(0.0034)	(0.0045)	(0.0058)	(0.0059)	(0.0089)	(0.0072)	(0.0080)	(0.0109)	(0.0117)	(0.0107)	(0.0153)	(0.0137)	(0.0114)	(0.0185)
SG	0.0016	0.0025	0.0064	0.0008	−0.0025	−0.0016	−0.0157*	−0.0133**	−0.0015	−0.0315***	−0.0174***	−0.0153	−0.0385***	−0.0188**	−0.0388***
	(0.0042)	(0.0032)	(0.0064)	(0.0081)	(0.0051)	(0.0070)	(0.0101)	(0.0067)	(0.0131)	(0.0123)	(0.0074)	(0.0155)	(0.0121)	(0.0093)	(0.0166)
EF	0.0026	0.0006	0.0035	−0.0025	−0.0023	0.0052	−0.0128**	−0.0148**	−0.0106	−0.0190***	−0.0233***	−0.0213*	−0.0233***	−0.0264***	−0.0413***
	(0.0033)	(0.0033)	(0.0048)	(0.0062)	(0.0048)	(0.0084)	(0.0067)	(0.0064)	(0.0114)	(0.0080)	(0.0083)	(0.0141)	(0.0084)	(0.0098)	(0.0143)
L	0.0014	0.0015	0.0002	−0.0012	−0.0029	−0.0057	−0.0140**	−0.0129**	−0.0148***	−0.0236***	−0.0212**	−0.0209***	−0.0328***	−0.0272***	−0.0284***
	(0.0030)	(0.0029)	(0.0037)	(0.0106)	(0.0050)	(0.0061)	(0.0082)	(0.0068)	(0.0063)	(0.0100)	(0.0092)	(0.0076)	(0.0107)	(0.0101)	(0.0088)

Note: This table presents the slope estimates from OLS regressions of 1, 3, 6, 9 and 12–month ahead value-weighted returns on characteristic–sorted portfolios on the lagged value of the *ES*. Bootstrap robust standard errors are reported in parentheses. 1, 5 and 10 decile portfolios are formed from single sorts on market equity (ME), age, beta, variance (V), earnings (E/P), accruals (AC), operating profit (OP), dividend yield (DY), fixed over total assets (PPE), research and development over total assets (RD), book-to-market equity ratio (BM), investment (I), net share issues (NSI), sales growth (SG), external finance over total assets (EF) and market leverage. ***, **, and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table V

Employee Sentiment and Industry Portfolios

Industry	Panel A: Equal-Weighted					Panel B: Value-Weighted				
	t+1	t+3	t+6	t+9	t+12	t+1	t+3	t+6	t+9	t+12
NonDurables	−0.0067**	−0.0087*	−0.0134**	−0.0285***	−0.0291***	−0.0021	−0.0038	−0.0046	−0.0071	−0.0039
	(0.0036)	(0.0053)	(0.0078)	(0.0103)	(0.0121)	(0.0038)	(0.0068)	(0.0064)	(0.0073)	(0.0095)
Durables	−0.0092**	−0.0273***	−0.0229**	−0.0488***	−0.0389**	−0.0114***	−0.0186**	−0.0198**	−0.0458***	−0.0330*
	(0.0050)	(0.0084)	(0.0133)	(0.0183)	(0.0224)	(0.0047)	(0.0090)	(0.0117)	(0.0166)	(0.0202)
Manufacture	−0.0099**	−0.0182**	−0.0164	−0.0479***	−0.0376**	−0.0068**	−0.0157***	−0.0166**	−0.0373***	−0.0280**
	(0.0044)	(0.0078)	(0.0131)	(0.0185)	(0.0211)	(0.0034)	(0.0063)	(0.0080)	(0.0120)	(0.0142)
Energy	−0.0212***	−0.0339**	−0.0109	−0.0599*	−0.0494	−0.0141***	−0.0221**	−0.0090	−0.0407***	−0.0284*
	(0.0074)	(0.0186)	(0.0266)	(0.0383)	(0.0390)	(0.0053)	(0.0100)	(0.0128)	(0.0172)	(0.0192)
HiTech	−0.0040	−0.0143**	−0.0102	−0.0387***	−0.0336**	−0.0068**	−0.0131***	−0.0062	−0.0231**	−0.0228**
	(0.0043)	(0.0068)	(0.0104)	(0.0143)	(0.0158)	(0.0037)	(0.0053)	(0.0074)	(0.0105)	(0.0125)
Telecom	−0.0062*	−0.0177**	−0.0142	−0.0294**	−0.0329**	−0.0056	−0.0100*	−0.0121**	−0.0178**	−0.0201**
	(0.0047)	(0.0078)	(0.0125)	(0.0166)	(0.0173)	(0.0045)	(0.0075)	(0.0072)	(0.0094)	(0.0099)
Shops	−0.0059*	−0.0109**	−0.0172**	−0.0386***	−0.0348***	−0.0054*	−0.0061	−0.0127**	−0.0261***	−0.0241***
	(0.0039)	(0.0058)	(0.0104)	(0.0133)	(0.0146)	(0.0036)	(0.0057)	(0.0062)	(0.0069)	(0.0079)
Healthcare	−0.0001	−0.0012	−0.0064	−0.0412**	−0.0383**	−0.0033	−0.0054	−0.0120*	−0.0277***	−0.0275***
	(0.0058)	(0.0091)	(0.0149)	(0.0200)	(0.0211)	(0.0039)	(0.0068)	(0.0076)	(0.0109)	(0.0115)
Utilities	−0.0038	−0.0031	−0.0067	−0.0142*	−0.0025	−0.0028	−0.0024	−0.0084	−0.0155*	−0.0097
	(0.0043)	(0.0055)	(0.0091)	(0.0091)	(0.0120)	(0.0042)	(0.0059)	(0.0093)	(0.0102)	(0.0132)

This table presents the slope estimates from OLS regressions of the h–month ahead returns (h=1,3,6,9,12) on the industry portfolios on the lagged value of the *ES*. Bootstrap robust standard errors are reported in parentheses. Panel A contains the results using equal-weighted returns, whereas Panel B shows the results from value-weighted returns. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, using a one-sided test.