A picture is worth a thousand words:

Measuring investor sentiment by combining machine learning and photos from news*

Khaled Obaid

Kuntara Pukthuanthong†

Forthcoming, *Journal of Financial Economics*

ABSTRACT

By applying machine learning to the accurate and cost-effective classification of photos based on sentiment, we introduce a daily market-level investor sentiment index (Photo Pessimism) obtained from a large sample of news photos. Consistent with behavioral models, Photo Pessimism predicts market return reversals and trading volume. The relation is strongest among stocks with high limits to arbitrage and during periods of elevated fear. We examine whether Photo Pessimism and pessimism embedded in news text act as complements or substitutes for each other in predicting stock returns and find evidence that the two are substitutes.

Keywords: Investor sentiment; Behavioral finance; Return predictability; Machine learning; Deep learning; Big data

JEL classification: C53; G10; G17

---

* We thank David Hirshleifer (the editor), and the referee for their helpful comments. We particularly want to thank Richard Roll for his time and suggestions and also Dominique Badoer, Dave Berger, John Howe, Xing Huang, Inder Khurana, Heikki Lehkonen, Venky Nagar, Michael O'Doherty, John Pinfold, and Martin Stefan and seminar participants at the University of Missouri-Columbia, the University of Missouri-St. Louis, Missouri State University, Saint Louis University, California State University-East Bay, the Midwest Finance Association Meeting (2019), the Southern Finance Association Meeting (2019), the GSU-RFS FinTech Conference (2020), and the Financial Management Association Annual Meeting (2020) for helpful feedback. We are grateful for financial support from the Hillcrest Behavior Finance Award, Winemiller Excellence Award, and the University of Missouri-Columbia and California State University-East Bay finance departments. An earlier version of this paper was circulated under the title "A Picture is Worth a Thousand Words: Market Sentiment from Photos." The data are available at: https://www.kuntara.net/.

† Khaled Obaid is an Assistant Professor at the College of Business and Economics, California State University-East Bay, Hayward, CA 94542; email khaled.obaid@csueastbay.edu; tel. (510) 885-2064. Kuntara Pukthuanthong is the Robert J. Trulaske Jr. Professor of Finance at the Robert J. Trulaske Sr. College of Business, University of Missouri, Columbia, MO 65211; email pukthuanthongk@missouri.edu; tel. (573) 884-9785.

A good sketch is better than a long speech.

—— Napoleon Bonaparte

1. Introduction

Numerous studies document how investor sentiment helps researchers understand and predict market returns over time (Hirshleifer and Shumway, 2003; Edmans et al., 2007; Tetlock, 2007; Spiegel, 2008; Cochrane, 2011) and stock returns cross-sectionally (Baker and Wurgler, 2006; Kozak et al., 2018). In this study, we develop a daily market-level investor sentiment index (Photo Pessimism) from US news photos and study how visual content in news relates to financial markets.

We make three key contributions to the literature. First, we demonstrate the importance of visual content in making predictions about market returns. We construct a daily investor sentiment index, Photo Pessimism (*PhotoPes*), calculated as the proportion of news photos predicted to be negative on a given day. We observe that *PhotoPes* negatively predicts the next day's market returns and is positively related to market returns over the remaining trading week. This reversal highlights that our measure has a noninformational impact on returns. Consistent with an investor sentiment proxy, we show that *PhotoPes* predicts an increase in the next day's trading volume and better predicts returns on stocks with high compared to low limits to arbitrage.

Second, we demonstrate how to overcome key hurdles of studying the importance of visual content in financial markets by employing machine learning techniques for large-scale photo classification. Photojournalism has increased in popularity due to modern technology and the demand for quick information. In light of studies suggesting that photos may convey emotional information more effectively than words, we believe it is important to examine how sentiment extracted from photos in the news media relates to market activities (Chemtob et al., 1999). However, because of the complexity of analyzing photos and the cost of manually sifting through many photos, conducting such a study is expensive, error prone, and tedious. Relying on surveys or crowd-sourcing websites (e.g., Amazon Mechanical Turk, MTurk) to evaluate photos has been a mainstream method for extracting information from photos. In addition to the high cost, economists are cautious about survey data as they can be subjective and not verifiable by "objective external measurement" (Vissing-Jorgensen, 2003). Singer (2002), for instance, documents that survey respondents tend to have less

incentive to answer questions carefully and truthfully when the questions are related to a sensitive subject and answers depend on respondents' perception. Our approach of using machine learning mitigates this concern as sentiment embedded in photos is uniformly extracted by a machine. To be specific, we apply convolutional neural networks (CNNs), a machine learning technique popular for classifying photos, to accurately, verifiably, and cost effectively classify a large sample of news photos based on sentiment.

Third, we compare the predictive ability of *PhotoPes* and pessimism embedded in news text. We show that pessimism embedded in news photos and pessimism embedded in news text substitute for one another. Moreover, we provide evidence suggesting that photos grab attention away from text during periods when photos are salient. Specifically, during periods when news photos are overwhelmingly negative or positive, the pessimism in news photos dominates. In contrast, during periods when news photos are neutral or mixed, the pessimism embedded in the text dominates. Tetlock (2007) and Garcia (2013), for example, measure investor sentiment by performing textual analysis on news; however, to the best of our knowledge, no one has examined the possibility of capturing useful and novel information about investors' beliefs from photos in news and how such information interacts with the information embedded in text.

Finance researchers have embraced the use of machine learning (Mullainathan and Spiess, 2017). Machine learning techniques are often used to classify textual content. For instance, Manela and Moreira (2017) use machine learning to construct a news implied volatility index from text in the Wall Street Journal (WSJ). Buehlmaier and Whited (2018) construct a measure of financial constraints by analyzing financial reports and show that their measure is related to access to capital and stock returns. Lately, asset pricing researchers have applied machine learning to financial data to predict risk premiums (Gu et al., 2020) and to search for true risk factors (Feng et al., 2020).

Recent developments in machine learning have introduced techniques that make the task of analyzing a large number of photos possible. In this study, we apply CNNs to construct an investor sentiment index from a large sample of news photos. In particular, we use a pretrained Google Inception (v3) model. Although the model is not specifically trained to identify sentiment, it contains a lot of domain knowledge on images. To tailor the model to classify the sentiment of photos, we use transfer learning (i.e., we feed the pretrained model

a sample of additional training photos specifically labeled by sentiment and replace the final fully connected layer of the original model with a new layer containing only the two classes of interest: negative and positive sentiments). The model considers many aspects of the photo, including objects, colors, and facial expressions, to make predictions. After the model is trained, we verify its accuracy against a test sample. We achieve 87.1% accuracy with our test sample. Next, we use the model to make investor sentiment predictions on photos from the WSJ. Using these sentiment predictions, we then construct a daily investor sentiment index, *Photo Pessimism* (*PhotoPes*), calculated as the proportion of photos predicted to be negative on a given day.

The photos we use in this study come from the WSJ Online Archive between September 2008 and September 2020. Photos from this source allow us to focus on the most widely distributed news, which is the type of content that Shiller (2005), Tetlock (2007), and Garcia (2013) argue can play an important role in financial markets.[1]

Through various tests, we show that *PhotoPes* exhibits characteristics of an investor sentiment proxy. First, we test how *PhotoPes* is related to major US equity indices and exchange-traded funds (ETFs). Behavioral models, such as De Long et al. (1990), expect that investor sentiment should predict market return reversals. That is, when sentiment is high (low), irrational investors will increase (decrease) their demand for assets driving up (down) prices away from fundamentals. Because of limits to arbitrage, the mispricing might not be corrected immediately (Pontiff, 1996; Shleifer and Vishny, 1997). However, over time, rational investors will take advantage of mispricing, leading prices to return to their fundamental levels. We observe that *PhotoPes* is negatively related to the next day's market returns and positively related to market returns over the remaining trading week. This reversal highlights that our measure has a noninformational impact on returns. In terms of magnitude for the relation between *PhotoPes* and the next day's market returns, the average impact of a one standard deviation shift in *PhotoPes* on the CRSP value-weighted (VWRETD) index is 4.2 bps. A trading strategy that takes advantage of the reversal pattern for *PhotoPes* and pessimism embedded in news text earns a 5.45% annual five-factor alpha. Overall, the results are statistically and economically meaningful.

[1] We collect photos from Getty Images between 1926 and 2018 and test how *PhotoPes* relates to market returns over long-horizons. We find our main premise holds. We report the results in the Internet Appendix and describe caveats.

Second, we show that *PhotoPes* and the pessimism embedded in text (*TextPes*) are significantly correlated, suggesting some commonalities in the type of information both variables capture. Next, we examine how pessimism embedded in photos and text interact. We show that pessimism embedded in news photos and pessimism embedded in news text substitute for one another. In addition, our evidence suggests that photos capture attention from text during periods when photos are salient.

Third, we strive to enhance our understanding of which type of information is more effectively transmitted by photos in the context of news and financial markets. We find that the coefficient for pessimism embedded in photos is over two times larger during periods of elevated fear compared to baseline periods, while the coefficient for *TextPes* is roughly unchanged. This evidence is consistent with photos being more effective at conveying traumatic news compared to text (Chemtob et al., 1999). The impact of *PhotoPes* on market returns during high levels of fear is most striking: the average impact of a one standard deviation shift in *PhotoPes* during fear periods on the VWRETD is 10.3 bps.

Fourth, we further validate *PhotoPes* as a proxy for investor sentiment by showing it has a larger effect on stocks that are difficult to arbitrage. Based on the prediction that stocks that are expensive to arbitrage are most sensitive to investor sentiment shocks (Baker and Wurgler, 2006), we construct portfolios based on firm idiosyncratic volatility and size. We find that *PhotoPes* has the strongest effect on the returns of the highest idiosyncratic volatility and the smallest firm portfolios. In terms of magnitude, the average impact of a one standard deviation shift in *PhotoPes* on the highest idiosyncratic volatility value-weighted quintile portfolio is 7.1 bps.

Fifth, we provide additional insights to help us better understand the channel through which *PhotoPes* relates to market returns. We find that high or low *PhotoPes* can predict an increase in the next day's abnormal trading volume. This evidence is further validation that *PhotoPes* captures investor sentiment (De Long et al., 1990).

Finally, we perform a battery of robustness tests. Most notably, we show that that our main results are robust to different variable construction criteria and even after controlling for extreme returns. In addition, we show that *PhotoPes* generates a significant out-of-sample *R*-squared.

Our paper is related to the literature on investor sentiment. In light of the multidimensionality of investor sentiment, researchers have been looking for different approaches to measure investor sentiment (Zhou, 2018). For example, researchers have used news (Tetlock, 2007), Google Search data (Da et al., 2015), Twitter data (Chen et al., 2014), company financial reports (Loughran and McDonald, 2011; Jiang et al., 2019), weather (Hirshleifer and Shumway, 2003), and sporting events (Edmans et al., 2007) to proxy for investor sentiment (Hirshleifer, 2001).

More specifically, our paper extends the literature on investor sentiment and news. News is a plausible proxy for investors' beliefs because the press has demand-side incentive to cater content to their readers' beliefs (Shiller, 2005). Mullainathan and Shleifer (2005) summarize the communications, psychology, memory, and information processing literature that supports the notion that people receive utility from content consistent with their beliefs. Gentzkow and Shapiro (2010) provide empirical evidence for this theory. Specifically, they show that media slant is largely attributed to consumer preference. Tetlock (2007) and Garcia (2013) show that sentiment embedded in news text predicts market returns and trading volume. We extend this literature by revealing that news photos contain content relevant to financial markets.

Our paper also extends the literature on the psychology of visual stimuli to an application in finance and economics. The evidence on visual stimuli in news is mixed. On the one hand the picture superiority effect, a known concept in psychology, is the phenomenon in which pictures are better recalled than words (Nelson et al., 1976; Paivio, 1991). Newhagen and Reeves (1992) extend the picture superiority effect to news media by finding that visual content in news is better recalled than text content in news. Moreover, Garcia and Stark (1991) and Powell et al. (2015) find that images play an attention-grabbing role in news. On the other hand, researchers find that visual content is most effective when the message is simple to understand, whereas text is most effective when the message is complex (Chaiken and Eagly, 1976). We reconcile these two strands of literature in the context of financial decision-making.

Finally, our paper extends the literature on the value of visual content in predicting important outcomes in financial markets. Several studies document how a mere photo is able to predict important outcomes, such as political elections, personal loan decisions, firm market value, and CEO compensation (Todorov et al., 2005;

Duarte et al., 2012; Halford and Hsu, 2014; Graham et al., 2016). Recently, Bazley et al. (2021) document how displaying financial information in red reduces investors' appetite for risk and optimism. Blankespoor et al. (2017) show how perceptions about management in video presentations relate to firm value. To the best of our knowledge, we are the first to use machine learning to develop an investor sentiment proxy from news photos to predict market returns.

The rest of the paper is structured as follows. Section 2 describes the photo classification model, data, and key variables. Section 3 presents the empirical results. Section 4 concludes the paper.

2. Data

We begin this section by discussing the technology used to construct *PhotoPes*. Second, we discuss the sample of photos from the WSJ. Third, we review the descriptive statistics for *PhotoPes* and other key variables.

2.1. *Photo classification*

Major advances in the field of computer vision enable us to create reliable models for photo classification. Below, we delve into some of the details about the sentiment prediction model we utilized in this study.

The main task of machine learning photo classification models is to be able to identify the content of the photo with minimal human involvement. CNNs are a type of deep neural network that is useful for photo classification (Krizhevsky et al., 2012). Recent studies use CNN photo classification models to identify solar panel installations (Yu et al., 2019), locations of slave camps (Scoles, 2019), and poverty levels in underdeveloped countries (Jean et al., 2016) by analyzing satellite images. Similar to us, You et al. (2015) employ CNNs for image sentiment analysis and achieve high accuracy on photos in social media.

For this study, we are interested in a model that is able to predict sentiment that a photo is likely to elicit from investors. We build a photo classification model based on Google Inception (v3) (Szegedy et al., 2016). Google Inception (v3) is a CNN model that performs very well at classifying photos across 1,000

categories in the ImageNet academic competition, ILSVRC (ImageNet Large-Scale Visual Recognition Challenge), and is widely used in practice and research.[2,3]

TensorFlow, an open-source software library popular for machine learning applications developed by the Google Brain Team, provides a pretrained Google Inception (v3) model. We start with a pretrained Google Inception (v3) model (trained on the ImageNet data set) and use transfer learning to fine-tune the model for our specific application. The pretrained Google Inception (v3) has 1,000 different classes since it is trained for the purpose of ILSVRC. Training photo classification models from scratch requires a huge training sample and is computationally expensive. For example, the pretrained Google Inception (v3) model is trained on 1,331,167 prelabeled images. Transfer learning allows us to reuse the domain knowledge stored in the pretrained model to ease the construction of our desired model by replacing only the final fully connected layer with a new layer that has the desired number of classes (Yang et al., 2013) and retraining the parameters in the final fully connected layer with a much smaller training sample. The output we get from the fine-tuned model is one of two probabilities: the probability that photos have positive sentiment and the probability that photos have negative sentiment.

For the following reasons, we do not develop a more refined classification model. First, having more refined classifications requires subjective judgement. Most textual sentiment analyses, such as in Loughran and McDonald (2011), generate binary classifications for words (e.g., positive and negative words).[4] Second, we attempt to create a photo classification model with finer classes, but we are not able to get such a model to converge. This is because the model could not detect a clear distinction between the various finer classes.

To perform transfer learning, we need a training set that consists of photos labeled by sentiment. We use the DeepSent data set for training.[5] DeepSent is a collection of photos that are collected and labeled by

[2] The top-five and top-one error rates on the validation set from the ImageNet data set are 3.5% and 17.3%, respectively. The top-one error rate is the percentage of times the model did not produce the correct class as its top prediction by probability. The top-five error rate is the percentage of times the model did not produce the correct class as its top-five predictions by probability.

[3] The ImageNet data set is available for free download at: http://image-net.org/download.

[4] "Disease" and "crash," for example, are clearly negative, but the degree of their negativity is less obvious. Context and an individual's subjective interpretation of the word itself also affect the degree of negative sentiment.

[5] DeepSent contains 1,269 labeled photos available for download at: https://www.cs.rochester.edu/u/qyou/DeepSent/deepsentiment.html.

sentiment by You et al. (2015). The main advantage of using the DeepSent data set is that the sentiment labels are verified via a crowd-sourcing website (e.g., MTurk) to ensure the correctness of the labels.[6] The DeepSent data set offers a choice to select photos that have three, four, or all five participants in the MTurk survey to agree on the sentiment associated with the photo. For increased reliability, we use photos in which all five MTurk survey participants agree on the photo's sentiment. This restriction reduces the training sample to 882 photos.

The model is trained using a learning rate of 0.01 and 500 learning steps.[7] We set the train batch size to 100 photos. We reserve 10% of the training photos for the validation sample, and 20% for the test sample (photos are randomly assigned into these sets).[8] The training set is the set of photos used to adjust the weights in the final fully connected layer during the training process. The validation set is the set of photos not used to adjust the weights on the last fully connected layer, but their sole purpose is to help minimize overfitting by verifying that any increase in the training accuracy is not made at the expense of out-of-sample performance. Finally, the test set is a set of photos that is never seen by the model during the training process and that is used to compute a final accuracy score for the model. In addition to using a validation set to limit overfitting, we also enlarge the training set by using augmentation techniques (e.g., flip, scale) and adding regularization techniques in the model like dropout and label smoothing (Szegedy et al., 2016).

In Fig. 1, we plot the training and validation accuracy over the training steps for the model. We achieve 87.1% test accuracy.[9] This value for accuracy is similar to those obtained from other photo sentiment classification models in the literature. For example, Campos et al. (2017) compare various modifications to CNN models trained on DeepSent and report the test set accuracy values, which range between 78.3% and

[6] The process of verifying labels is an expensive and time-consuming task. However, not authenticating labels can lead to poorly performing models because the training data might be noisy, so we take the extra step of confirming the labels.

[7] The learning rate is the amount the weights in the model can change after each learning step. Learning steps refer to the number of times we pass our training set through our model. These are the common values used in the image classification applications of CNNs (You et al., 2015).

[8] We avoid using a large proportion of our training sample for the test set to ensure that we have enough photos in the training set. Some variation is present in the literature, but most studies assign the majority of the sample to the training set (usually between 70% and 80%), and the remainder is divided between the validation and test sets. For example, Yu et al. (2019) use a 77/3/20 split.

[9] Test accuracy is computed as the proportion of photos in the test set that the trained model is able to classify correctly, $Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative}$.

83.0%. To better test the performance of our model, we also calculate recall (86.2%), precision (94.3%), and F1 (90.1%).[10] Precision measures how accurate our model is at identifying positive photos out of all the photos that were predicted to be positive (this is an important metric when the cost of a false positive is high). Recall measures how accurate the model is at identifying positive photos out of all the positive photos in our sample (this is an important metric when the cost of a false negative is high). F1 is simply the harmonic mean of precision and recall.

One concern is that the photos in the DeepSent training set are from social media and might not closely resemble the types of professional news photos in the WSJ samples; thus, the model trained with the DeepSent training set might not be able to accurately classify professional photos. To address this concern, we randomly select 100 photos from our WSJ sample and classify each photo (classification is preformed in MTurk by five individuals).[11] We ask these individuals to rate the content of each photo based on sentiment. We pass these photos through the model to get predictions and compare the predictions to the responses we collect from MTurk. We summarize the results from this analysis in the following confusion matrix:[12]

		Actual	
		Positive	Negative
Prediction	Positive	64	19
	Negative	5	12

Based on the confusion matrices above, we calculate the performance of our model for classifying our sample of professional news photos from the WSJ. The accuracy is 76.0%; recall is 92.8%; precision is 77.1%; and F1 is 84.2%. Given that we have imbalanced classes in this sample, it is crucial to pay attention to the F1. Our accuracy, recall, precision, and F1 numbers are close to the ones You et al. (2015) report based on photo classification algorithms trained using the DeepSent data set (see Table 1 in You et al. (2015) for a summary of

[10] $Recall = \frac{True\ Positive}{True\ Positive+False\ Negative}$, $Precision = \frac{TruePositive}{True\ Positive+False\ Positive}$, and $F1 = \frac{2*Recall*Precision}{Recall+Precision}$

[11] We require MTurk "workers" to have a HIT approval rate of greater than 95% and to be located in the US.

[12] In the field of machine learning, a confusion matrix is a popular way of summarizing the performance of classification models. The numbers in the table represent the number of photos in our test sample that fall in each of the four buckets: true positives, false positives, false negatives, and true negatives.

their results). Overall, this table shows that our model trained with the DeepSent training set performs well at classifying news photos.

Table A1 in the Appendix presents the top-20 photos in our sample that are predicted to have the highest probability of depicting negative (top) or positive (bottom) sentiment. $PhotoNeg_{it}$ is the probability that photo *i* on day *t* depicts negative sentiment. Overall, these example photos help confirm that the photo classification model is working correctly.

2.2. The Wall Street Journal *sample*

The WSJ is a daily newspaper that focuses on major events and targets an audience with special interests in economics and the financial markets. The WSJ offers its subscribers online access to past articles, as far back as December 1997. However, articles prior to September 2008 are seldom included with photos. One reason behind the lack of photos prior to September 2008 is that WSJ does not own the rights to many of the photos and instead licenses the photos for a limited time from other news and media agencies, such as the Associated Press, Reuters, and Getty Images. Once the license expires, the photos are taken down.

Between September 2008 and September 2020, we collect the headline and the summary of each article, any associated photos, and the time stamps of when the article is published from the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion." We include international topics as long as they are related to economics (e.g., "Asia Business"). These sections cover major events related to companies or industries and general market conditions. See the top-20 photos with the highest probabilities of negative or positive sentiment from the WSJ in Table A1 of the Appendix.

We collect a total of 148,823 articles spanning 3,048 trading days. We classify these photos by sentiment using the photo classification model we discussed in Section 2.1.

2.3. *Variable construction*

Our main variable, *PhotoPes*, is calculated as the proportion of photos predicted to be negative on a given date. The formula for *PhotoPes* on day *t* is:

$$PhotoPes_t = \frac{\sum_i (Neg_{it})}{n_t}, \quad (1)$$

where Neg_{it} is an indicator variable for whether photo *i* on day *t* is predicted to have negative sentiment. The denominator, n_t, corresponds to the number of photos on date *t*. We provide the data on our website (https://www.kuntara.net/).

PhotoPes is not simply a binary measure. Although individual photos are classified as either negative or positive, *PhotoPes* is a continuous measure. Days with more negative photos have a higher value for *PhotoPes* compared to days with fewer negative photos. This is the same idea behind many sentiment measures in the literature. We show our baseline results using another version of *PhotoPes*, which is constructed using the predicted likelihood ($PhotoNeg_{it}$) instead of an indicator variable (Neg_{it}) for whether photo *i* on day *t* is negative and show that our results are consistent (Table 2, Panel B). We also show that if we adjust the cutoff for the indicator variable (Neg_{it}) from 50% to 55%, our results continue to hold (Table A2, Panel A).

One of the goals in this paper is to compare the pessimism embedded in photos and text. Motivated by Manela and Moreira (2017) and Cong et al. (2018), we analyze the headlines and summary of articles. Conceptually, the machine learning methodology we use to identify pessimism in photos includes highly nonlinear relationships among features. To facilitate a fair comparison between pessimism embedded in photos and text, we avoid classifying text using the lexicon approach. Instead, we use the sentiment tool in Stanford's CoreNLP software to evaluate the pessimism in each sentence and take the average pessimism score across all sentences in the text as the pessimism score for the article, *TextNeg* (Source: https://stanfordnlp.github.io/CoreNLP/). The sentiment tool is based on the recursive neural tensor network (RNTN) and is trained on a data set containing 215,154 phrases with fine-grained sentiment labels (Scale: {" Negative " = 1; " Neutral " = 0.5; " Positive " = 0}). This RNTN model performs especially well for shorter phrases and pushes the accuracy on short phrases to 85.4% (Socher et al., 2013). Since the headline and the summary section of articles from the WSJ are often brief, this tool is appropriate. The formula for *TextPes* is as follows:

$$TextPes_t = \frac{\sum_i (TextNeg_{it})}{n_t}, \quad (2)$$

where $TextNeg_{it}$ is the pessimism score for each article *i* on day *t* from the CoreNLP model. The denominator, n_t, corresponds to the number of articles in date *t*. *PhotoPes* and *TextPes* are winsorized at the 1% level. Our results remain the same without winsorizing (Table A2, Panel B).

We prefer the RNTN approach over the dictionary-based approach for the following reasons. First, the method we use to identify sentiment in images includes highly non-linear relations among features, whereas resorting to a basic methodology for the text can bias results in our favor by allowing images to capture more subtle sentiment characteristics, which would not be captured by the simple methodology utilized for the text. For instance, the dictionary approach gives the same weight to positive words like "good" and "great" and does not consider context and combination of words while the RNTN approach does. Thus, we believe it is appropriate to extract sentiment with methodologies of comparable complexity. Second, we analyze the headline and the summary section of articles, which are much shorter than the full text of news articles. The dictionary approach works best for long text like 10-Ks or full press releases because it avoids situations where none of the words overlap with words in the dictionary (Loughran and McDonald, 2011). In our data, 92.47% of headline and the summary section of articles do not contain any net negative or positive words that belong in the Loughran and McDonald dictionary. In contrast, only 63.66% of headline and the summary section of articles contain neutral sentiment using the machine learning approach.

2.4. *Descriptive statistics*

Table 1, Panel A, reports the summary statistics for the pessimism variables. We have 3,048 trading days between September 2008 and September 2020. On average, *PhotoPes* is 0.228, or 22.8% of photos are predicted to have negative sentiment on a given day. On average, *TextPes* is 0.686, indicating that, on average, the headline and the summary text of articles are made up of negative sentences (Scale: {" Negative " = 1; " Neutral " = 0.5; " Positive " = 0}). All the variables have significant first-order autocorrelation and thus are persistent (we address autocorrelation concerns in our tests).

Table 1, Panel B, reports summary statistics for the major US equity indices and ETFs that we use in our tests. Table 1, Panel C, reports how *PhotoPes* relates to *TextPes*. We calculate pairwise correlations and corresponding *p*-values between *PhotoPes* and *TextPes*. We find that *PhotoPes* and *TextPes* are positively correlated

(correlation coefficient = 0.079, *p*-value < 0.01). The positive correlation suggests that *PhotoPes* and *TextPes* are related. However, the correlation is not very high, indicating some information present in photos is distinct from that in the text in the headline and in the summary section of WSJ articles. At the article level, 62.65% of articles have consistent tone between their text and photo. Fig. 2 presents a time series of our *PhotoPes* and *TextPes* in both daily and monthly frequency. We observe that both series have large spikes during periods of elevated fear including the subprime mortgage crisis between 2008 and 2011 and the COVID-19 crisis in 2020. For example, we note particularly large spikes in both variables in 2011, which can be attributed to the high tension between the US and Al Qaeda as evident from the US government terminating the head of Al Qaeda on May 9, 2011. In addition, following the Tōhoku earthquake and tsunami in Japan on March 11, 2011 there was a tsunami warning on the west coast of the US. On March 16, 2011, wholesale food prices in the US rose by the largest monthly increase since November 1974, with an increase of 3.9%. Rumors about inflation getting out of control only made the situation worse. Taken together, these major news events are often accompanied by pessimistic photos that are used in our measure. Finally, we observe that *PhotoPes* and *TextPes* move closer together during market turmoil compared to normal times.[13]

3. Results

Behavioral models break from rational investor and market efficiency models by making two assumptions. First, behavioral models take into consideration that some investors are irrational and able to affect prices (De Long et al., 1990). Biases, such as extrapolation (Tversky and Kahneman, 1983) and overconfidence (Fischhoff et al., 1977), may lead irrational investors to increase demand for financial assets, pushing prices beyond economic fundamentals. Second, limits to arbitrage prevent rational investors from fully and instantly correcting price deviations from fundamentals (Pontiff, 1996; Shleifer and Vishny, 1997). One of the main predictions of behavioral models is market return reversal: when there is a positive (negative) sentiment spike, irrational investors will increase (decrease) demand for assets, driving prices away from

[13] We find that the correlation coefficient between *PhotoPes* and *TextPes* is highest during market turmoil. The correlation peaked in the last quarter of 2008 to about 0.5, and then declined to 0.2 in January 2009. Similarly, in the second quarter of 2020, the correlation jumped to 0.36 during the midst of the COVID pandemic.

fundamental levels. Behavioral models predict that this increase (decrease) in demand will lead to higher (lower) returns that will reverse over time as the market corrects to its fundamental level.

3.1. *News sentiment embedded in photos and text*

In this section, we discuss our main results. First, we show that the pessimism embedded in photos predicts market return reversals, consistent with behavioral model predictions (De Long et al., 1990). Second, we study how the pessimism embedded in photos interacts with the pessimism embedded in text. Third, we explore which type of news content is more effectively transmitted by photos compared to text. Fourth, we construct three real-world trading strategies to highlight the benefit of analyzing news photos.

3.1.1. *The impact of* PhotoPes *on market returns*

Table 2 presents our main results from the time series regression of market returns on lags of *PhotoPes*. Our specification is similar to that in Tetlock (2007) and Garcia (2013). Specifically, we run the following, with Newey and West (1987) *t*-statistics:

$$R_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(R_t) + \beta_3 L5(R_t^2) + \beta_4 X_t + \varepsilon_t, \quad (3)$$

where R_t denotes daily log returns on the CRSP value-weighted (VWRETD) index, the S&P 500 Index (SPX), the SPDR S&P 500 ETF (SPY), the Dow Jones Industrial Average Index (INDU), and the SPDR Dow Jones Industrial Average ETF (DIA). $PhotoPes_t$ is the proportion of photos that is predicted to be negative at time *t*; $L5$ transforms a variable into a row vector consisting of five lags of that variable; and X_t is a set of exogenous variables that includes an intercept, day-of-the-week indicators (except for Monday), and an indicator variable for whether time *t* is in a recession period. The timing of news sentiment measures at *t*-1 is based on news on day *t*-1 that is published and in the public domain on the morning of day *t*. We examine ETFs in addition to the major equity indices to confirm that our results are not driven by the illiquid components of the index (Da et al., 2015).

In Panel A in Table 2, $PhotoPes_{t-1}$ is negatively related to market returns in all specifications. In other words, days with a higher proportion of photos predicted to contain negative sentiment predict lower next day's returns, on average, compared to days with a lower proportion of photos predicted to contain negative sentiment. The relation is statistically significant at the 5% level for INDU and DIA and the 10% level

for VWRETD, SPX, and SPY. The magnitude of the effect is economically meaningful: the average impact of a one standard deviation shift in *PhotoPes* on the next day's VWRETD is 4.2 bps, which is nearly the size of the unconditional average daily returns of VWRETD (see Table 1, Panel B, for descriptive statistics). We examine the lags of *PhotoPes* to determine whether a reversal of the initial decline occurs in the following four trading days. We note that the reversal is concentrated on lags two and five. The magnitude of the reversal during the trading week after the photo is published ranges between 9.8 and 10.9 bps. The chi-square tests show that the reversal (the sum of the coefficients between *t*-2 and *t*-5) is statistically significant at the 1% level for INDU and at the 5% level for VWRETD, SPX, SPY, and DIA. Moreover, the chi-square test shows that the sum of the *PhotoPes* coefficients between *t*-1 and *t*-5 is indistinguishable from zero, suggesting that the initial decline at time *t*-1 is reversed over the remainder of the trading week. The return-reversal pattern is consistent with transient downward price pressure from pessimistic investors. We can rule out that *PhotoPes* contains fundamental information since we show that the initial decline in returns is followed by a complete reversal.

An important question is whether news photos directly cause investors to become more or less pessimistic. Although we cannot rule out the possibility that news editors select photos to reflect their readers' beliefs, we provide some evidence that suggests news photos directly affect investors' beliefs. We examine how *PhotoPes* relates to the Advisors Sentiment report. The Advisors Sentiment report "surveys the market views of over 100 independent investment newsletters and reports the findings as the percentage of advisors that are bullish, those bearish and those that expect a correction."[14] We regress the net bearish scores on *PhotoPes* on days *t*-1 to *t*-5 and find that *PhotoPes* predicts an increase in the net bearish score in the following trading week.[15] This exercise suggests that when news contains many negative photos, financial advisors become more bearish and less bullish over the following trading week.

Next, we modify *PhotoPes* by replacing the negative sentiment indicator variable, Neg_{it}, with the predicted likelihood of negative sentiment in photos. The predicted likelihood is the classification model's confidence about the prediction that photo *i* on day *t* contains negative sentiment; thus, the modified *PhotoPes*

[14] Source: https://www.investorsintelligence.com/x/advisors_sentiment.html.
[15] The results are available upon request from the authors.

is going to place a higher weight on photos that the photo classification model is more certain about. In Panel B of Table 2, we continue to find that $PhotoPes_{t-1}$ is negatively related to market returns. The economic magnitude and the statistical significance of the relation is stronger when we use the modified *PhotoPes* compared to our baseline results in Panel A. For example, the average impact of a one standard deviation shift in the modified *PhotoPes* on the next day's VWRETD is 5.2 bps. The relation is statistically significant at the 1% level for INDU and the 5% level for VWRETD, SPX, SPY, and DIA. The chi-square tests show that the reversal between *t*-2 and *t*-5 is statistically significant at the 1% level for VWRETD, SPX, INDU, and DIA and at the 5% level for SPY. For the sake of brevity, we use only the non-modified *PhotoPes* in our remaining tests.

3.1.2. PhotoPes *and sentiment embedded in text*

In this section, we compare the predictive abilities of *PhotoPes* and *TextPes*. Moreover, we also examine whether photos and news text are complements or substitutes. Specifically, does the news media use photos to enhance the sentiment embedded in text (complements), or are photos used to convey alternative information to the text (substitutes)? This is an important question to address given the increasingly multimodal news media.

In Table 3, we examine how *PhotoPes*, *TextPes*, and their interaction relates to market returns. We run the following regression:

$$R_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(TextPes_t) + \beta_3 (PhotoPes \times TextPes)_{t-1} + \beta_4 L5(R_t) + \beta_5 L5(R_t^2) + \beta_6 X_t + \varepsilon_t, \quad (4)$$

where $TextPes_t$ is the average pessimism score from the CoreNLP model for all articles at time *t* (defined in Section 2.3). To better understand how the two variables interact, we calculate *TextPes* and *PhotoPes* based on the same articles (i.e., we do not include articles that do not have photos in the calculation of *TextPes*).

After controlling for the pessimism in news text in Table 3, we continue to observe $PhotoPes_{t-1}$ is negatively related to market returns for VWRETD, SPX, SPY, INDU, and DIA. Although we show that *TextPes* without controlling for *PhotoPes* predicts return reversal in Table 4, the coefficients for $TextPes_{t-1}$ in Table 3 are negative, but not statistically significant at the 10% level, after adding *PhotoPes* to the model. It is important to note that one possible reason for *TextPes* loading insignificantly in this regression is because we compute

TextPes based on only articles with photos. News editors may include photos in articles when the topic is more strongly communicated using visual information as opposed to text, thus favoring our *PhotoPes* variable.

Next, we shift our focus to the interaction term of *PhotoPes* and *TextPes*. We assess the complementary or substitutive effects of pessimism in news variables on the next day's returns by relying on the interaction term, $PhotoPes \times TextPes_{t-1}$, and examining the marginal effect of one pessimism variable on the next day's return depending on the levels of the other pessimism variable (Aiken and West, 1991; Siggelkow, 2002). The media can use photos to reinforce the pessimism embedded in the text. If the pessimism embedded in photos enhances the pessimism embedded in the text, the marginal decrease in returns between the high and low levels of pessimism embedded in the photos should be higher when the pessimism embedded in the text is higher rather lower. To support the complementary perspective, we expect to find that the coefficient for the interaction term, $PhotoPes \times TextPes_{t-1}$, is negative. On the other hand, the media can use photos to convey an alternative dimension of investor sentiment that is not already reflected in the text. If pessimism embedded in the text and photo interact as substitutes, the marginal decrease in returns between the high and low levels of pessimism embedded in photos should be smaller when the pessimism embedded in the text is higher rather than lower. In other words, when there is a high level of *TextPes*, providing additional pessimism embedded in photos does not make a significant marginal contribution to market returns. To support the substitutive perspective, we expect to find that the coefficient for the interaction term, $PhotoPes \times TextPes_{t-1}$, is positive. In all five specifications of Table 3, the coefficients for $PhotoPes \times TextPes_{t-1}$ are positive and significant at the 10% level, supporting the substitutive hypothesis.[16]

Although *PhotoPes* dominates once we control for *TextPes*, we corroborate Tetlock (2007) and Garcia (2013) by showing that once we remove *PhotoPes* from our model, *TextPes* alone predicts the return reversal. Table 4 reports the results from the time series regression of market returns (VWRETD, SPX, SPY, INDU, and DIA) on lags of *TextPes* and controls (without controlling for *PhotoPes* and without requiring articles to have photos). In all five specifications of Table 4, $TextPes_{t-1}$ is negatively related to market returns at the 5% level

[16] Interaction plots are available on request.

for SPY, INDU, and DIA and at the 10% level for VWRETD and SPX. The magnitude of the effect is economically large: the average impact of a one standard deviation shift in *TextPes* on the next day's INDU is 8.5 bps, which is nearly 2.2 times the size of the unconditional average daily returns of INDU. This economic magnitude is in line with what Tetlock (2007) reports (8.1 bps). Moreover, the coefficients at *t*-2 continue to be negative and significant for SPY and INDU, suggesting that markets take more time to reflect information in news text compared to photos. The reversal for *TextPes* is concentrated in *t*-5, whereas the reversal for *PhotoPes* in our baseline results starts earlier at *t*-2 for most specifications. The magnitude of the reversal between *t*-2 and *t*-5 is between 8.0 bps and 8.1 bps for INDU and DIA and is statistically significant at the 10% level. Considering the lack of significant reversal between *t*-2 and *t*-5 for *TextPes* in specifications 1 to 3 but the significant positive coefficient at *t*-5 in all specifications, we conclude that the initial effect is only partially reversed. Overall, our results are consistent with *TextPes* containing both sentiment and fundamental information; thus, its relation to market returns is only partly transitory.

3.1.3. *Attention and* PhotoPes

We are interested in examining whether photos can play an attention-grabbing role in newspapers. Photos that evoke strong emotions from their audience could possibly detract from news text and in turn dominate the pessimism-return relation. Studies show that inclusion of photos in news can be attention-grabbing. For example, Garcia and Stark (1991) conduct an eye-tracking study and document how photos are the most common initial attraction to newspaper pages. Powell et al. (2015) find similar evidence by showing that content with text accompanied by photo or photo-alone is more attention-grabbing than content with text-alone. Motivated by these studies, we test how the relation between the pessimism embedded in news and market returns varies depending on the presence of salient news photos.

To test the attention-grabbing role of news photos, we run the following regression to differentiate between the effect of *PhotoPes* and *TextPes* on market returns along periods when photos are salient:

$$R_t = (E_t)[\beta_1 L5(PhotoPes_t) + \beta_2 L5(TextPes_t) + \beta_3 PhotoPes \times TextPes_{t-1} + \beta_4 L5(R_t) + \beta_5 L5(R_t^2)] + (1 - E_t)[\gamma_1 L5(PhotoPes_t) + \gamma_2 L5(TextPes_t) + \gamma_3 PhotoPes \times TextPes_{t-1} + \gamma_4 L5(R_t) + \gamma_5 L5(R_t^2)] + \beta_6 X_t + \varepsilon_t \,, \quad (5)$$

where E_t is an indicator variable that takes a value of one if day *t* is in the top or bottom decile of *PhotoPes*. We postulate that days when the majority of photos have the same sentiment (i.e., photos are overwhelmingly positive or negative), $E_t = 1$, readers will find the consistent message from photos to be salient and thus will focus on photos instead of text. On the other hand, days when photos have a neutral or mixed sentiment, $E_t = 0$, readers will find the mixed message from photos to be weak and thus will focus on the text.

We provide support for the attention-grabbing hypothesis in Table 5. We find that *PhotoPes* is negatively related to the next day's market returns during periods when photos are salient at the 5% level in all specifications. Moreover, during periods when photos are salient, *TextPes* is not statistically related to the next day's market returns. In contrast, we find that during periods when photos are not salient, *TextPes* is negatively related to the next day's market returns at the 5% level for SPX, SPY, INDU, and DIA, and at the 10% for VWRETD. Moreover, *PhotoPes* is not significantly related to the next day's market returns during periods when photos are not salient. Overall, the evidence in Table 5 suggest that during days when photos are salient, *PhotoPes* dominates and *TextPes* is not significant. In contrast, during days when photos are not salient, *TextPes* dominates and *PhotoPes* is not statistically significant. It is possible that photos grab attention to the whole article including the text; however, our evidence does not support this possibility since we find that the coefficient for $TextPes_{t-1}$ is insignificant (significant) during periods when photos are salient (not salient).[17]

3.1.4. *Which information is more effectively transmitted by photos?*

We attempt to answer what information news photos capture that text cannot. Prior studies suggest that photos can be a more effective medium at capturing traumatic events (Chemtob et al., 1999). We test whether the relation between market returns and pessimism embedded in photos and text varies during periods of elevated fear.

We proxy for periods of elevated fear using TRMI (Thomson Reuters MarketPsych Indices) measures that capture tone from a broad spectrum of news media sources and social media content for different topics

[17] In untabulated results, we document that the salience effect in Table 5 is not symmetric. The effect is strongest during days when most photos are positive (lowest *PhotoPes* decile). This evidence is consistent with Sicherman et al. (2015) who find that investors have selective attention. They find that investor attention to financial markets plummets by 9.5% during market declines. When *PhotoPes* is high (more negative photos), investors pay less attention to photos, thus the effect is smaller compared to days with low *PhotoPes*.

and emotions. TRMI uses a proprietary dictionary to classify the tone of different emotions and events in day *t* and quantify them on a 0 to 1 scale; the higher the score, the more prevalent is the event or emotion. To examine how the relation between pessimism in news using text and photos relates to market returns by fear level, we run the following regression:

$$R_t = (F_t)[\beta_1 L5(PhotoPes_t) + \beta_2 L5(TextPes_t) + \beta_3 PhotoPes \times TextPes_{t-1} + \beta_4 L5(R_t) + \beta_5 L5(R_t^2)] + (1 - F_t)[\gamma_1 L5(PhotoPes_t) + \gamma_2 L5(TextPes_t) + \gamma_3 PhotoPes \times TextPes_{t-1} + \gamma_4 L5(R_t) + \gamma_5 L5(R_t^2)] + \beta_6 X_t + \varepsilon_t, \quad (6)$$

where F_t is an indicator variable that takes a value of one if day *t* has an above-median fear score (computed as the average TRMI score of the following topics: fear and gloom).

We next examine how the effects of *PhotoPes* and *TextPes* on market returns vary by the level of fear portrayed in the news. In all five specifications in Table 6, $PhotoPes_{t-1}$ is negatively related to market returns during periods of both high and low levels of fear. However, the magnitude is much larger during high fear periods: for example, the average impact of a one standard deviation shift in *PhotoPes* on the next day's VWRETD is 10.3 bps during periods of high fear and only 3.7 bps for periods of low fear. The same cannot be said for *TextPes*. In all the specifications, $TextPes_{t-1}$ is negative, but not significant, and the magnitude of the coefficient is similar along periods of high or low levels of fear.[18]

Overall, the coefficient for the pessimism embedded in photos is roughly 2.8 times larger during periods of elevated fear compared to periods of little fear, while the coefficient for *TextPes* is similar during both periods. This evidence is consistent with photos being more effective than text at conveying fear or traumatic events like war (Chemtob et al., 1999).

3.1.5. *Applications*

We construct three real-world trading strategies to highlight the benefit of analyzing news photos using our sample of news from the WSJ. To ensure that the returns on these trading strategies are not driven by bid-

[18] Curious readers might be concerned that the F_t variable is simply picking out days with higher *PhotoPes*. This should not be a concern as the correlation between the TRMI fear and gloom scores and *PhotoPes* is quite low and insignificant. (correlation of -0.025 with *p*-value=0.16).

ask bounce or day-of-the-week effects, we use residuals from recursively regressing *PhotoPes* or *TextPes* on lagged returns (five lags) and day-of-the-week dummies available at time *t*-1, denoted by $PhotoPes^{\perp}$ or $TextPes^{\perp}$, respectively. The strategies require investors, each day, to either invest in SPY or in the risk-free asset (30-day Treasury bills), depending on lagged pessimism in news. The first strategy is based on the pessimism embedded in news photos: following days in which $PhotoPes^{\perp}$ is above its historical mean (expanding), we invest in the SPY at the market close of day *t*+3 and sell on the market close two days later (*t*+5). The second strategy is based on pessimism embedded in text: following days in which $TextPes^{\perp}$ is above its historical mean (expanding), we invest in the SPY at the market close of day *t*+3 and sell on the market close two days later (*t*+5). The third strategy involves pessimism from both text and photos: following days in which $PhotoPes^{\perp}$ and $TextPes^{\perp}$ are above their historical means, we invest in the SPY at the market close of day *t*+3 and sell on the market close two days later (*t*+5). The reason we pick the two-day investment period is that the reversal patterns for *PhotoPes* and *TextPes* overlap between *t*+4 and *t*+5. The reversal pattern for *PhotoPes* starts on day *t*+2 (Table 2); however, the reversal pattern for *TextPes* does not turn positive until day *t*+4 (Table 4).

In Panel A of Table 7, we report the means and standard deviations of daily excess returns (in percentages) and the Sharpe ratio of the three strategies involving $PhotoPes^{\perp}$, $TextPes^{\perp}$, or a combination of the two. We purchase the SPY in the first, second, and third strategies 1,992, 1,891, and 1,221 times, respectively. On average, the first strategy (based on $PhotoPes^{\perp}$) generates 5.8 bps in daily excess returns, which is higher than the 4.7 bps in excess returns for the buy-and-hold SPY strategy, and the 3.7 bps in daily excess returns for the $TextPes^{\perp}$ strategy. In addition, the $PhotoPes^{\perp}$ strategy generates a 0.88 Sharpe ratio compared to the 0.60 Sharpe ratio from the buy-and-hold SPY strategy and the 0.53 Sharpe ratio from the $TextPes^{\perp}$ strategy (Sharpe ratios are annualized). Next, we run time series regressions of daily excess returns from the $PhotoPes^{\perp}$ or the combined strategy on the Fama-French (1993) three factors (*Mkt_Rf*, *SMB*, and *HML*), the Carhart (1997) momentum factor (*MOM*), and the Da et al. (2014) short-run reversal factor (*ST_Rev*). Panel B of Table 7 reports the results.

We find that the combined strategy generates a positive and significant five-factor alpha of 5.45% per annum (2.1 bps per day).[19]

According to Fig. 3, by the end of our sample period, the first, second, and third strategies earned $4.76, $2.52, and $4.46 in excess cumulative returns for a $1 invested at the beginning of our sample period. The profit from the trading strategies based on *PhotoPes*$^{\perp}$ is economically higher than a buy-and-hold strategy of SPY, which earned $3.23 for a $1 invested at the beginning of our sample period.[20]

3.2. *Validation of* PhotoPes

3.2.1. *Limits to arbitrage*

Next, we focus on how limits to arbitrage affect the relation between *PhotoPes* and the market returns documented earlier. Limits to arbitrage suggest that correcting mispricing in the market is risky, and, thus, mispricing can take an extended period to correct (Shleifer and Vishny, 1997; Chu et al., 2020). De Long et al. (1990) suggest that investor sentiment should have the strongest effect on stocks that are the most difficult to arbitrage, while D'Avolio (2002) finds that arbitrage is riskier and costlier for riskier stocks than for safer stocks. Motivated by Wurgler and Zhuravskaya (2002) and Baker and Wurgler (2006), we focus on idiosyncratic volatility-sorted portfolios and size-sorted portfolios to test the prediction that *PhotoPes* should have the biggest impact on difficult-to-value or the riskiest stocks.

To test whether *PhotoPes* relates to stock returns differently depending on limits to arbitrage, we run the following regression:

$$R_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(R_t) + \beta_3 L5(R_t^2) + \beta_4 X_t + \varepsilon_t, \quad (7)$$

where R_t denotes the value-weighted daily returns on the highest and lowest quintile portfolios and the spread between the two portfolios (*High-low*) sorted on idiosyncratic volatility and size. More specifically, we sort stocks

[19] We continue to observe a positive and significant alpha if we use the Fama and French (2015) factors.

[20] Our trading strategy involves trading only two highly liquid assets—SPY and 30-day T-bills; thus, we think transaction costs are going to be small for our strategy. However, it is important to highlight that we do not take into account transaction costs including commissions, price impact costs, and capital gain taxes. The breakeven transaction cost that will eliminate our profit from the *PhotoPes* (*Combined*) trading strategy relative to holding the index is roughly 11.8 (7.6) bps per trade. The breakeven transaction cost that will make the total return from our *PhotoPes* (*Combined*) trading strategy equal to zero is 22.2 (19.0) bps per trade. We trade (switch between 100% in SPY and 100% in 30-day T-bills) 789 (866) times in our *PhotoPes* (*Combined*) trading strategy. If the cost per trade is above the breakeven transaction cost, our trading strategy will not be attractive.

in the CRSP universe into quintile portfolios based on idiosyncratic volatility estimated with the capital asset pricing model (CAPM) using the past 36-month returns (Panel A, Table 8). Next, we sort stocks in the CRSP universe into quintile portfolios based on idiosyncratic volatility estimated with a model consisting of the Fama and French (1993) three factors and the Carhart (1997) momentum factor using the past 36-month returns (Panel B). Finally, we sort stocks into quintile portfolios based on market capitalization of the firm (Panel C).

The first two specifications of Panel A in Table 8 show that $PhotoPes_{t-1}$ is negatively related to returns on the highest and lowest idiosyncratic volatility (CAPM)-sorted portfolios. The magnitude of the effect is larger and more significant for the highest idiosyncratic volatility-sorted portfolio compared to the lowest idiosyncratic volatility-sorted portfolio (over two times larger). Specifically, the average impact of a one standard deviation shift in *PhotoPes* on the next day's return on the highest and lowest idiosyncratic volatility-sorted portfolio is 7.1 (significant at the 5% level) and 3.5 (significant at the 10% level) bps, respectively. We examine the lags of *PhotoPes* to determine whether a reversal of the initial effect occurs in the following trading week. A clear and significant reversal pattern is significant at the 1% level for the highest idiosyncratic volatility-sorted portfolio and at the 5% level for the lowest idiosyncratic volatility-sorted portfolio. It is noteworthy that there is an overreaction in the reversal for the highest idiosyncratic volatility portfolio, but not for the lowest idiosyncratic portfolio. This overreaction can be attributed to stocks with the highest idiosyncratic volatility being more difficult to value, a fact that makes correcting mispricing difficult to do (Wurgler and Zhuravskaya, 2002).

To test whether the effect of sentiment on stock returns is stronger for the highest idiosyncratic volatility stocks compared to the lowest idiosyncratic volatility stocks, we regress the difference between the returns on the highest and lowest idiosyncratic volatility quintile portfolios (H-L) on lags of *PhotoPes* and controls. We show the results from this regression in the third specification of Panel A in Table 8. The coefficient for $PhotoPes_{t-1}$ is negative and significant at the 5% level, suggesting that *PhtooPes* has a stronger impact on the next day's returns for the highest compared with the lowest idiosyncratic volatility stocks.

In Panel B of Table 8, we find results similar to those in Panel A. More specifically, *PhotoPes* has a stronger impact on the next day's returns for the highest compared to the lowest idiosyncratic volatility-sorted

portfolios when we estimate idiosyncratic volatility using the model with the Fama and French (1993) and Carhart (1997) momentum factors.

In Panel C of Table 8, we examine the relation between *PhotoPes* and the returns for portfolios sorted on firm size. In the third specification of Panel C, we find that *PhotoPes* has a stronger impact on small companies compared to large companies. This evidence is consistent with our results using idiosyncratic volatility because small firms are more difficult to arbitrage compared to large firms (Baker and Wurgler, 2006). The key point of the results in Table 8 is that *PhotoPes* has a stronger impact on stocks that are difficult to arbitrage.

3.2.2. *Out-of-sample analysis*

Although the in-sample analysis provides more efficient parameter estimates and thus more precise return forecasts by utilizing all available data, Goyal and Welch (2008), among others, argue that out-of-sample tests are more appropriate to avoid the in-sample overfitting issue. Moreover, our out-of-sample tests are much less affected by the small-sample size distortions, such as the Stambaugh bias (Busetti and Marcucci, 2012). Hence, we examine the out-of-sample predictive performance of *PhotoPes*. Following Goyal and Welch (2008), Kelly and Pruitt (2013), Rapach et al. (2016), and many others, we evaluate the out-of-sample predictive performance based on the widely used Campbell and Thompson (2008) R^2_{OOS} statistic and the Clark and West (2007) MSPE-adjusted statistic. The R^2_{OOS} statistic measures the proportional reduction in the mean-squared prediction error (MSPE) for the predictive regression relative to the historical average benchmark:

$$R^2_{OOS} = 1 - \frac{\sum_{t=1}^{T}(R_t - \hat{R}_t)^2}{\sum_{t=1}^{T}(R_t - \bar{R}_t)^2}, \tag{8}$$

where $\hat{R}_t$ is the fitted value from a predictive regression of market returns on a one-period lag of *PhotoPes* estimated recursively with information available at time *t*-1. $\bar{R}_t$ denotes the historical average benchmark estimated through period *t*-1 from the constant expected return model ($R_t = \alpha + \varepsilon_t$). More specifically, we use the data from September 2008 to December 2009 as the initial estimation period so that the prediction evaluation period spans from January 2010 to September 2020. The initial in-sample estimation period balances between the desire to have enough observations to precisely estimate the initial parameters and the desire for a relatively long out-of-sample period for evaluation.

Goyal and Welch (2008) show that the historical average is a very stringent out-of-sample benchmark, and individual economic variables typically fail to outperform the historical average. The R^2_{OOS} statistic lies in the range (-∞, 1]. If $R^2_{OOS} > 0$, it means that the predicted $\hat{R}_t$ outperforms the historical average $\bar{R}_t$ in terms of MSPE.

The second statistic is the MSPE-adjusted statistic of Clark and West (2007) (henceforth, CW test). The null hypothesis is that the historical average MSPE is less than or equal to the predictive regression MSPE. The null hypothesis is tested against the one-sided (upper-tail) alternative hypothesis that the historical average MSPE is greater than the predictive regression MSPE, corresponding to H_0: $R^2_{OOS} \leq 0$ against H_A : $R^2_{OOS} > 0$. The MSPE-adjusted statistic accounts for the negative expected difference between the historical average MSPE and predictive regression MSPE under the null, so that it can reject the null even if the R^2_{OOS} statistic is negative. Our results in Table 9 show R^2_{OOS} is positive and significant according to the CW test. Across all test assets, R^2_{OOS} ranges between 0.123% and 0.302% and is significant at the 1% level according to the CW test.

To assess the economic usefulness of *PhotoPes* in predicting returns, we follow Campbell and Thompson (2008) and Rapach et al. (2016), among others, and compute the certainty equivalent return (CER) gain and Sharpe ratio for a mean-variance investor who allocates between equity index and risk-free assets (30-day Treasury bills) using the out-of-sample predictive regression forecasts, $\hat{R}_t$. At end of period *t*-1, the investor allocates w_{t-1} to the equity index during period *t* following:

$$w_{t-1} = \frac{1}{\gamma}\frac{\hat{R}_t}{\hat{\sigma}^2{}_t}, \tag{9}$$

where γ is the risk aversion coefficient of three, and $\hat{\sigma}^2{}_t$ is the variance forecast estimated over an expanding window of historical returns. The investor allocates the remaining capital, $1 - w_{t-1}$, in the risk-free asset.

The CER of the portfolio is:

$$CER = \hat{\mu} - \frac{1}{2}\gamma\hat{\sigma}^2, \tag{10}$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ are the average and variance, respectively, of the investor's portfolio over our evaluation window. The CER gain is the difference between the CER for an investor who allocates according to the

predictive regression forecast based on *PhotoPes*, $\hat{R}_t$, and the CER for an investor who allocates according to the historical average return forecast, $\bar{R}_t$. We multiply the CER gain by 250 to annualize the CER gain.

In Table 9, we report positive CER gain for all five indices between 1.042% and 1.439%. The CER gain is what an investor should be willing to pay in portfolio management fees to have access to the predictive regression forecast based on *PhotoPes* instead of the historical average return forecast. We also report the Sharpe ratio of the portfolio using the predictive regression forecast based on *PhotoPes* and the portfolio using the historical average return forecast. The Sharpe ratio is computed as the mean portfolio return net the risk-free rate divided by the standard deviation of the excess portfolio return. In Table 9, we note that the Sharpe ratios of the portfolio using the predictive regression forecast based on *PhotoPes* are notably higher than the Sharpe ratios of the portfolio using the historical average return forecast.

3.3. *The impact of* PhotoPes *on trading volume*

Another channel through which *PhotoPes* can affect market activity is trading volume. In Table 10, we examine whether the NYSE aggregate trading volume is related to *PhotoPes*. This channel helps us determine whether *PhotoPes* is a proxy for trading costs or investors' beliefs. Behavioral models, such as those of De Long et al. (1990) and Campbell et al. (1993), predict that a shock to sentiment indicates disagreement between rational and noise investors, and disagreement leads to an increase in trading. As the market absorbs these orders, we should see that a high or low value for *PhotoPes* is related to an increase in trading volume. On the other hand, if *PhotoPes* is a proxy for transaction costs, we expect to find a negative relation between *PhotoPes* and trading volume (Tetlock, 2007).

To remove time trends in trading volume, we model trading volume as follows:

$$V_t = \beta L5(V_t) + \gamma X_t + \varepsilon_t, \tag{11}$$

where V_t denotes the log of the aggregate daily NYSE trading volume. We take the residual from the above equation, ε_t, normalize it to have unit variance and mean of zero, and use it as the key dependent variable in the regressions below ($\bar{V}_t$). This procedure is intended to remove any calendar or day-of-the-week effects, in addition to the time trend. These are not the intended effects we attempt to explain using *PhotoPes*. Gallant et

al. (1992) and Garcia (2013) use a similar method to remove the time trend and other effects in the daily volume data. To test how *PhotoPes* is related to trading volume, we follow Tetlock (2007) and run the following model:

$$\bar{V}_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(|PhotoPes_t|) + \beta_3 L5(R_t) + \beta_4 L5(R_t^2) + \varepsilon_t, \quad (12)$$

where $\bar{V}_t$ is standardized (zero mean, unit variance) residual from Eq. (11) and $|PhotoPes_t|$ is the absolute value of *PhotoPes* after it is standardized to a mean of zero (high values of $|PhotoPes_t|$ will indicate days with either unusually large positive or unusually negative sentiment). Table 10 shows the coefficients for $|PhotoPes_{t-1}|$ are positive and significant at the 1% level, suggesting that either high or low values of pessimism in photos are able to predict an increase in the next day's abnormal trading volume. In terms of magnitude, a one standard deviation increase in *PhotoPes* is associated with a moving future trading volume of 0.080 standard deviations. Because we find that low or high values of *PhotoPes* predict an increase in trading volume, our evidence is consistent with the behavioral story and inconsistent with *PhotoPes* capturing transaction costs. Overall, our evidence on trading volume is consistent with that of Tetlock (2007) and Garcia (2013).

3.4. *Robustness*

Table A2 reports the results from the main regression for Table 2, except we make modifications to *PhotoPes*. In the original *PhotoPes*, a photo is labeled negative if the probability cutoff for Neg_{it} is above 50%. In the regressions for Panel A of Table A2, we adjust the cutoff for Neg_{it} from 50% to 55%. Looking at the five specifications of Panel A, we note that the return reversal results documented earlier, for instance, in Table 2, still hold with similar magnitudes.[21] In the regressions for Panel B of Table A2, we modify *PhotoPes* by not winsorizing the variable. We note that the results are similar.

One could be concerned that periods of high volatility are driving our results. Cox and Peterson (1994) show that extreme price returns are followed by short-term price reversals. Moreover, Connolly and Stivers (2003) find that weeks with extreme return-dispersion shocks tend to have a high number of macroeconomic news releases. In addition to including a recession indicator, we alleviate this concern in the following ways. First, we use GARCH (1,1)-adjusted-returns as our dependent variable. We calculate GARCH-adjusted returns

[21] In Table 2, Panel B, our results are robust to using the predicted likelihood a photo relays a negative sentiment instead of an indicator variable based on a specific cutoff.

by normalizing returns by the estimated GARCH (1,1) volatility, $\hat{\sigma}$. The normalization of the returns gives us a time series of returns with volatility normalized to one. In Panel C of Table A2, we continue to find that *PhotoPes* negatively predicts the next day's GARCH-adjusted returns in all five test assets. Moreover, we find significant reversal between lags 2 and 5 in all five specifications.

Past extreme negative returns might be predictive for future large negative returns. To alleviate concerns that the explanatory power of *PhotoPes* is subsumed by this effect, we remove the 1% (Panel D in Table A2) of the most extreme returns from our sample (trim 0.5% from both ends of the returns distribution). We continue to find that $PhotoPes_{t-1}$ is negatively related to market returns. Examining the chi-square tests, we note significant reversal (the sum of coefficients between $t-2$ and $t-5$) for all five test assets. In all specifications, we find the sum of the coefficients between $t-1$ and $t-5$ is not significantly differently from zero, suggesting a complete reversal within five days after the news photo is published. However, it is noteworthy that the results when we remove extreme returns or adjust for volatility (in Panels C and D) present an oscillating sign; specifically, there is a significant negative coefficient for day $t-3$. We hesitate to make any conjecture or draw any conclusion without a deep investigation into why controlling for extreme returns results in oscillating signs in the reversal window.

4. Conclusion

In this paper, we use a machine learning technique to extract information from a large sample of news media images and translate that information into a daily investor sentiment index, *PhotoPes*. We make three important contributions to the literature. First, we document that *PhotoPes* predicts market return reversals. This return reversal pattern is consistent with sentiment-induced transient mispricing (De Long et al., 1990; Campbell et al., 1993). This relation is strongest for difficult-to-arbitrage stocks. Moreover, we show that *PhotoPes* can predict market trading volume.

Second, we show that the pessimism embedded in photos and the pessimism embedded in news text act as substitutes for each other in predicting returns. Moreover, the pessimism embedded in news photos serves to grab attention away from text during periods when photos are salient. Our evidence shows that *PhotoPes* is especially useful for predicting market returns during periods of elevated fear.

Third, we demonstrate the benefit of using cutting-edge photo classification techniques to study how the information obtained from a large sample of news photos is relevant to the context of financial markets. Relying on surveys or crowd-sourcing websites (e.g., MTurk) to evaluate photos has shortcomings that make it infeasible to study large amounts of photos. However, with the continued focus on artificial intelligence, the machine learning techniques for analyzing photos are bound to grow in popularity and improve. This technological development will improve our ability to translate the rich information embedded in the billions of photos uploaded online into insights about central issues in financial research that have implications in both corporate finance and asset pricing.

In an increasingly multimodal media reality, future researchers should attempt to advance the machine learning technique of photo classification. Doing so could help bridge the gap between photo and text classification models, to better capture sentiment, and to determine the other important content embedded in news photos.

References

Aiken, L.S., West, S.G., 1991. Multiple regression: Testing and interpreting interactions. Sage Publications, Newbury Park, CA.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. Journal of Finance 61 (4), 1645–1680.

Bazley, W.J., Cronqvist, H., Mormann, M.M., 2021. Visual Finance: The Pervasive Effects of Red on Investor Behavior. Management Science.

Blankespoor, E., Hendricks, B.E., Miller, G.S., 2017. Perceptions and price: Evidence from CEO presentations at IPO roadshows. Journal of Accounting Research 55 (2), 275–327.

Buehlmaier, M., Whited, T., 2018. Are financial constraints priced? Evidence from textual analysis. Review of Financial Studies 31 (7), 2693–2728.

Busetti, F., Marcucci, J., 2012. Comparing forecast accuracy: A Monte Carlo investigation. International Journal of Forecasting 29 (1), 13–27.

Calomiris, C., Mamaysky, H., 2019. How news and its context drive risk and returns around the world. Journal of Financial Economics 133 (2), 299–336.

Campbell, J.Y., Grossman, S.J., Wang, J., 1993. Trading volume and serial correlation in stock returns. Quarterly Journal of Economics 108 (4), 905–939.

Campbell, J.Y., Thompson, S.B., 2008. Predicting the equity premium out of sample: Can anything beat the historical average? Review of Financial Studies 21 (4), 1509–1531.

Campos, V., Jou, B., Giro-i-Nieto, X., 2017. From pixels to sentiment: Fine-tuning CNNs for visual sentiment prediction. Image and Vision Computing 65 (September), 15–22.

Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of Finance 52 (1), 57–82.

Chaiken, S., Eagly, A. H., 1976. Communication modality as a determinant of message persuasiveness and message comprehensibility. Journal of Personality and Social Psychology 34 (4), 605–614.

Chemtob, C., Roitblat, H., Hamada, R., Muraoka, M., Carlson, J., Bauer, G., 1999. Compelled attention: The effects of viewing trauma-related stimuli on concurrent task performance in posttraumatic stress disorder. Journal of Traumatic Stress 12 (2), 309–326.

Chen, H., De, P., Hu, Y., Hwang, B., 2014. Wisdom of crowds: The value of stock opinions transmitted through social media. Review of Financial Studies 27 (5), 1367–1403.

Chu, Y., Hirshleifer, D., Ma, L., 2020. The causal effect of limits to arbitrage on asset pricing anomalies. Journal of Finance 75 (5),2631–2672.

Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138 (1), 291–311.

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

Cong, L.W., Liang, T., Zhang, X., 2018. Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information. Working Paper, Cornell University.

Connolly, R., Stivers, C., 2003. Momentum and reversals in equity-index returns during periods of abnormal turnover and return dispersion. Journal of Finance 58 (4), 1521–1556.

Cox, D.R., Peterson, D.R., 1994. Stock returns following large one-day declines: Evidence on short-term reversals and longer-term performance. Journal of Finance 49 (1), 255–267.

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

Da, Z., Engelberg, J., Gao, P., 2015. The sum of All FEARS investor sentiment and asset prices. Review of Financial Studies 28 (1), 1–32.

Da, Z., Liu, Q., Schaumburg, E., 2014. A closer look at the short-term return reversal. Management Science 60 (3), 658–674.

De Long, B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Noise trader risk in financial markets. Journal of Political Economy 98 (4), 703–738.

Duarte, J., Siegel, S., Young, L., 2012. Trust and credit: The role of appearance in peer-to-peer lending. Review of Financial Studies 25 (8), 2455–2483.

Edmans, A., Garcia, D., Norli, Ø., 2007. Sports sentiment and stock returns. Journal of Finance 62(4), 1967–1998.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33 (1), 3–56.
Fama, E., French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116 (1), 1–22.
Feng, G., Giglio, S., Xiu, D., 2020. Taming the factor zoo: A test of new factors. Journal of Finance 75 (3), 1327–1370.
Fischhoff, B., Slovic, P., Lichtenstein, S., 1977. Knowing with certainty: The appropriateness of extreme confidence. Journal of Experimental Psychology: Human Perception and Performance 3 (5), 552–564.
Gallant, A.R., Rossi, P.E., Tauchen, G., 1992. Stock prices and volume. Review of Financial Studies 5 (2), 199–242.
Garcia, D., 2013. Sentiment during recessions. Journal of Finance 68 (3), 1276–1300.
Garcia, M., Stark, P., 1991. Eyes on the news. St. Petersburg, FL: Poynter Institute for Media Studies.
Gentzkow, M., Shapiro, J.M., 2010. What drives media slant? Evidence from U.S. daily newspapers. Econometrica 78 (1), 35–71.
Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies 21 (4), 1455–1508.
Graham, J.R., Harvey, C.R., Puri, M., 2017. A corporate beauty contest. Management Science 63 (9), 2773–3145.
Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. Review of Financial Studies 33 (5), 2223–2273.
Halford, J.T., Hsu, H.S., 2014. Beauty is wealth: CEO attractiveness and firm value. Financial Review 55 (4), 529–556.
Hirshleifer, D., 2001. Investor psychology and asset pricing. Journal of Finance 56 (4), 1533–1597.
Hirshleifer, D., Shumway, T., 2003. Good day sunshine: Stock returns and the weather. Journal of Finance 58 (3), 1009–1032.
Jean, N., Burke, M., Xie, M., Davis, W., Lobell, D., Ermon, S., 2016. Combining satellite imagery and machine learning to predict poverty. Science 353 (6301), 790–794.
Jiang, F., Lee, J., Martin, X., Zhou, G., 2019. Manager sentiment and stock returns. Journal of Financial Economics 132 (1), 126–149.
Kelly, B., Pruitt, S., 2013. Market expectations in the cross-section of present values. Journal of Finance 68 (5), 1721–1756.
Kozak, S., Nagel, S., Santosh, S., 2018. Interpreting factor models. Journal of Finance 73 (3), 1183–1223.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems, 1097–1105. ACM, New York.
Loughran, T., McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. Journal of Finance 66 (1), 35–65.
Manela, A., Moreira, A., 2017. News implied volatility and disaster concerns. Journal of Financial Economics 123 (1), 137–162.
Mullainathan, S., Shleifer, A., 2005. The market for news. American Economic Review 95 (4), 1031–1053.
Mullainathan, S., Spiess, J., 2017. Machine learning: An applied econometric approach. Journal of Economic Perspectives, 31 (2), 87–106.
Nelson, D.L., Reed, U.S., Walling, J.R., 1976. Pictorial superiority effect. Journal of Experimental Psychology: Human Learning & Memory 2 (5), 523–-528.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.
Newhagen, J. E., Reeves, B., 1992. This evening's bad news: Effects of compelling negative television news images on memory. Journal of Communication, 42 (2), 25–41.
Paivio, A., 1991. *Images in mind*. New York: Harvester Wheatsheaf.
Pontiff, J., 1996. Costly arbitrage: Evidence from closed-end funds. Quarterly Journal of Economics 111 (4), 1135–1151.
Powell, T.E., Boomgaarden, H.G., De Swert, K., de Vreese, C.H., 2015. A clearer picture: The contribution of visuals and text to framing effects. Journal of Communication 65 (6), 997–1017.

Rapach, D., Ringgenberg, M., Zhou, G., 2016. Short Interest and Aggregate Stock Returns. Journal of Financial Economics 121(1), 46–65.
Sicherman, N., Loewenstein, G., Seppi, D.J., Utkus, S.P., 2015. Financial Attention. Review of Financial Studies 29(4), 863–897.
Siggelkow, N., 2002. Misperceiving interactions among complements and substitutes: Organizational consequences. Management Science 48 (7), 900–916.
Scoles, S., 2019. Researchers spy signs of slavery from space. Science 363 (6429), 804.
Shiller, R.J., 2005. *Irrational exuberance*. Princeton University Press, Princeton, NJ.
Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. Journal of Finance 52(1), 35–55.
Singer, E., 2002. The use of incentives to reduce nonresponse in household surveys. In Survey nonresponse, eds. Groves, R.M., Dillman, D.A., Eltinge, J.L., Little, R.J.A., 163–178. Wiley, Hoboken, NJ.
Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C., 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 conference on empirical methods in natural language processing, 1631–1642.
Spiegel, M., 2008. Forecasting the equity premium: Where we stand today. Review of Financial Studies 21 (4), 1453–1454.
Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2818–2826.
Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. Journal of Finance 62(3), 1139–1168.
Todorov, A., Mandisodza, A.N., Goren, A., Hall, C.C., 2005. Inferences of competence from faces predict election outcomes. Science 308 (5728), 1623–1626.
Tversky, A., Kahneman, D., 1983. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review 90 (4), 293–315.
Vissing-Jorgensen, A., 2003. Perspectives on behavioral finance: Does irrationality disappear with wealth? Evidence from expectations and actions. NBER Macroeconomics Annual 18 (2003), 139–208.
Wurgler, J., Zhuravskaya, K., 2002. Does arbitrage flatten demand curves for stocks? Journal of Business 75 (4), 583–608.
Yang, L., Hanneke, S., Carbonell, J., 2013. A theory of transfer learning with applications to active learning. Machine Learning 90 (2), 161–189.
You, Q., Luo, J., Jin, H., Yang, J., 2015. Robust image sentiment analysis using progressively trained and domain transferred deep networks. In The Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI).
Yu, J., Wang, Z., Majumdar, A., Rajagopal, R., 2019. DeepSolar: A machine learning framework to efficiently construct solar deployment database in the United States. Joule 2 (12), 2605-2617.
Zhou, G., 2018. Measuring investor sentiment. Annual Review of Financial Economics 10 (2018), 239–259.

Table 1. Summary statistics.

Panel A reports summary statistics for photo pessimism (*PhotoPes*) and text pessimism (*TextPes*). Panel B reports the sample statistics for the daily returns on the CRSP value-weighted (*VWRETD*) index, the S&P 500 Index (*SPX*), the SPDR S&P 500 ETF (*SPY*), the Dow Jones Industrial Average Index (*INDU*), and the SPDR Dow Jones Industrial Average ETF (*DIA*). Panel C reports the correlations between these variables. *PhotoPes* is calculated as the proportion of photos predicted to be negative on a given date. Text pessimism (*TextPes*) is calculated as the average pessimism score generated from the sentiment tool in Stanford's CoreNLP software. We use the headline and summary of each article for calculating *TextPes* and use news photos that belong to articles from the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion." The sample period ranges from September 2008 to September 2020. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Panel A: Summary statistics of sentiment variables

Variable	*N*	*Mean*	*Median*	*P25*	*P75*	*Std dev*
PhotoPes	3,048	0.228	0.222	0.180	0.270	0.077
TextPes	3,048	0.686	0.681	0.646	0.722	0.056

Panel B: Summary statistics of market returns

R_t(%)	*N*	*Mean*	*P50*	*P25*	*P75*	*Std dev*
VWRETD	3,048	0.045	0.081	-0.391	0.586	1.332
SPX	3,048	0.042	0.070	-0.380	0.570	1.335
SPY	3,048	0.049	0.070	-0.370	0.580	1.327
INDU	3,048	0.039	0.060	-0.390	0.550	1.283
DIA	3,048	0.047	0.070	-0.370	0.550	1.296

Panel C: Correlations between sentiment variables

	PhotoPes
TextPes	0.079***
	<0.01

Table 2. The impact of *PhotoPes* on market returns.

This table reports β_1 from the following time series regression:

$$R_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(R_t) + \beta_3 L5(R_t^2) + \beta_4 X_t + \varepsilon_t,$$

where $PhotoPes_t$ is calculated as the proportion of photos predicted to be negative at time *t*, $L5$ denotes five lags, and X_t contains a set of exogenous variables including a constant term, day-of-the-week dummies (except for Monday), and a recession dummy. In the regressions for Panel A, *PhotoPes* is calculated based on an indicator variable that the photo is predicted to be negative, while in the regressions for Panel B, *Predicted Likelihood PhotoPes* is calculated based on a predicted likelihood that the photo is negative. *PhotoPes* is winsorized at the 1% level and standardized to have a zero mean and unit variance. We use news photos that belong to articles from the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion." R_t is log daily return on the CRSP value-weighted ($VWRETD_t$) index, the S&P 500 Index (SPX_t), the SPDR S&P 500 ETF (SPY_t), the Dow Jones Industrial Average Index ($INDU_t$), and the SPDR Dow Jones Industrial Average ETF (DIA_t). The sample period ranges from September 2008 to September 2020. Newey and West (1987) standard errors are applied to compute the *t*-statistics. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Panel A: PhotoPes

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		SPX_t		SPY_t		$INDU_t$		DIA_t	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat
$PhotoPes_{t-1}$	-0.042*	-1.837	-0.041*	-1.803	-0.040*	-1.787	-0.046**	-2.182	-0.047**	-2.183
$PhotoPes_{t-2}$	0.055**	2.004	0.051*	1.886	0.046*	1.726	0.043*	1.687	0.038	1.502
$PhotoPes_{t-3}$	-0.033	-1.324	-0.030	-1.213	-0.030	-1.294	-0.024	-1.053	-0.025	-1.142
$PhotoPes_{t-4}$	0.030	1.299	0.024	1.047	0.026	1.143	0.030	1.387	0.033	1.487
$PhotoPes_{t-5}$	0.057**	2.137	0.059**	2.228	0.056**	2.119	0.057**	2.193	0.054**	2.103
Sum t-1 to t-5	0.067		0.063		0.058		0.060		0.053	
Sum t-2 to t-5	0.109		0.104		0.098		0.106		0.100	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	2.272	0.132	2.081	0.149	1.700	0.192	1.979	0.160	1.644	0.200
$\chi^2(1)$[Sum t-2 to t-5=0]	6.615**	0.010	6.200**	0.013	5.466**	0.019	6.973***	0.008	6.257**	0.012
Adj. *R*-squared	0.033		0.038		0.029		0.042		0.040	
N	3,044		3,044		3,044		3,044		3,044	

Panel B: Predicted Likelihood PhotoPes

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		SPX_t		SPY_t		$INDU_t$		DIA_t	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat
$PhotoPes_{t-1}$	-0.052**	-2.240	-0.052**	-2.269	-0.051**	-2.228	-0.055***	-2.581	-0.056**	-2.569
$PhotoPes_{t-2}$	0.039	1.487	0.034	1.345	0.028	1.139	0.029	1.255	0.024	1.031
$PhotoPes_{t-3}$	-0.020	-0.829	-0.017	-0.687	-0.017	-0.716	-0.011	-0.472	-0.011	-0.485
$PhotoPes_{t-4}$	0.019	0.789	0.016	0.676	0.019	0.769	0.026	1.120	0.029	1.242
$PhotoPes_{t-5}$	0.068***	2.596	0.070***	2.686	0.067**	2.576	0.062**	2.516	0.060**	2.390
Sum t-1 to t-5	0.054		0.051		0.046		0.051		0.046	
Sum t-2 to t-5	0.106		0.103		0.097		0.106		0.102	

	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	1.777	0.182	1.646	0.200	1.311	0.252	1.760	0.185	1.412	0.235
$\chi^2(1)$[Sum t-2 to t-5=0]	7.202***	0.007	7.033***	0.008	6.24**	0.012	8.013***	0.005	7.199***	0.007
Adj. *R*-squared	0.032		0.038		0.029		0.042		0.040	
N	3,044		3,044		3,044		3,044		3,044	

Table 3. Pessimism in photos and text.

This table reports $\beta_1, \beta_2,$ and β_3 from the following time series regression:

$$R_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(TextPes_t) + \beta_3 (PhotoPes \times TextPes)_{t-1} + \beta_4 L5(R_t) + \beta_5 L5(R_t^2) + \beta_6 X_t + \varepsilon_t,$$

where R_t is log daily return on the CRSP value-weighted ($VWRETD_t$) index, the S&P 500 Index (SPX_t), the SPDR S&P 500 ETF (SPY_t), the Dow Jones Industrial Average Index ($INDU_t$), and the SPDR Dow Jones Industrial Average ETF (DIA_t). $PhotoPes_t$ is calculated as the proportion of photos predicted to be negative at time *t*. $TextPes_t$ is calculated as the average pessimism score for headline and summary of each article generated from the sentiment tool in Stanford's CoreNLP software. $PhotoPes \times TextPest_{t-1}$ is the interaction between $PhotoPes_{t-1}$ and $TextPes_{t-1}$. $L5$ transforms a variable into a row vector consisting of five lags of that variable, and X_t contains a set of exogenous variables including a constant term, day-of-the-week dummies (except for Monday), and a recession dummy. We use news photos that belong to articles from the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion." *PhotoPes* and *TextPes* are winsorized at the 1% level and standardized to have a zero mean and unit variance. The sample period ranges from September 2008 to September 2020. Newey and West (1987) standard errors are applied to compute the *t*-statistics. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		SPX_t		SPY_t		$INDU_t$		DIA_t	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat
$PhotoPes_{t-1}$	-0.052**	-2.359	-0.049**	-2.229	-0.049**	-2.220	-0.054***	-2.600	-0.054***	-2.596
$TextPes_{t-1}$	-0.027	-0.816	-0.038	-1.183	-0.041	-1.241	-0.042	-1.390	-0.043	-1.372
$PhotoPesxTextPes_{t-1}$	0.038*	1.942	0.033*	1.754	0.034*	1.788	0.034*	1.899	0.034*	1.917
$PhotoPes_{t-2}$	0.056**	2.090	0.052**	1.980	0.048*	1.817	0.045*	1.801	0.040	1.595
$PhotoPes_{t-3}$	-0.027	-1.089	-0.025	-1.012	-0.025	-1.076	-0.020	-0.878	-0.021	-0.949
$PhotoPes_{t-4}$	0.032	1.394	0.026	1.125	0.028	1.228	0.031	1.454	0.034	1.581
$PhotoPes_{t-5}$	0.051*	1.936	0.053**	2.018	0.049*	1.896	0.051**	2.003	0.049*	1.897
$TextPes_{t-2}$	-0.040	-1.162	-0.045	-1.277	-0.042	-1.197	-0.047	-1.444	-0.042	-1.295
$TextPes_{t-3}$	-0.024	-0.640	-0.014	-0.387	-0.016	-0.432	-0.003	-0.078	-0.004	-0.119
$TextPes_{t-4}$	-0.016	-0.494	-0.019	-0.578	-0.020	-0.629	-0.014	-0.451	-0.021	-0.671
$TextPes_{t-5}$	0.090**	2.488	0.093***	2.603	0.095***	2.652	0.083**	2.497	0.089***	2.701
Sum t-1 to t-5 *PhotoPes*	0.060		0.057		0.051		0.053		0.048	
Sum t-2 to t-5 *PhotoPes*	0.112		0.106		0.100		0.107		0.102	
Sum t-1 to t-5 *TextPes*	-0.017		-0.023		-0.024		-0.023		-0.021	
Sum t-2 to t-5 *TextPes*	0.010		0.015		0.017		0.019		0.022	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5 *PhotoPes*=0]	1.866	0.172	1.740	0.187	1.376	0.241	1.633	0.201	1.318	0.251
$\chi^2(1)$[Sum t-2 to t-5 *PhotoPes*=0]	6.983***	0.008	6.412**	0.011	5.693**	0.017	7.187***	0.007	6.483**	0.011
$\chi^2(1)$[Sum t-1 to t-5 *TextPes*=0]	0.332	0.565	0.562	0.453	0.629	0.428	0.658	0.417	0.553	0.457
$\chi^2(1)$[Sum t-2 to t-5 *TextPes*=0]	0.050	0.823	0.146	0.702	0.181	0.671	0.262	0.609	0.328	0.567
Adj. *R*-squared	0.037		0.043		0.034		0.046		0.044	
N	3,044		3,044		3,044		3,044		3,044	

Table 4. The impact of *TextPes* on market returns.

This table reports β_1 from the following time series regression:

$$R_t = \beta_1 L5(TextPes_t) + \beta_2 L5(R_t) + \beta_3 L5(R_t^2) + \beta_4 X_t + \varepsilon_t,$$

where $TextPes_t$ is calculated as the average pessimism score for the headline and the summary of each article generated from the sentiment tool in Stanford's CoreNLP software on time *t*, $L5$ transforms a variable into a row vector consisting of five lags of that variable, and X_t contains a set of exogenous variables including a constant term, day-of-the-week dummies (except for Monday), and a recession dummy. *TextPes* is winsorized at the 1% level and standardized to have a zero mean and unit variance. We use news articles that belong to articles from the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion." R_t is log daily return on the CRSP value-weighted ($VWRETD_t$) index, the S&P 500 Index (SPX_t), the SPDR S&P 500 ETF (SPY_t), the Dow Jones Industrial Average Index ($INDU_t$), and the SPDR Dow Jones Industrial Average ETF (DIA_t). The sample period ranges from September 2008 to September 2020. Newey and West (1987) standard errors are applied to compute the *t*-statistics. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		SPX_t		SPY_t		$INDU_t$		DIA_t	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat
$TextPes_{t-1}$	-0.071*	-1.663	-0.083*	-1.904	-0.087**	-1.977	-0.085**	-2.002	-0.086**	-1.998
$TextPes_{t-2}$	-0.056	-1.466	-0.061	-1.590	-0.065*	-1.661	-0.062*	-1.686	-0.061	-1.628
$TextPes_{t-3}$	-0.007	-0.150	0.001	0.018	-0.004	-0.078	0.005	0.108	-0.003	-0.069
$TextPes_{t-4}$	0.021	0.527	0.017	0.436	0.021	0.540	0.027	0.709	0.029	0.756
$TextPes_{t-5}$	0.107**	2.280	0.116**	2.448	0.124***	2.612	0.110**	2.408	0.116**	2.521
Sum t-1 to t-5		-0.006		-0.010		-0.011		-0.005		-0.005
Sum t-2 to t-5		0.065		0.073		0.076		0.080		0.081
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	0.054	0.816	0.112	0.738	0.120	0.729	0.041	0.840	0.030	0.862
$\chi^2(1)$[Sum t-2 to t-5=0]	1.697	0.193	2.171	0.141	2.454	0.117	2.812*	0.094	2.931*	0.087
Adj. *R*-squared		0.028		0.035		0.028		0.038		0.039
N		3,044		3,044		3,044		3,044		3,044

Table 5. The impact of pessimism embedded in photos and text on market returns during salient *PhotoPes*.

This table reports $\beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2$, and γ_3 from the following time series regression:

$$R_t = (E_t)[\beta_1 L5(PhotoPes_t) + \beta_2 L5(TextPes_t) + \beta_3 (PhotoPes \times TextPes)_{t-1} + \beta_4 L5(R_t) + \beta_5 L5(R_t^2)] + (1 - E_t)[\gamma_1 L5(PhotoPes_t) + \gamma_2 L5(TextPes_t) + \gamma_3 (PhotoPes \times TextPes)_{t-1} + \gamma_4 L5(R_t) + \gamma_5 L5(R_t^2)] + \beta_6 X_t + \varepsilon_t,$$

where E_t is an indicator variable for whether period *t* is in the top or bottom decile of *PhotoPes*, R_t is the log daily return on the CRSP value-weighted ($VWRETD_t$) index, the S&P 500 Index (SPX_t), the SPDR S&P 500 ETF (SPY_t), the Dow Jones Industrial Average Index ($INDU_t$), and the SPDR Dow Jones Industrial Average ETF (DIA_t). $PhotoPes_t$ is calculated as the proportion of photos predicted to be negative at time *t*. $TextPes_t$ is calculated as the average pessimism score for the headline and the summary of each article generated from the sentiment tool in Stanford's CoreNLP software. $(PhotoPes \times TextPes)_{t-1}$ is the interaction between $PhotoPes_{t-1}$ and $TextPes_{t-1}$. $L5$ transforms a variable into a row vector consisting of five lags of that variable, and X_t contains a set of exogenous variables including a constant term, day-of-the-week dummies (except for Monday), and a recession dummy. We use news photos that belong to articles from the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion." *PhotoPes* and *TextPes* are winsorized at the 1% level and standardized to have a zero mean and unit variance. The sample period ranges from September 2008 to September 2020. Newey and West (1987) standard errors are applied to compute the *t*-statistics. **p* < 0.1; ***p* < 0.05; ****p* < 0.01.

	(1)				(2)				(3)			
	$VWRETD_t$				SPX_t				SPY_t			
	E_t=Salient photo period				E_t= Salient photo period				E_t= Salient photo period			
Variables	β	*t*-stat	γ	*t*-stat	β	*t*-stat	γ	*t*-stat	β	*t*-stat	γ	*t*-stat
$PhotoPes_{t-1}$	-0.070**	-2.479	-0.015	-0.332	-0.064**	-2.295	-0.016	-0.365	-0.063**	-2.260	-0.015	-0.342
$TextPes_{t-1}$	0.047	0.900	-0.070*	-1.883	0.031	0.606	-0.081**	-2.220	0.030	0.585	-0.080**	-2.137
$(PhotoPes \times TextPes)_{t-1}$	0.034	1.524	0.070	1.450	0.029	1.312	0.065	1.362	0.030	1.403	0.060	1.270
$PhotoPes_{t-2}$	0.100***	3.282	-0.034	-0.813	0.099***	3.316	-0.041	-0.978	0.094***	3.173	-0.044	-1.079
$PhotoPes_{t-3}$	-0.020	-0.659	-0.017	-0.411	-0.017	-0.565	-0.020	-0.498	-0.015	-0.526	-0.019	-0.489
$PhotoPes_{t-4}$	0.046*	1.710	0.009	0.216	0.042	1.571	-0.001	-0.033	0.042	1.632	0.003	0.066
$PhotoPes_{t-5}$	0.047	1.469	-0.009	-0.225	0.049	1.538	-0.004	-0.098	0.045	1.383	-0.006	-0.166
$TextPes_{t-2}$	0.044	0.684	-0.066*	-1.795	0.043	0.664	-0.071*	-1.912	0.051	0.761	-0.072*	-1.931
$TextPes_{t-3}$	-0.128**	-2.373	0.013	0.316	-0.113**	-2.174	0.022	0.538	-0.114**	-2.146	0.020	0.486
$TextPes_{t-4}$	-0.079	-1.412	0.002	0.057	-0.083	-1.488	-0.001	-0.018	-0.079	-1.488	-0.002	-0.050
$TextPes_{t-5}$	0.182***	3.133	0.059	1.520	0.186***	3.255	0.064	1.638	0.184***	3.135	0.065*	1.694
Sum t-1 to t-5 *PhotoPes*	0.103		-0.066		0.109		-0.082		0.103		-0.081	
Sum t-2 to t-5 *PhotoPes*	0.173		-0.051		0.173		-0.066		0.166		-0.066	
Sum t-1 to t-5 *TextPes*	0.066		-0.062		0.064		-0.067		0.072		-0.069	
Sum t-2 to t-5 *TextPes*	0.019		0.008		0.033		0.014		0.042		0.011	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5 *PhotoPes*=0]	2.905*	0.088	0.602	0.438	3.224*	0.073	0.958	0.328	2.851*	0.091	0.941	0.332
$\chi^2(1)$[Sum t-2 to t-5 *PhotoPes*=0]	9.777***	0.002	0.480	0.489	9.696***	0.002	0.823	0.364	8.943***	0.003	0.833	0.362
$\chi^2(1)$[Sum t-1 to t-5 *TextPes*=0]	0.437	0.509	1.842	0.175	0.405	0.525	2.247	0.134	0.525	0.469	2.443	0.118
$\chi^2(1)$[Sum t-2 to t-5 *TextPes*=0]	0.041	0.840	0.028	0.867	0.115	0.734	0.079	0.778	0.187	0.666	0.052	0.820
Adj. *R*-squared	0.075				0.086				0.071			
N	3,044				3,044				3,044			

	(4)				(5)			
	$INDU_t$				DLA_t			
	E_t= Salient photo period				E_t= Salient photo period			
Variables	β	*t*-stat	γ	*t*-stat	β	*t*-stat	γ	*t*-stat
$PhotoPes_{t-1}$	-0.066**	-2.537	-0.034	-0.797	-0.066**	-2.505	-0.034	-0.806
$TextPes_{t-1}$	0.016	0.338	-0.079**	-2.285	0.016	0.345	-0.075**	-2.124
$(PhotoPesxTextPes)_{t-1}$	0.030	1.471	0.060	1.358	0.031	1.536	0.056	1.253
$PhotoPes_{t-2}$	0.083***	2.901	-0.039	-0.999	0.077***	2.696	-0.044	-1.134
$PhotoPes_{t-3}$	-0.011	-0.411	-0.017	-0.444	-0.010	-0.354	-0.021	-0.556
$PhotoPes_{t-4}$	0.041	1.638	0.016	0.418	0.040	1.603	0.025	0.626
$PhotoPes_{t-5}$	0.041	1.320	0.003	0.080	0.036	1.111	0.003	0.070
$TextPes_{t-2}$	0.049	0.808	-0.076**	-2.169	0.064	1.055	-0.079**	-2.188
$TextPes_{t-3}$	-0.089*	-1.844	0.032	0.814	-0.088*	-1.801	0.032	0.819
$TextPes_{t-4}$	-0.068	-1.308	0.002	0.046	-0.073	-1.394	-0.008	-0.219
$TextPes_{t-5}$	0.159***	2.987	0.058	1.629	0.165***	2.929	0.062*	1.774
Sum t-1 to t-5 *PhotoPes*	0.088		-0.071		0.077		-0.071	
Sum t-2 to t-5 *PhotoPes*	0.154		-0.037		0.143		-0.037	
Sum t-1 to t-5 *TextPes*	0.067		-0.063		0.084		-0.068	
Sum t-2 to t-5 *TextPes*	0.051		0.016		0.068		0.007	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5 *PhotoPes*=0]	2.256	0.133	0.772	0.380	1.777	0.183	0.770	0.380
$\chi^2(1)$[Sum t-2 to t-5 *PhotoPes*=0]	8.235***	0.004	0.279	0.598	7.128***	0.008	0.281	0.596
$\chi^2(1)$[Sum t-1 to t-5 *TextPes*=0]	0.529	0.467	2.309	0.129	0.860	0.354	2.590	0.108
$\chi^2(1)$[Sum t-2 to t-5 *TextPes*=0]	0.338	0.561	0.107	0.744	0.598	0.439	0.027	0.869
Adj. *R*-squared	0.096				0.093			
N	3,044				3,044			

Table 6. The impact of pessimism embedded in photos and text on market returns during periods of elevated fear.

This table reports β_1, β_2, β_3, γ_1, γ_2, and γ_3 from the following time series regression:

$$R_t = (F_t)[\beta_1 L5(PhotoPes_t) + \beta_2 L5(TextPes_t) + \beta_3 PhotoPes \times TextPes_{t-1} + \beta_4 L5(R_t) + \beta_5 L5(R_t^2)] + (1 - F_t)[\gamma_1 L5(PhotoPes_t) + \gamma_2 L5(TextPes_t) + \gamma_3 (PhotoPes \times TextPes)_{t-1} + \gamma_4 L5(R_t) + \gamma_5 L5(R_t^2)] + \beta_6 X_t + \varepsilon_t,$$

where F_t is an indicator variable for whether period *t* has an above-median fear score (computed as the average TRMI score of the following topics: fear and gloom), R_t is log daily return on the CRSP value-weighted (*VWRETD*$_t$) index, the S&P 500 Index (*SPX*$_t$), the SPDR S&P 500 ETF (*SPY*$_t$), the Dow Jones Industrial Average Index (*INDU*$_t$), and the SPDR Dow Jones Industrial Average ETF (*DIA*$_t$). $PhotoPes_t$ is calculated as the proportion of photos predicted to be negative at time *t*. $TextPes_t$ is calculated as the average pessimism score for the headline and the summary of each article generated from the sentiment tool in Stanford's CoreNLP software. $PhotoPes \times TextPes_{t-1}$ is the interaction between $PhotoPes_{t-1}$ and $TextPes_{t-1}$. $L5$ transforms a variable into a row vector consisting of five lags of that variable, and X_t contains a set of exogenous variables including a constant term, day-of-the-week dummies (except for Monday), and a recession dummy. We use news photos that belong to articles from the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion." *PhotoPes* and *TextPes* are winsorized at the 1% level and standardized to have a zero mean and unit variance. The sample period ranges from September 2008 to September 2020. Newey and West (1987) standard errors are applied to compute the *t*-statistics. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

	(1)				(2)				(3)			
	VWRETD$_t$				*SPX*$_t$				*SPY*$_t$			
	F_t=Fear period				F_t=Fear period				F_t=Fear period			
Variables	β	*t*-stat	γ	*t*-stat	β	*t*-stat	γ	*t*-stat	β	*t*-stat	γ	*t*-stat
PhotoPes$_{t-1}$	-0.103**	-2.006	-0.037*	-1.893	-0.093*	-1.839	-0.036*	-1.822	-0.092*	-1.793	-0.036*	-1.848
TextPes$_{t-1}$	-0.018	-0.291	-0.019	-0.720	-0.031	-0.517	-0.025	-0.956	-0.035	-0.563	-0.026	-0.987
(PhotoPesxTextPes)$_{t-1}$	0.087**	2.063	0.009	0.525	0.079*	1.920	0.005	0.294	0.076*	1.830	0.006	0.334
PhotoPes$_{t-2}$	0.065	1.100	0.050**	2.284	0.056	0.956	0.049**	2.272	0.053	0.879	0.049**	2.257
PhotoPes$_{t-3}$	0.018	0.357	-0.054**	-2.423	0.023	0.470	-0.054**	-2.455	0.018	0.374	-0.056**	-2.550
PhotoPes$_{t-4}$	0.093*	1.914	0.001	0.059	0.077	1.587	0.001	0.059	0.078*	1.646	0.004	0.194
PhotoPes$_{t-5}$	0.080	1.431	0.027	1.214	0.082	1.482	0.029	1.307	0.075	1.374	0.030	1.349
TextPes$_{t-2}$	-0.063	-0.919	-0.041	-1.510	-0.071	-1.031	-0.044*	-1.649	-0.059	-0.862	-0.046*	-1.697
TextPes$_{t-3}$	-0.079	-1.094	0.013	0.517	-0.062	-0.885	0.018	0.721	-0.065	-0.920	0.019	0.756
TextPes$_{t-4}$	-0.062	-0.937	0.007	0.263	-0.069	-1.046	0.006	0.235	-0.066	-1.022	0.007	0.267
TextPes$_{t-5}$	0.134*	1.916	0.061**	2.122	0.138**	2.006	0.062**	2.161	0.138**	2.025	0.061**	2.150
Sum t-1 to t-5 *PhotoPes*	0.153		-0.013		0.145		-0.011		0.132		-0.009	
Sum t-2 to t-5 *PhotoPes*	0.256		0.024		0.238		0.025		0.224		0.027	
Sum t-1 to t-5 *TextPes*	-0.088		0.021		-0.095		0.017		-0.087		0.015	
Sum t-2 to t-5 *TextPes*	-0.070		0.040		-0.064		0.042		-0.052		0.041	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5 *PhotoPes*=0]	2.259	0.133	0.104	0.747	2.033	0.154	0.065	0.799	1.724	0.189	0.048	0.827
$\chi^2(1)$[Sum t-2 to t-5 *PhotoPes*=0]	7.078***	0.008	0.452	0.502	6.283**	0.012	0.496	0.481	5.629**	0.018	0.567	0.451
$\chi^2(1)$[Sum t-1 to t-5 *TextPes*=0]	1.306	0.253	0.618	0.432	1.562	0.211	0.406	0.524	1.332	0.249	0.360	0.548
$\chi^2(1)$[Sum t-2 to t-5 *TextPes*=0]	0.665	0.415	1.307	0.253	0.580	0.446	1.438	0.231	0.372	0.542	1.432	0.231
Adj. *R*-squared	0.046				0.053				0.044			
N	3,044				3,044				3,044			

	(4)				(5)			
	$INDU_t$				DIA_t			
	F_t=Fear period				F_t=Fear period			
Variables	β	*t*-stat	γ	*t*-stat	β	*t*-stat	γ	*t*-stat
$PhotoPes_{t-1}$	-0.101**	-2.102	-0.043**	-2.252	-0.102**	-2.090	-0.043**	-2.260
$TextPes_{t-1}$	-0.031	-0.563	-0.028	-1.134	-0.030	-0.505	-0.027	-1.099
$PhotoPesxTextPes_{t-1}$	0.084**	2.145	0.003	0.215	0.083**	2.107	0.003	0.217
$PhotoPes_{t-2}$	0.039	0.701	0.050**	2.410	0.029	0.511	0.051**	2.473
$PhotoPes_{t-3}$	0.026	0.555	-0.049**	-2.368	0.022	0.487	-0.050**	-2.424
$PhotoPes_{t-4}$	0.070	1.541	0.013	0.622	0.074	1.626	0.015	0.733
$PhotoPes_{t-5}$	0.071	1.322	0.034	1.624	0.068	1.255	0.033	1.614
$TextPes_{t-2}$	-0.079	-1.245	-0.044*	-1.682	-0.072	-1.118	-0.042	-1.637
$TextPes_{t-3}$	-0.047	-0.714	0.027	1.101	-0.048	-0.728	0.025	1.050
$TextPes_{t-4}$	-0.050	-0.815	0.005	0.186	-0.061	-0.978	0.002	0.067
$TextPes_{t-5}$	0.127**	2.017	0.051*	1.913	0.135**	2.185	0.054**	1.995
Sum t-1 to t-5 *PhotoPes*	0.105		0.005		0.091		0.006	
Sum t-2 to t-5 *PhotoPes*	0.206		0.048		0.193		0.049	
Sum t-1 to t-5 *TextPes*	-0.080		0.011		-0.076		0.012	
Sum t-2 to t-5 *TextPes*	-0.049		0.039		-0.046		0.039	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5 *PhotoPes*=0]	1.198	0.274	0.019	0.891	0.916	0.338	0.036	0.850
$\chi^2(1)$[Sum t-2 to t-5 *PhotoPes*=0]	5.23**	0.022	1.905	0.168	4.583**	0.032	2.065	0.151
$\chi^2(1)$[Sum t-1 to t-5 *TextPes*=0]	1.307	0.253	0.169	0.681	1.086	0.297	0.169	0.681
$\chi^2(1)$[Sum t-2 to t-5 *TextPes*=0]	0.396	0.529	1.350	0.245	0.309	0.578	1.307	0.253
Adj. *R*-squared	0.058				0.058			
N	3,044				3,044			

Table 7. Performance of trading strategies based on *PhotoPes*.

Panel A reports the mean, standard deviation, and Sharpe ratio (*SR*) of excess daily returns (in percentages) for our trading strategies involving *PhotoPes*, *TextPes*, *Combined* (a combination of the two plus the SPY index), and *Index* (SPY). N indicates the number of days holding the SPY index. See Section 3.1.5 for details about the strategies. Panel B reports estimates from time series regressions of daily excess returns from the *PhotoPes* and combined strategies on the Fama-French (1993) three factors (*Mkt_Rf*, *SMB*, and *HML*), the Carhart (1997) momentum factor (*UMD*), and the Da et al. (2014) short-run reversal factor (*ST_Rev*). Newey and West (1987) standard errors are applied to compute the *t*-statistics. The sample period ranges from September 2008 to September 2020. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Panel A: Summary statistics of trading strategies					
Strategy	*N*	*Mean*	*t-stat*	*Std dev*	*SR*
PhotoPes	1,992	0.058	3.251	1.119	0.052
TextPes	1,891	0.037	2.085	1.166	0.032
Combined	1,221	0.054	3.547	0.980	0.055
Index	3,034	0.047	2.246	1.325	0.036

Panel B: Time series regression				
	(1)		(2)	
	Combined strategy$_t$		*PhotoPes strategy*$_t$	
Variables	β	*t*-stat	β	*t*-stat
Alpha	0.021*	1.742	0.014	1.302
Mkt_Rf_t	51.0***	13.347	69.4***	22.912
SMB_t	-1.830	-0.317	-6.612	-1.418
HML_t	-15.6***	-3.517	-13.9***	-3.799
UMD_t	-4.947*	-1.743	-5.489**	-2.121
ST_Rev_t	6.754*	1.708	3.252	1.002
Adj. *R*-squared	0.545		0.706	
N	3,034		3,034	

Table 8. Validation test.

This table reports the results of β_1 from the following time series regression:

$$R_t = \beta_1 L5(PhotoPes_t) + \beta_3 L5(R_t) + \beta_4 L5(R_t^2) + \beta_5 X_t + \varepsilon_t,$$

where R_t is the log value-weighted daily return on the top (*High*) and bottom (*Low*) quintile portfolios sorted on monthly idiosyncratic volatility using the CAPM (Panel A), idiosyncratic volatility using the Fama-French (1993) factors (*FF*) and the Carhart (1997) momentum factor (*UMD*) (Panel B), and firm size (Panel C). *High-low* is the difference between highest and lowest quintile portfolio return. $PhotoPes_t$ is calculated as the proportion of photos predicted to be negative at time *t*; $L5$ transforms a variable into a row vector consisting of five lags of that variable; and X_t contains a set of exogenous variables including a constant term, day-of-the-week dummies (except for Monday), and a recession dummy. *PhotoPes* is winsorized at the 1% level and standardized to have a zero mean and unit variance. We use news photos that belong to articles from the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion." The sample period ranges from September 2008 to September 2020. Newey and West (1987) standard errors are applied to compute the *t*-statistics. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

	Panel A: VW Idvol (CAPM)$_t$					
	(1)		(2)		(3)	
	High		*Low*		*High-low*	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat
PhotoPes$_{t-1}$	-0.071**	-2.182	-0.035*	-1.852	-0.040**	-1.986
PhotoPes$_{t-2}$	0.107***	2.833	0.042*	1.902	0.061***	2.810
PhotoPes$_{t-3}$	-0.044	-1.241	-0.030	-1.481	-0.008	-0.381
PhotoPes$_{t-4}$	0.061*	1.745	0.029	1.518	0.028	1.292
PhotoPes$_{t-5}$	0.064*	1.778	0.047**	2.043	0.019	0.867
Sum t-1 to t-5	0.117		0.053		0.060	
Sum t-2 to t-5	0.188		0.088		0.100	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	3.431*	0.064	2.054	0.152	2.454	0.117
$\chi^2(1)$[Sum t-2 to t-5=0]	9.441***	0.002	6.304**	0.012	7.121***	0.008
Adj. *R*-squared	0.026		0.030		0.028	
N	3,044		3,044		3,044	

	Panel B: VW Idvol (FF+UMD)$_t$					
	(1)		(2)		(3)	
	High		*Low*		*High-low*	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat
PhotoPes$_{t-1}$	-0.066**	-2.073	-0.037*	-1.828	-0.035*	-1.854
PhotoPes$_{t-2}$	0.102***	2.759	0.045*	1.942	0.052**	2.512
PhotoPes$_{t-3}$	-0.049	-1.387	-0.027	-1.300	-0.014	-0.722
PhotoPes$_{t-4}$	0.063*	1.871	0.032	1.633	0.028	1.344
PhotoPes$_{t-5}$	0.062*	1.791	0.049**	2.051	0.015	0.736
Sum t-1 to t-5	0.112		0.062		0.046	
Sum t-2 to t-5	0.178		0.099		0.081	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	3.412*	0.065	2.691	0.101	1.560	0.212
$\chi^2(1)$[Sum t-2 to t-5=0]	9.026***	0.003	7.362***	0.007	5.133**	0.023
Adj. *R*-squared	0.028		0.032		0.027	
N	3,044		3,044		3,044	

	Panel C: VW size$_t$					
	(1)		(2)		(3)	
	Large		*Small*		*Large-small*	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat
PhotoPes$_{t-1}$	-0.041*	-1.825	-0.073**	-2.452	0.033**	1.984
PhotoPes$_{t-2}$	0.046*	1.776	0.063*	1.914	-0.011	-0.704
PhotoPes$_{t-3}$	-0.028	-1.163	-0.026	-0.810	-0.005	-0.296
PhotoPes$_{t-4}$	0.025	1.110	0.023	0.779	0.000	0.026
PhotoPes$_{t-5}$	0.055**	2.080	0.075**	2.325	-0.016	-1.088

Sum t-1 to t-5		0.057		0.062		0.001
Sum t-2 to t-5		0.098		0.135		-0.032
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	1.805	0.179	1.311	0.252	0.001	0.970
$\chi^2(1)$[Sum t-2 to t-5=0]	5.796**	0.016	6.648***	0.010	1.482	0.224
Adj. *R*-squared		0.036		0.043		0.013
N		3,044		3,044		3,044

Table 9. Out-of-sample analysis and flexible model.

This table reports R^2_{OOS} (out-of-sample R^2) as a percentage and the associated *p*-values using the MSPE-adjusted statistic in Clark and West (2007) for recursively estimated predictive regression of the CRSP value-weighted (*VWRETD*) index, the S&P 500 Index (*SPX*), the SPDR S&P 500 ETF (*SPY*), the Dow Jones Industrial Average Index (*INDU*), or the SPDR Dow Jones Industrial Average ETF (*DIA*) on a one-period lag of *PhotoPes*. We use data between September 2008 and December 2009 for the initial estimation period. The sample period ranges from September 2008 to September 2020. Clark-West (2007) MSPE-adjusted statistic that tests the null of $R^2_{OOS} \leq 0$ against the alternative $R^2_{OOS} > 0$ is applied to test for significance. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. The CER gain is the annualized certainty equivalent return gain for the investor. The Sharpe ratio (*PhotoPes*) is the average portfolio return using the predictive regression forecast based on *PhotoPes* net the risk-free rate divided by the standard deviation of the excess portfolio return. The Sharpe ratio (historical) is the average portfolio return using the historical average return forecast net the risk-free rate divided by the standard deviation of the excess portfolio return.

Return	R^2_{OOS} *(%)*	*CER gain (%)*	*Sharpe ratio (PhotoPes)*	*Sharpe ratio (historical)*
VWRETD	0.251***	1.439	0.470	0.380
SPX	0.162***	1.150	0.441	0.304
SPY	0.302***	1.290	0.490	0.400
INDU	0.123***	1.042	0.396	0.224
DIA	0.158***	1.420	0.490	0.380

Table 10. *PhotoPes* and trading volume.

This table reports β_1 to β_2 from the following time series regression:

$$\bar{V}_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(|PhotoPes_t|) + \beta_3 L5(R_t) + \beta_4 L5(R_t^2) + \varepsilon_t,$$

where $\bar{V}_t$ is the standardized (zero mean, unit variance) residual from regressing the log of aggregate NYSE trading volume on its own lags (five lags) and month- and day-of-the-week dummies; R_t is the log daily return on the CRSP value-weighted index; $PhotoPes_t$ is calculated as the proportion of photos predicted to be negative at time *t*; and $|PhotoPes_t|$ is the absolute value of standardized *PhotoPes* (with zero mean and unit variance). In both panels, $L5$ transforms a variable into a row vector consisting of five lags of that variable. *PhotoPes* is winsorized at the 1% level and standardized to have a zero mean and unit variance. We use news photos that belong to articles from the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion." The sample period ranges from September 2008 to September 2020. Newey and West (1987) standard errors are applied to compute the *t*-statistics. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

	$\bar{V}_t$	
Variables	β	*t*-stat
$PhotoPes_{t-1}$	-0.003	-0.157
$PhotoPes_{t-2}$	-0.006	-0.278
$PhotoPes_{t-3}$	0.014	0.686
$PhotoPes_{t-4}$	-0.015	-0.786
$PhotoPes_{t-5}$	-0.011	-0.582
$\lvert PhotoPes_{t-1} \rvert$	0.080***	2.912
$\lvert PhotoPes_{t-2} \rvert$	0.013	0.453
$\lvert PhotoPes_{t-3} \rvert$	0.008	0.237
$\lvert PhotoPes_{t-4} \rvert$	0.059**	2.059
$\lvert PhotoPes_{t-5} \rvert$	0.015	0.511
Adj. R-squared	0.023	
N	3,044	

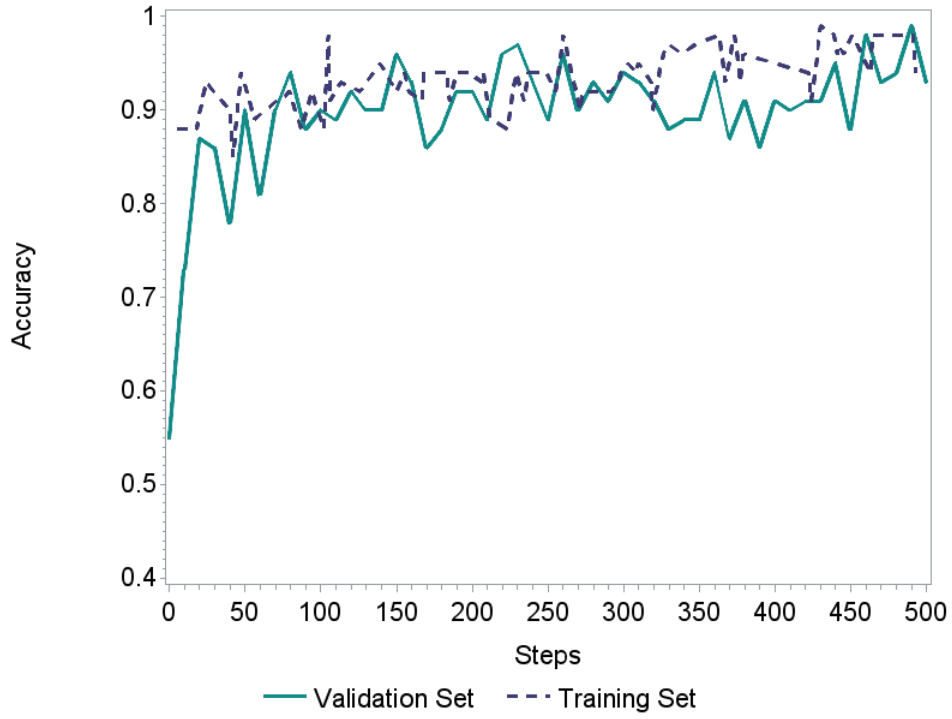


Fig. 1. Model accuracy.

This figure shows the training and validation accuracy for the first 500 training steps using the Google Inception (v3) model and the DeepSent training data set. Accuracy is computed as the proportion of photos that the model is able to classify correctly, $Accuracy = \frac{True\ Positive+True\ Negative}{True\ Positive+True\ Negative+False\ Positive+False\ Negative}$. Steps refers to learning steps or the number of times we pass the training set through the model. The training set is the set of photos used to adjust the weights in the final fully connected layer during the training process. The validation set is the set of photos not used to adjust the weights on the last fully connected layer, but its sole purpose is to help minimize overfitting by verifying that any increase in the training accuracy is not made at the expense of out-of-sample performance. We assign 10% of the photos to the validation set, 70% to the training set, and 20% to the testing set.

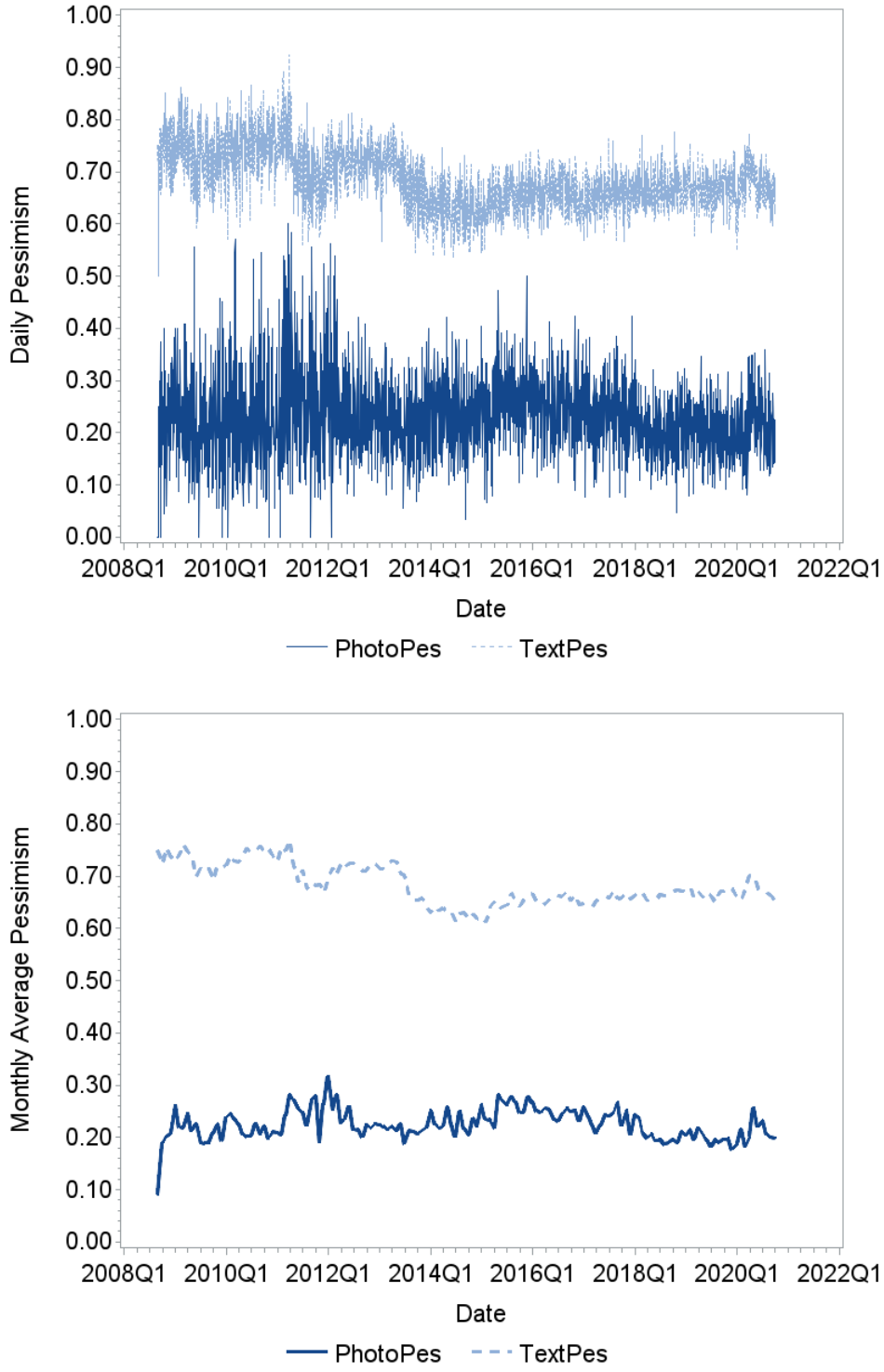


Fig. 2. The time series of *PhotoPes* and *TextPes*.

These figures show the time series of *PhotoPes* and *TextPes* on a daily (top) and monthly (bottom) basis for our sample period between September 2008 and September 2020.

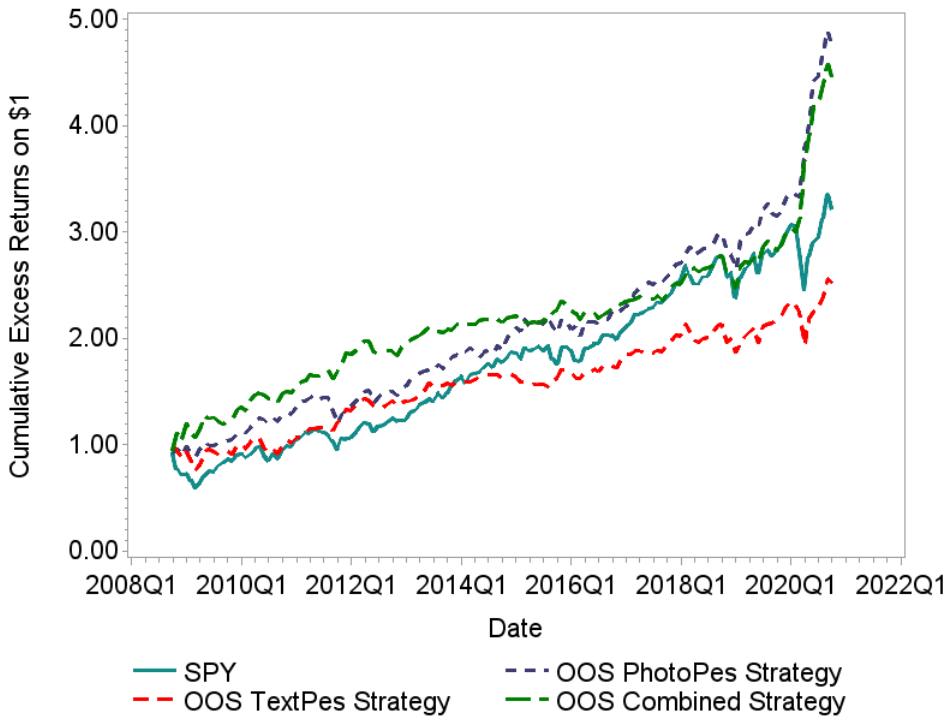


Fig. 3. Trading strategy.

This graph depicts the accumulated dollar excess return from out-of-sample trading strategies based on pessimism from news compared to the passive strategy of holding the SPY. See Section 3.1.5 for details about the trading strategies.

Table A1. Photos with the highest probability of negative and positive sentiment.

This table presents the 20 photos that have the highest *PhotoNeg* or the highest probability of having negative sentiment (top) and the 20 photos that have the lowest *PhotoNeg* or the highest probability of having positive sentiment (bottom). *TextNeg*, presented at the end of the table, is the pessimism score for the text from the headline and article summary associated with each photo. Pessimism in text is measured using Stanford's CoreNLP software. News photos come from WSJ articles that appeared in the following WSJ sections: "Business," "Economy," "Markets," "Politics," and "Opinion."

Twenty photos with the highest *PhotoNeg* or highest probability of having negative sentiment.


a1


a2

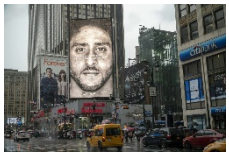

a3


a4


a5


a6


a7


a8


a9


a10


a11


a12


a13


a14


a15


a16


a17


a18


a19


a20

Twenty photos with the lowest *PhotoNeg* or highest probability of having positive sentiment.


b1


b2

b3


b4


b5

b6

b7


b8

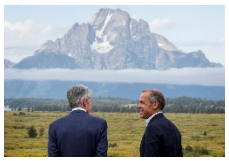

b9


b10


b11


b12

b13


b14


b15


b16


b17


b18

b19


b20

Photo ID	*PhotoNeg*	*TextNeg*	Photo ID	*PhotoNeg*	*TextNeg*
a1	1	0.50	b1	0	0.50
a2	1	1.00	b2	0	0.50
a3	1	1.00	b3	0	0.50
a4	1	1.00	b4	0	0.50
a5	1	0.75	b5	0	0.50
a6	1	0.25	b6	0	0.63
a7	1	1.00	b7	0	0.50
a8	1	0.58	b8	0	0.75
a9	1	1.00	b9	0	0.50
a10	1	0.50	b10	0	0.75
a11	1	0.75	b11	0	0.63
a12	1	1.00	b12	0	0.50
a13	1	1.00	b13	0	0.25
a14	1	0.00	b14	0	0.75
a15	1	1.00	b15	0	0.88
a16	1	1.00	b16	0	0.25
a17	1	0.50	b17	0	0.25
a18	1	0.83	b18	0	0.75
a19	1	1.00	b19	0	0.63
a20	1	1.00	b20	0	0.50

Table A2. Robustness test.

This table reports results of β_1 from the following time series regression:

$$R_t = \beta_1 L5(PhotoPes_t) + \beta_2 L5(R_t) + \beta_3 L5(R_t^2) + \beta_5 X_t + \varepsilon_t,$$

where R_t is log daily return on the CRSP value-weighted ($VWRETD_t$) index, the S&P 500 Index (SPX_t), the SPDR S&P 500 ETF (SPY_t), the Dow Jones Industrial Average Index ($INDU_t$), and the SPDR Dow Jones Industrial Average ETF (DIA_t). $PhotoPes_t$ is calculated as the proportion of photos predicted to be negative at time *t* (modified in several ways as discussed below); $L5$ transforms a variable into a row vector consisting of five lags of that variable; and X_t contains a set of exogenous variables including a constant term, day-of-the-week dummies (except for Monday), and a recession dummy. Panel A presents results when the probability cutoff for Neg_{it} changes from 50% to 55%. Panel B presents results when *PhotoPes* is not winsorized. In the regressions for Panel C, R_t is normalized log-returns. We normalize returns by dividing them by the estimate's volatility from the GARCH (1,1) model. In the regressions for Panel D, we remove the 1% most extreme returns in our sample. *PhotoPes* in all but Panel B is winsorized at the 1% level. We use news photos that come from articles that appeared in the following sections: "Business," "Economy," "Markets," "Politics," and "Opinion." *PhotoPes* is standardized to have a zero mean and unit variance. The sample period ranges from September 2008 to September 2020. Newey and West (1987) standard errors are applied to compute the *t*-statistics. *p < 0.1; **p < 0.05; ***p < 0.01.

Panel A: Cutoff = {0.55,0.45}										
	(1)		(2)		(3)		(4)		(5)	
	VWRETD$_t$		*SPX$_t$*		*SPY$_t$*		*INDU$_t$*		*DIA$_t$*	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat
PhotoPes$_{t-1}$	-0.040*	-1.705	-0.038*	-1.652	-0.038*	-1.680	-0.044**	-2.059	-0.046**	-2.140
PhotoPes$_{t-2}$	0.043*	1.685	0.038	1.515	0.036	1.463	0.032	1.355	0.029	1.223
PhotoPes$_{t-3}$	-0.024	-0.977	-0.022	-0.885	-0.025	-1.040	-0.017	-0.727	-0.019	-0.830
PhotoPes$_{t-4}$	0.038	1.605	0.033	1.391	0.035	1.520	0.039*	1.784	0.043*	1.932
PhotoPes$_{t-5}$	0.049*	1.783	0.051*	1.874	0.049*	1.837	0.047*	1.818	0.045*	1.759
Sum t-1 to t-5	0.066		0.062		0.057		0.057		0.052	
Sum t-2 to t-5	0.106		0.100		0.095		0.101		0.098	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	2.116	0.146	1.900	0.168	1.675	0.196	1.803	0.179	1.520	0.218
$\chi^2(1)$[Sum t-2 to t-5=0]	6.118**	0.013	5.591**	0.018	5.223**	0.022	6.366**	0.012	5.969**	0.015
Adj. *R*-squared	0.029		0.035		0.028		0.039		0.038	
N	3,044		3,044		3,044		3,044		3,044	

Panel B: No winsorization										
	(1)		(2)		(3)		(4)		(5)	
	VWRETD$_t$		*SPX$_t$*		*SPY$_t$*		*INDU$_t$*		*DIA$_t$*	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat
PhotoPes$_{t-1}$	-0.044*	-1.819	-0.043*	-1.792	-0.043*	-1.797	-0.049**	-2.194	-0.049**	-2.188
PhotoPes$_{t-2}$	0.061**	2.126	0.057**	2.025	0.055*	1.915	0.049*	1.842	0.044*	1.661
PhotoPes$_{t-3}$	-0.025	-0.911	-0.021	-0.781	-0.025	-0.971	-0.014	-0.545	-0.017	-0.646
PhotoPes$_{t-4}$	0.025	1.039	0.018	0.764	0.022	0.965	0.023	0.995	0.026	1.146
PhotoPes$_{t-5}$	0.053*	1.880	0.056**	2.031	0.052*	1.903	0.054**	2.029	0.051*	1.910
Sum t-1 to t-5	0.070		0.067		0.061		0.063		0.055	

Sum t-2 to t-5	0.114		0.110		0.104		0.112		0.104	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	2.195	0.138	2.094	0.148	1.750	0.186	2.014	0.156	1.684	0.194
$\chi^2(1)$[Sum t-2 to t-5=0]	6.408**	0.011	6.178**	0.013	5.525**	0.019	7.02***	0.008	6.245**	0.012
Adj. *R*-squared	0.030		0.036		0.028		0.039		0.038	
N	3,044		3,044		3,044		3,044		3,044	

Panel C: GARCH-adjusted returns

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		SPX_t		SPY_t		$INDU_t$		DIA_t	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat
$PhotoPes_{t-1}$	-0.036*	-1.948	-0.036*	-1.959	-0.035*	-1.935	-0.042**	-2.348	-0.043**	-2.372
$PhotoPes_{t-2}$	0.042**	2.203	0.041**	2.172	0.041**	2.162	0.042**	2.205	0.044**	2.273
$PhotoPes_{t-3}$	-0.042**	-2.219	-0.042**	-2.178	-0.043**	-2.267	-0.042**	-2.187	-0.043**	-2.261
$PhotoPes_{t-4}$	0.028	1.542	0.024	1.354	0.026	1.465	0.032*	1.763	0.033*	1.841
$PhotoPes_{t-5}$	0.037*	1.936	0.037**	1.963	0.036*	1.886	0.031	1.643	0.029	1.525
Sum t-1 to t-5	0.029		0.024		0.025		0.021		0.020	
Sum t-2 to t-5	0.065		0.060		0.060		0.063		0.063	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	0.673	0.412	0.528	0.468	0.498	0.480	0.370	0.543	0.330	0.566
$\chi^2(1)$[Sum t-2 to t-5=0]	3.884**	0.049	3.539*	0.060	3.392*	0.066	3.906**	0.048	3.778*	0.052
Adj. *R*-squared	0.009		0.009		0.009		0.010		0.010	
N	3,044		3,044		3,044		3,044		3,044	

Panel D: Trim 1% of the most extreme returns

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		SPX_t		SPY_t		$INDU_t$		DIA_t	
Variables	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat	β	*t*-stat
$PhotoPes_{t-1}$	-0.047**	-2.238	-0.046**	-2.221	-0.045**	-2.192	-0.047**	-2.384	-0.043**	-2.201
$PhotoPes_{t-2}$	0.040*	1.837	0.040*	1.844	0.040*	1.860	0.028	1.372	0.032	1.625
$PhotoPes_{t-3}$	-0.044*	-1.909	-0.038*	-1.660	-0.037*	-1.685	-0.034	-1.615	-0.035*	-1.703
$PhotoPes_{t-4}$	0.041**	2.062	0.037*	1.902	0.043**	2.217	0.037**	1.979	0.043**	2.326
$PhotoPes_{t-5}$	0.035	1.547	0.042*	1.883	0.032	1.480	0.035*	1.685	0.018	0.898
Sum t-1 to t-5	0.025		0.035		0.033		0.019		0.015	
Sum t-2 to t-5	0.072		0.081		0.078		0.066		0.058	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t-1 to t-5=0]	0.436	0.509	0.808	0.369	0.722	0.396	0.266	0.606	0.192	0.661
$\chi^2(1)$[Sum t-2 to t-5=0]	3.982**	0.046	5.019**	0.025	4.698**	0.030	3.804*	0.051	3.047*	0.081
Adj. *R*-squared	0.018		0.023		0.023		0.024		0.025	
N	3,012		3,012		3,012		3,013		3,012	