A Picture is Worth a Thousand Words:

Measuring Investor Sentiment by Combining Machine Learning and Photos from News*

Khaled Obaid

Kuntara Pukthuanthong†

January 2021

Abstract

By applying machine learning to accurately and cost effectively classify photos based on sentiment, we introduce a daily market-level investor sentiment index (*Photo Pessimism*) from a large sample of news photos. *Photo Pessimism* predicts market return reversal and increase in trading volume. The relation is strongest among stocks with high limits to arbitrage and during high uncertainty periods. *Photo Pessimism* subsumes pessimism embedded in text suggesting photos play an attention-grabbing role. *Photo Pessimism* from finance-focused news has over two times stronger predictive power than *Photo Pessimism* from general news.

JEL classification: C53; G10; G17

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

* We thank Dominique Badoer, Dave Berger, John Howe, Xing Huang, Inder Khurana, Heikki Lehkonen, Venky Nagar, Michael O'Doherty, John Pinfold, Richard Roll, Martin Stefan, and seminar participants at the University of Missouri-Columbia, University of Missouri-St. Louis, Missouri State University, Saint Louis University, California State University-East Bay, Midwest Finance Association Meeting (2019), Southern Finance Association Meeting (2019), GSU-RFS FinTech Conference (2020), and Financial Management Association Annual Meeting (2020) for helpful feedback. We are grateful for financial supports from Hillcrest Behavior Finance Award, Willmiller Excellence Award, and University of Missouri-Columbia and California State University-East Bay finance departments. We appreciate David Hirshleifer for his support and encouragement. An earlier version of this paper was circulated under the title "A Picture Is Worth a Thousand Words: Market Sentiment From Photos".

† Khaled Obaid (E-mail: khaled.obaid@csueastbay.edu) is from California State University-East Bay, and Kuntara Pukthuanthong (E-mail: pukthuanthongk@missouri.edu) is from the University of Missouri-Columbia.

‘A good sketch is better than a long speech.’ – *Napoleon Bonaparte*

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

Our key contributions are threefold. First, we demonstrate the importance of visual content in helping explain and predict the market risk premium. 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* is negatively related to contemporaneous market returns and positively related to future market returns. This reversal highlights that our measure has a non-informational impact on returns. Consistent with an investor sentiment proxy, we show that *PhotoPes* predicts an increase in trading volume and better explains and predicts returns on stocks with high compared to low limits to arbitrage. Our evidence is in line with Shiller’s argument that news is a good proxy for investors’ beliefs because the press has demand-side incentive to cater content to their readers’ beliefs (Shiller, 2005).

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. Using photos in news has gained dominance due to modern technology and demand for quick information. In light of studies suggesting that photos may convey emotional information more effectively than words, it is important to examine how sentiment information extracted from photos in news relates to market activities (Chemtob, Roitblat, Hamada, Muraoka, Carlson, and Bauer, 1999). However, due to the complexity of analyzing photos and cost of manually sifting through many photos, conducting such study is expensive, error prone, and tedious. Relying on surveys and crowd sourcing website (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 it 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 depend on their 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 popular 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 show that pessimism embedded in news photos subsumes pessimism embedded in news text suggesting photos play an attention-grabbing role. Although studies such as Tetlock (2007) and Garcia (2013) measure investor sentiment by performing textual analysis on news, to the best of our knowledge, no one has examined whether it is possible to capture useful and novel information about investors' beliefs from photos in news and how such information interacts with information embedded in text. We provide evidence supporting the importance of capturing pessimism embedded in news photos, especially during periods of elevated fear in news.

Machine learning has seen explosive use among academics in finance (Mullainathan and Spiess, 2017). Machine learning techniques are often used to classify textual content. For instance, Manela and Moreira (2017) use machine learning to classify words in the *Wall Street Journal* based on market volatility to construct a news implied volatility index. 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 apply machine learning on financial data to predict risk premiums (Gu, Kelly, and Xiu, 2020) and search for true risk factors (Feng, Gilio, and Xiu, 2020).

Recent developments in machine learning introduce techniques that make the task of analyzing large amounts of photos possible. In this study, we apply CNNs, a popular machine learning technique for classifying photos, to construct an investor sentiment index from a large sample of photos in the press. In particular, we use a pre-trained Google Inception v3 model. Although the model is not specifically trained to identify sentiment, it contains a lot of domain knowledge on images. In order to tailor the model for sentiment classification, we use transfer learning (i.e., we feed the pre-trained model with a sample of additional training photos specifically labeled on 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 sentiment). The model considers many aspects of the photo to make predictions, including objects, colors, and facial expressions. After the model is trained, we verify its accuracy on a test sample. We achieve 87.1% test sample accuracy. Next, we use the model to make sentiment predictions on photos from *Getty Images* (Editorial News Section), one of the largest distributors of photos to national and international news media, and *The Wall Street Journal* (*WSJ*). Using these sentiment predictions, we then construct a daily investor sentiment index, *Photo Pessimism* (*PhotoPes*), calculated as the proportion of the photos predicted to be negative on a given day.

The photos we use in this study come from two sources: *Getty Images'* Editorial News Section (sample period between 1926 and 2018) and the *WSJ* Online Archive (sample period between 2008 and 2020). Photos from these sources help us focus on the most widely distributed news that are the type of content that Shiller (2005), Tetlock (2007), and Garcia (2013) argue can play an important role in financial markets. The reason we collect photos from *Getty Images* in addition to the *WSJ* is to help us study the relation between news photos and market returns over a substantially longer time horizon.

Through various tests, we show that *PhotoPes* exhibits characteristics of an investor sentiment proxy. First, we test how *PhotoPes* is related to the CRSP value-weight (VWRETD) index, CRSP equal-weight (EWRETD) index, SPDR S&P 500 ETF Trust (SPY), SPDR Dow Jones Industrial Average ETF (DIA), and iShares Russell 2000 ETF (IWM) returns. Behavioral models, such as De Long, Shleifer, Summers, and Waldmann (1990), expect that investor sentiment should predict market return reversals. That is, when sentiment is high (low), irrational investors will increase (decrease) demand for assets driving up (down) prices away from fundamentals. Due to 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 contemporaneous market returns and positively related to future market returns especially for sample of photos from the *WSJ*. This reversal highlights that our measure has a non-informational impact on returns. In terms of magnitude for the contemporaneous relation between *PhotoPes* and market returns, the average impact of one standard deviation shift in *PhotoPes* on the VWRETD is 1.9 basis points for the *Getty Images* sample and 4.1 basis points for the *WSJ* sample. Consistent with behavioral models (e.g. De Long et al., 1990; Campbell, Grossman, and Wang, 1993), this negative relation between *PhotoPes* and contemporaneous market returns is temporary and is reversed later in the trading week. The magnitude of the reversal (over three-day period) is 2.5 basis points for the *Getty Images* sample and 8.6 basis points for the *WSJ* sample. Using the *WSJ* sample, a trading strategy based on both *PhotoPes* and pessimism embedded in news text earns 5.19% annual 5-factor alpha. Overall, the results using the sample from the *WSJ* are statistically and economically stronger than the results using *Getty Images* sample. The stronger results for the *WSJ* sample can stem from the content of the *WSJ* being more relevant to investors.

Second, we show that *PhotoPes* and investor pessimism embedded in text (*TextPes*) are significantly correlated, suggesting that there are some commonalities in the type of information they capture. However, *PhotoPes* dominates once we control for pessimism embedded in text, suggesting that *PhotoPes* contains novel and important information about investors' beliefs. Moreover, our evidence suggests that photos attract attention away from text resulting in delayed market response to pessimism embedded in news text.

Third, we strive to enhance our understanding of which type of information is more effectively or only transmitted by photos in the context of news and financial markets. We find that the coefficient on pessimism embedded in photos is three times larger during periods of elevated fear in news, while the coefficient on *TextPes* is only 1.3 times larger during periods of elevated fear in news. This evidence is consistent with photos being more effective at conveying emotionally charged news compared to text (Chemtob et al., 1999). The impact of *PhotoPes* on market return during high fear periods is most striking—the average impact of one standard deviation shift in *PhotoPes* during fear-inducing periods on the VWRETD is 10.5 basis points for the *WSJ* sample.

Fourth, we further validate *PhotoPes* as a proxy for investor sentiment by showing it has larger effect on stocks whose valuations 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 volatility. We find that *PhotoPes* has strongest effect on the returns of highest volatility portfolio and during periods of high uncertainty. In terms of magnitude, the average impact of one standard deviation shift in *PhotoPes* on the highest volatility value-weighted quintile portfolio is 4.4 basis points for the *Getty Images* sample and 10.2 basis points for the *WSJ* sample.

Fifth, we provide additional insights to help us better understand the channel through which *PhotoPes* relates to market returns. We find that the relation between *PhotoPes* and market returns is primarily driven by days when most photos are predicted to be positive. This evidence is in line with the explanation that positive sentiment has the most impact on returns because short selling impediments limit the impact of negative sentiment on returns (Miller, 1977). Moreover, we find that a shock to *PhotoPes* is able to predict increase in abnormal trading volume and this relation is driven primarily by negative sentiment. This evidence is a further validation that *PhotoPes* is a proxy for sentiment. According to behavioral models, a shock to investor sentiment should lead to increase in trading volume driven by uninformed noise traders (e.g. De Long et al., 1990; Campbell et al., 1993).

Finally, we perform a battery of robustness tests. Most notably, we show that that our main results are robust to different variable and sample construction criteria, controlling for extreme returns, and using more flexible models (neural networks).

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, Engelberg, and Gao, 2015), Twitter data (Chen, De, Hu, and Hwang, 2014), company financial reports (Loughran and McDonald, 2011; Jiang Lee, Martin, and Zhou, 2019), weather (Hirshleifer and Shumway, 2003), and sporting events (Edmans et al., 2007) to proxy for investor sentiment.[1]

More specifically, our paper extends 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 literature from communications, psychology,

[1] See Hirshleifer (2001) for review of this area.

memory, and information processing 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 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 showing that there is relevant content to financial markets in news photos.

Our paper also extends the literature on the psychology of visual stimuli to an application in finance and economics. On one hand, a known concept in psychology, Picture Superiority Effect, is the phenomenon in which pictures are more likely to be remembered than words.[2] It is based on the notion that "human memory is extremely sensitive to the symbolic modality of presentation of event information" (Yuille, 2014). On the other hand, researchers find text is better than images when you need time to make a thorough and important decision. Some find visual stimulation is not reliable and text is better than words in term of learning. Some find words are more effective than images when the goal is to change views especially a long-term view (Powell, Boomgaarden, De Swert, and de Vreese, 2015). 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, Mandisodza, Goren, and Hall, 2005; Duarte, Siegel, and Young, 2012; Halford and Hsu, 2014; Graham, Harvey, and Puri, 2016). Recently, Bazley, Cronqvist, and Mormann (2017) document how displaying financial information in red reduces investors' appetite for risk and optimism. Blankespoor, Hendricks, and Miller (2017) show how perception of management in video presentations relates to firm value. All of these papers entirely use surveys and crowd sourcing website (Amazon Mechanical Turk or MTurk) to evaluate photos. To the best of our knowledge, we are the first to use machine learning to develop our own investor sentiment proxy from news photos to predict market returns.

---

[2] See Curran (2011), Shepard (1967), McBride and Dosher (2002), Defetyer, Russo, McPartlin (2009), Whitehouse, Maybery, and Durkin (2006), and Ally, Gold, and Budson (2009).

1. Data

We begin this section by discussing the technology used to construct *PhotoPes*. Second, we discuss the sample of photos from *Getty Images* and *WSJ*. Third, we review the descriptive statistics for *PhotoPes* and how it relates to other investor sentiment measures in the literature.

1.1. Photo classification

There have been major advances in the field of computer vision that enable us to create reliable models for photo classification. We delve into some details about the sentiment prediction model 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 networks that are useful for photo classification (Krizhevsky, Sutskever, and Hinton, 2012). Recent studies use CNN photo classification models to identify solar panel installations (Yu, Wang, Majumdar, and Rajagopal, 2019), locations of slave camps (Scoles, 2019), and poverty levels in underdeveloped countries (Jean, Burke, Xie, Davis, Lobell, and Ermon, 2016) by analyzing satellite images. Similar to our application, You, Luo, Jin, and Yang (2015) employ CNN 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 invoke in investors. We build a photo classification model based on Google Inception v3 (Szegedy, Vanhoucke, Ioffe, Shlens, and Wojna, 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.[3,4]

TensorFlow, an open-source software library popular for machine learning applications developed by Google Brain Team, provides a pre-trained Google Inception v3 model. We start with a pre-trained Google Inception v3 model (trained on the ImageNet dataset) and use transfer learning to fine-tune the model for our specific application. The pre-trained Google Inception v3 has 1,000 different classes since it is trained for the

[3] 3.5% top-5 error and 17.3% top-1 error on the validation set from the ImageNet dataset. Top-1 error rate is the percentage of time the model did not produce the correct class as its top prediction by probability. Top-5 error rate is the percentage of time the model did not produce the correct class as its top 5 predictions by probability.

[4] The ImageNet dataset is available for free download at http://image-net.org/download

purpose of ILSVRC. Training photo classification models from scratch requires a huge training sample and is computationally expensive. For example, the pre-trained Google Inception v3 model is trained on 1,331,167 pre-labelled images. Transfer learning allows us to reuse the domain knowledge stored in the pre-trained 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, Hanneke, and Carbonell, 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 two probabilities—probability that photos have positive sentiment and probability that photos have negative sentiment.

We do not develop a more refined classification model due to the following reasons. First, having more refined classifications requires subjective judgement. Most textual sentiment analyses such as Loughran and McDonald (2011) come up with binary classifications for words (e.g. positive and negative words).[5] 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 on sentiment. We use *DeepSent* dataset for training.[6] *DeepSent* is a collection of photos that are collected and labeled on sentiment by You et al. (2015). The main advantage of using the *DeepSent* dataset is that the sentiment labels are verified via a crowd sourcing website (MTurk) to ensure that the labels are correct.[7] In the *DeepSent* dataset, there is a choice to select photos that have three, four, or all five participants in the MTurk survey to agree on sentiment label of the photo. For increased reliability, this study uses photos where all five MTurk survey participants agree on the sentiment label (clean labels). This restriction reduces the training sample to 882 photos.

[5] Words like "disease" and "crash" are clearly negative but saying one is more or less negative is not obvious. Even if one is more or less negative, how much is one word more or less negative than the other is subjective and depends on the context and individual's interpretation.

[6] *DeepSent* contains 1,269 labeled photos available for download from this link: https://www.cs.rochester.edu/u/qyou/Deep Sent/deepsentiment.html.

[7] The process of verifying labels is expensive and time-consuming task. Not verifying labels can lead to poorly performing models because the training data might be noisy.

The model is trained using a learning rate of 0.01 and 5,000 learning steps.[8] We set the train batch size to 100 photos. We reserve 10% of the training photos for the validation sample, and 10% for the test sample.[9,10] 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 that are 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 are never seen by the model during the training process and are used to compute a final accuracy score of 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 add regularization techniques in the model like dropout and label smoothing (Szegedy et al., 2016).

Figure 1 plots the training and validation accuracy over the training steps for the model that uses photos with clean labels. We achieve 87.1% test accuracy for the clean label model.[11] These accuracy figures are similar to those from other photo sentiment classification models in the literature. For example, Campos, Jou, and Giro-i-Nieto (2017) compare various modifications to CNN models trained on *DeepSent* and report test set accuracy figures ranging between 78.3% and 83.0%. To better test the performance of the model we use in this paper, we also calculate recall (86.2%), precision (94.3%), and F1 (90.1%).[12] Precision measures how accurate is our model at identifying positive photos out of all the photos that were predicted to be positive (important metric when the cost of a false positive is high). Recall measures how accurate is the model at identifying

---

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

[9] 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. There are some variations in the literature, but most papers assign the majority of the sample to training (usually around 80%), and the rest are divided between validation and test set. For example, Yu, Wang, Majumdar, and Rajagopal (2019) which we cite in this paper, use 77/3/20 split.

[10] Photos are randomly assigned into these sets.

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

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

positive photos out of all the positive photos in our sample (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 might not closely resemble the types of professional photos in the *Getty Images* or *WSJ* samples; thus, the model trained with *DeepSent* training set might not be able to accurately classify professional photos. To address this concern, we randomly select 100 photos using stratified sampling from our *Getty Images* sample and classify each photo in MTurk by five individuals.[13] These pictures represent the various decades in our sample. 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 for the *Getty Images* test set in the following confusion matrix:[14]

		Actual	
		Positive	Negative
Prediction	Positive	60	6
	Negative	17	17

We perform a similar exercise to the one above, except we use a test set from the *WSJ* sample:

		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 editorial news photos from *Getty Images* and the *WSJ*. The accuracy is 77.0% (76.0%), recall is 77.9% (92.8%), precision is 90.9% (77.1%), and F1 is 83.9% (84.2%) for *Getty Images* (*WSJ*). Given that we have imbalanced

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

[14] 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.

classes in this sample, it is crucial to pay attention to the F1. You et al. (2015) construct competing photo classification algorithms using *DeepSent* dataset and report accuracy, recall, precision, and F1 numbers close to the ones we report using our *Getty Images* test sample.[15] Overall, this table shows that our model trained with the *DeepSent* training set performs well at classifying news photos.

1.2. *Getty Images* sample

We select photos daily between January 1926 and June 2018 from *Getty Images* Editorial News Section. Their website describes photos in this section as follows: "*show real-world people, places, events and things and are intended to be used only in connection with events that are newsworthy or of general interest*."[16] Focusing on the Editorial News Section helps us focus on the most widely distributed news that can play an important role in financial markets. The *Getty Images* news database contains photos that are taken by photographers working for *Getty Images* in addition to photos taken by photojournalists working for various news outlets. News outlets such as *Bloomberg*, *The Los Angeles Times*, and *The Washington Post* license their photos through *Getty Images*. *Getty Images* makes its news photos database available for licensing to news organization and other interested parties. In terms of timing, which is important for our specifications, photos in the *Getty Images* database are published as early as a few minutes from the time the events occur.

Each day, we sort photos by popularity, and select the 20 most popular photos. In addition to collecting the photos, we gather information on the popularity ranking of the photo, date the photo is taken, photo identification number, and description of the photo and associated event. The popularity rank considers purchase history and number of views.[17] Ideally, we would select photos based on popularity score distribution instead of ranks. However, *Getty Images* does not make the actual popularity scores available. Moreover, we do not have access to the actual number of purchases and views for each image.[18] In Section 2.1.5, we use an alternative sample selection criterion based on the top 5% of popularity rank distribution in a given *month*. The main benefit of this alternative selection criterion is that we focus on the most influential photos We also

---

[15] See Table 1 in You et al. (2015) for a summary of their results.
[16] Source: https://www.gettyimages.com/faq/workingfiles
[17] Although the popularity score observed when we collect the photos is not known on the day the photo is first available, we argue it is a close proxy for the importance of the event and associated photo.
[18] We have made several attempts to get this information but it is strictly proprietary by Getty.

perform our analysis using photos from the *WSJ* to address the concerns associated with using *Getty Images* (e.g. finance-focused news, popularity rank).

To help filter out irrelevant photos or photos related to unimportant or frivolous events, we apply two filters to the sample of *Getty Image* photos we collect. First, we require that a given day has at least 15 total photos available. This rule is important especially earlier in the sample where there are days with sparse photos that have no apparent link to financial markets or important events. Requiring at least 15 photos be available on a given day helps us include only days when important event(s) is (are) captured by photos.[19] Second, we require that the photo description contains at least one negative or positive word according to the Loughran and McDonald (2011) dictionaries (LM). This filter removes photos with descriptions not in English and also helps increase the chance that photo is related to a news story with relevance to financial markets. For example, we postulate that when the photo description contains the word "liability" or "recession", the photo is more likely to be relevant to financial markets compared to a photo that does not contain any of the words in the Loughran and McDonald (2011) dictionaries. In total, we have 220,136 (top 20), 169,886 (top 15), and 74,044 (top 10) qualifying photos that pass the two filters and are used to construct our main variable. Below table summarizes the number of photos after each step of screening process:

	# of *Getty Image* photos after each screening process		
	Top 10	Top 15	Top 20
Original number of photos	310,433	453,869	586,832
Photos on trading day	214,276	314,242	406,837
At least 15 photos in a day	194,570	291,775	384,370
At least one word from LM dictionary	74,044	169,886	220,136

We process all qualifying photos we select from *Getty Images* into the photo classification model we train in the earlier section to get predictions on likelihood of positive or negative sentiment in the photo—all the photos from *Getty Images* are never labeled manually and are not used to train the model.[20] We use the model that is trained with clean labels (5 out of 5 survey participants agree on sentiment label). The reason we choose

[19] When we lower this requirement to 10, our main results continue to hold.
[20] Before photos are used in training or prediction, they are standardized to have the same number of pixels.

to use the photo classification model trained with clean labels is that clean labels help the model achieve high test accuracy and will lead to more reliable predictions.

Table A1 in the Appendix presents the top twenty photos in our sample that are predicted to have the highest probability to contain negative (top) or positive (bottom) sentiment. More examples from the *WSJ* and the influential photos datasets are provided in Tables A2 and A3 of the Appendix, respectively. *PhotoNeg* is the probability that the photo has negative sentiment. These predictions are made using the model with clean labels. Overall, these example photos help confirm that the photo classification model is working correctly.

1.3. *The Wall Street Journal* sample

The *WSJ*, founded in 1889, is a daily newspaper that focuses on major events and targets an audience with special interest in economics and 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 included without photos. The reason behind the lack of photos prior to September 2008 is that *WSJ* does not own the rights to the majority of the photos and is licensing the photo for a limited time from other news and media agencies such as the *Associated Press*, *Reuters*, and *Getty Images*. Once the license expire, photos are taken down. Although we believe the *WSJ* is a better source of data for our study compared to *Getty Images*, we think it is beneficial to perform our study using both samples since the *WSJ* sample is restricted to a much shorter sample period compared to *Getty Images* (sample period is between September 2008 and September 2020 for the *WSJ* compared to between January 1926 and June 2018 for *Getty Images*).[21]

Between September 2008 and September 2020, we collect headline and summary of each article, any associated photos and timestamps of when the article is published from the following 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

[21] Although our *WSJ* sample period is shorter than our sample from *Getty Images*, our *WSJ* sample is comparable with similar studies including Tetlock (2007), which has a sample period of 16 years and Da et al (2015) which has a sample period of 8 years.

general market conditions. See the top twenty photos with the highest probabilities of positive or negative sentiment from the *WSJ* in Table A2 of the Appendix.

We collect a total of 148,823 articles spanning 3,066 trading days. Out of these 148,823 articles, 20.4% of these articles have a photo licensed from *Getty Images*. We classify these photos using the clean model we discuss in Section 1.1 to get predictions on sentiment.

1.4. Variable construction

Our main variable, *PhotoPes*, is calculated as the proportion of photos predicted to be negative on a given date. For the *Getty Images* sample, we focus on the top 10 most popular, but later in the paper, we show that the main results hold if we include the top 15 or top 20 most popular photos in our variable. We weight photos from *Getty Images* based on their popularity (i.e., $\frac{1}{Popularity\ Rank_{it}}$).[22] The purpose behind weighting photos by popularity is to give more popular photos higher weights compared to less popular photos since those photos are more likely to capture events that attract more market attention. The formula for *PhotoPes* on day *t* for the *Getty Images* sample is the following:

$$PhotoPes_t = \frac{\sum_i (Neg_{it} \times \frac{1}{Popularity\ Rank_{it}})}{\sum_i \frac{1}{Popularity\ Rank_{it}}}, \tag{1}$$

where Neg_{it} is an indicator variable for whether photo *i* is predicted to have negative sentiment.[23] The denominator is the sum of the weights.

Because we include all photos from the *WSJ* that belong in “Business”, “Economy”, “Markets”, “Politics”, and “Opinion” sections, there is no popularity ranking to take into consideration for *PhotoPes*. Thus, we make slight modification when calculating *PhotoPes* for the *WSJ* sample:

$$PhotoPes_t = \frac{\sum_i (Neg_{it})}{n_t}, \tag{2}$$

[22] The most popular photo on a given day will have a popularity rank of 1, and the 10th most popular photo in a given day will have a popularity rank of 10.

[23] The probability cutoff for Neg_{it} is 50%. In Section 2.4.1, we show that our main results hold if we use different cutoffs (55%) and if we use predicted likelihood in place of the indicator variable Neg_{it}.

where Neg_{it} is an indicator variable for whether photo i is predicted to have negative sentiment. The denominator, n_t, corresponds to the number of photos in date *t*.

PhotoPes is not simply a binary measure. Although individual photos are classified as either positive or negative, *PhotoPes* is a continuous measure. Days with more negative photos have higher value for *PhotoPes* compared to days with fewer negative photos. This is the same idea behind a lot of the sentiment measures in the literature. To address the concern that sentiment in some photos can be ambiguous or neutral, we create an alternative *PhotoPes* that only includes photos with higher probability of negative or positive sentiment (>55%). This robustness test excludes photos with ambiguous or neutral sentiment and our results continue to hold (see Tables 13 and 14).

One of the goals in this paper is to better understand the type of information that is embedded in news photos by comparing pessimism embedded in photos and text. When comparing photo and text, we exclusively use the *WSJ* sample. We avoid using photo descriptions from the *Getty Images* sample because photo descriptions often fail to capture the overall tone of the article. For example, we might have a photo of a politician making a speech. The photo can convey emotions expressed by the politician and crowd. However, the photo description might simply state the name of the politician and the fact that the politician was giving a speech. To alleviate this concern, we use the *WSJ* sample and capture pessimism embedded in news text of articles and compare directly with the pessimism embedded in the photo associated with that article. Motivated by Manela and Moreira (2017) and Cong Liang, and Zhang (2018), we analyze the headlines and summary of articles.

Conceptually, the methodology we use to identify pessimism in photos includes highly non-linear relationships among features. In order to facilitate a fair comparison between pessimism embedded in photo and text, we use a comparable technique to classify the pessimism embedded in text of the headlines and summary of articles. 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 text as the pessimism score for the article (*TextNeg*).[24] The sentiment tool is based on recursive neural tensor network (RNTN) and trained on a dataset containing 215,154 phrases with fine-grained sentiment labels.[25] This RNTN model performs especially well for shorter phrases pushing the state of the art on short phrases to 85.4% accuracy (Socher, Perelygin, Wu, Chuang, Manning, Ng, and Potts, 2013). Since the headline and summary section of articles from the *WSJ* are often made up of a few short sentences, this tool is appropriate. The formula for *TextPes* is the following:

$$TextPes_t = \frac{\sum_i (TextNeg_{it})}{n_t}, \tag{3}$$

where $TextNeg_{it}$ is the pessimism score for each article from the CoreNLP model. The denominator, n_t, corresponds to the number of photos in date *t*. *PhotoPes* and *TextPes* are winsorized at 1%. Our results remain the same without winsorizing (see Tables 13 and 14).

1.5. Descriptive statistics

Table 1 Panel A gives summary statistics of the pessimism variables from the *Getty Images* (top) and the *WSJ* (bottom) samples. For the *Getty Images* sample, we have 19,243 trading days between January 1926 and June 2018 after removing days with sparse photos. On average, *PhotoPes* is 11.6%; however, this number can be difficult to interpret because of the weighting by popularity. Alternatively, without weighting by popularity (*PhotoPes'*), 30.9% of the 10 most popular qualified photos are predicted to have negative sentiment on a given day.

For the *WSJ* sample, we have 3,066 trading days between September 2008 and September 2020. On average, *PhotoPes* is 22.8% or 22.8% of photos from the *WSJ* sample are predicted to have negative sentiment on a given day. On average, *TextPes* is -0.630 indicating that on average, headlines and summary text of articles are made up of negative sentences.[26] All the variables have significant first-order autocorrelation and thus are persistent (we address autocorrelation concerns in our tests).

Table 1 Panel B reports how *PhotoPes* relates to *TextPes*. We calculate pairwise correlations and corresponding p-values for between *PhotoPes* and *TextPes*. We find that *PhotoPes* and *TextPes* are positively

[24] Source: https://stanfordnlp.github.io/CoreNLP/
[25] Scale: {"Very negative" = 0; "Negative" = -1; "Neutral" = -2; "Positive" = -3; "Very positive" = -4}
[26] Scale: {"Very negative" = 0; "Negative" = -1; "Neutral" = -2; "Positive" = -3; "Very positive" = -4}

correlated (correlation coefficient=0.073 and p-value<0.01). The positive correlation suggests that *PhotoPes* and *TextPes* are related. However, the correlation is not very high, indicating there is some information provided by photos that is distinct from the text in the headline and summary section of *WSJ* articles.

2. 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, Slovic, and Lichtenstein, 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 deviation 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 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 fundamental level.

2.1. News sentiment embedded in photos and text

We present our main results. First, we show that pessimism embedded in photos predicts market return reversal consistent with behavioral model predictions (De Long et al., 1990). Second, we show that pessimism embedded in photos subsumes pessimism embedded in text. Third, we explore which 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.

2.1.1. The impact of *PhotoPes* on market returns (*Getty Images* sample)

Table 2 presents our main results from time-series regression of market returns on *PhotoPes* from the *Getty Images* sample. The specific model we run is the following, with Newey and West (1987) t-statistics: [27]

[27] Our specification is similar to that in Tetlock (2007) and Garcia (2013). They denote the timing of their textual sentiment measures at *t-1* based on the information on day *t-1* that is ready to be printed in the evening of day *t-1*, but the information or news is publicly available in the morning on day *t*. Our *PhotoPes* is denoted at time *t* based on when photos are publicly available, which is sometimes during day *t*. That is, their "*t*" is timing of information occurrence whereas our "*t*" is based on the timing when information is public.

$$R_t = \beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2) + \beta_5 X_t + \varepsilon_t, \tag{4}$$

where R_t denotes daily log-returns on VWRETD or EWRETD, $PhotoPes_t$ is the proportion of the most popular photos that are predicted to be negative on time *t*, L_s denotes an *s*-lag operator (we set *s*=5), and X_t is a set of exogenous variables that include an intercept, day-of-the-week indicators (except for Monday), and an indicator variable for whether time *t* is in a recession period. We include a recession indicator to address the concern that our results are driven by extreme negative returns predicting future large negative returns. We continue to address concerns regarding extreme returns in Panels B, C, and D.

In the two specifications of Panel A in Table 2, $PhotoPes_t$ is constructed using the top 10 most popular photos. In the first specification, $PhotoPes_t$ is negatively related to VWRETD. The effect is significant at the 5% level. In other words, days with higher proportion of photos predicted to contain negative sentiment have lower returns, on average, compared to days with lower proportion of photos predicted to contain negative sentiment. The magnitude of the effect is economically meaningful—the average impact of one standard deviation shift in *PhotoPes* on the VWRETD is 1.9 basis points for the *Getty Images* sample, which is over 50% of the unconditional average daily returns of VWRETD (see Table A4 for descriptive statistics). We examine the lags of *PhotoPes* to determine whether a reversal of the initial effect occurs in the following 5 days. We note that the coefficient on $PhotoPes_{t-3}$ is positive and significant at the 5% level indicating a reversal occurs in the third day after the initial decline. The magnitude of the reversal over the following three days is 2.5 basis points. The Chi-Square tests show that the sum of the *PhotoPes* coefficients between lags one and three statistically significant at the 10% level. Moreover, the Chi-Square test shows that the sum of the *PhotoPes* coefficients between *t* and *t-5* is indistinguishable from zero thus suggesting that the initial decline at time *t* is reversed over the following 5 days. Overall, our baseline results suggest that *PhotoPes* has non-informational effect on returns.

As a robustness check, we also consider equal-weighting and report the results in the second specification of Panel A in Table 2. We see that $PhotoPes_t$ is negatively related to EWRETD. The effect is significant at the 5% level. The magnitude of the effect is like what is reported earlier for the VWRETD. However, the reversal pattern is stronger economically and statistically for the EWRETD result compared to the VWRETD—the magnitude of the reversal over the following three days is 3.0 basis points and the Chi-Square test shows that the sum of the *PhotoPes* coefficients between lags one and three is statistically significant at the 5% level.

Readers might wonder whether our results vary with the number of top photos we incorporate into our measure. To address this concern, we vary the number of photos we allow into our measure to include the top 15 and 20 most popular photos instead of just the top 10 most popular photos. The number of observations increases from 16,430 (top 10) to 18,513 (top 15) and 19,213 (top 20) because incorporating more photos into our measure increases the chance that we end up with photos that pass our inclusion criteria (discussed earlier) on a given day. However, the downside of including more photos into our measure is that we might be including photos that are not popular and thus might not be associated with important events.

In the third and fourth specifications of Panel A in Table 2, $PhotoPes_t$ is constructed using the top 15 most popular photos. We note that $PhotoPes_t$ is negatively related to VWRETD and EWRETD at the 5% level. We note that the coefficient on $PhotoPes_{t-1}$ is positive and significant in both specifications indicating a reversal occurs in the next day after the initial decline. We observe a similar result in the fifth and sixth specifications of Panel A in Table 2 where $PhotoPes_t$ is constructed using the top 20 most popular photos.

We are 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 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 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. Panel B of Table 2 presents the estimates of the coefficients on $PhotoPes_t$ and its lags using unit-variance VWRETD and EWRETD. We continue to find that $PhotoPes_t$ is negatively related to normalized VWRETD and EWRETD. The magnitude of the reversal over the following three days is 2.4 and 2.1 basis points for VWRETD and EWRETD, respectively. Both Chi-

Square tests show that the sum of the *PhotoPes* coefficients between lags one and three are statistically significant at the 10% level.

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 C in Table 2) and 2% (Panel D in Table 2) most extreme returns from our sample. We continue to find that $PhotoPes_t$ is negatively related to VWRETD and EWRETD after trimming the series by removing the 1% and 2% most extreme returns. However, the reversal results are statistically weak. Although the statistical power of the reversal pattern is weak for the *Getty Images* sample especially after we remove the most extreme returns from our sample, we show in the next section that the reversal pattern is much stronger once we focus on the *WSJ* sample.

2.1.2. The impact of ***PhotoPes*** on market returns (***WSJ*** sample)

In the previous section, we show *PhotoPes* constructed using the *Getty Images* sample is negatively related to contemporaneous market returns and positively related to future market returns. However, the predictability results falter once we remove extreme returns from our sample. The main advantage of the *Getty Images* sample is the long-time horizon. However, the *Getty Images* sample suffers from the following concerns which we aim to address by using the *WSJ* sample. First, although we know the date the photo is taken, we do not have exact timestamps. The lack of timestamps makes it difficult to know whether there is contemporaneous or predictive relation between *PhotoPes* and market returns. Second, although we attempt to focus on photos from *Getty Images* with a finance focus (by requiring at least one word from the Loughran and McDonald dictionary to be present in the photo description), some readers will be unconvinced by this method. The *WSJ* is one of the most widely circulated newspapers with finance-focus. Third, the popularity rank that we use to determine which photos are included concerns readers because we do not know the popularity rank with certainty at the time the photo is taken and the measure is proprietary to *Getty Images*. Finally, one of our main goals in this paper is to compare pessimism embedded in news photos and text. We are better able to compare photo and text using the *WSJ* sample because photo descriptions from the *Getty Images* sample often fail to capture the overall tone of the article.

Table 3 presents our main results from time-series regression of market returns on *PhotoPes* and controls for the *WSJ* sample. The specific model we run is the same as the one from the previous section, except we construct *PhotoPes* using photos from the *WSJ* sample. In addition to examining the relation between *PhotoPes* and market returns using the VWRETD and EWRETD, we also use returns on the SPDR S&P 500 ETF Trust (SPY), SPDR Dow Jones Industrial Average ETF (DIA), and iShares Russell 2000 ETF (IWM) using the data until September 2020. The data to compute VWRETD and EWRETD is available only until 2019.

In Panel A in Table 3, $PhotoPes_t$ is negatively related to contemporaneous market returns regardless of what index we use. The relation between $PhotoPes_t$ and IWM is the largest in terms of magnitude (6.3 basis points, which is 50% larger than the unconditional average of IWM).[28] While the relation between $PhotoPes_t$ and EWRETD is the smallest in magnitude (3.9 basis points, which is 75% of the unconditional average of EWRETD). The relation is statistically significant at the 5% level for DIA and IWM and at the 10% level for VWRETD, EWRETD, and SPY. We examine the lags of *PhotoPes* to determine whether a reversal of the initial decline occurs in the following 5 days. We note that the reversal for the *WSJ* sample is concentrated between lags three and five. The magnitude of the reversal between three and five days after the photo is published ranges between 7.2 and 12.7 basis points. The Chi-Square tests show that the reversal is statistically significant at the 1% level for SPY, DIA, and IWM and at the 5% level for VWRETD and EWRETD. Moreover, the Chi-Square test shows that the sum of the *PhotoPes* coefficients between *t* and *t-5* is indistinguishable from zero thus suggesting that the initial decline at time *t* is reversed over the following 5 days.

Panel B of Table 3 presents the estimates of the coefficients on $PhotoPes_t$ and its lags using unit-variance VWRETD, EWRETD, SPY, DIA and IWM. We continue to find that $PhotoPes_t$ is negatively related to normalized returns in all specifications. Moreover, we continue to find that the reversal is concentrated

[28] See Table A4 for descriptive statistics on returns.

between lags three and five. The Chi-Square tests show that the reversal is statistically significant at the 5% level for SPY, DIA, and IWM and at the 10% level for VWRETD and EWRETD.

Panels C and D of Table 3 present our results from time-series regression of market returns on *PhotoPes* after we remove 1% (Panel C) and 2% (Panel D) of most extreme returns from our sample. Unlike our results using the *Getty Images* sample, we find clear evidence for return reversal after removing most extreme returns using the *WSJ* sample. In all specifications in Panels C and D, we find that $PhotoPes_t$ is negatively related to contemporaneous market returns. The magnitude of the reversal between three and five days after the photo is published ranges between 6.5 and 12.8 basis points. The Chi-Square tests show that the reversal is statistically significant for all specifications in Panels C and D.

Overall, we confirm our earlier results showing that *PhotoPes* predicts short-term return reversal using the sample of news photos from the *WSJ*. The results using the *WSJ* sample are statistically and economically stronger than the results using the *Getty Images* sample. The stronger results using the *WSJ* sample can stem from using more finance-focused news as opposed to general news like in the *Getty Images* sample.

2.1.3. *PhotoPes* and sentiment embedded in text (*WSJ* sample)

In this section, we examine whether photos and news text are complements or substitutes. More specifically, are photos used by news media to enhance the sentiment embedded in text (complements) or are photos used to convey alternative information to text (substitutes)? If photos enhance the pessimism embedded in text, the pessimism-returns relation becomes stronger when pessimism embedded in photos is consistent with pessimism embedded in text. Moreover, we are interested in examining whether photos can play an attention-grabbing role. If photos play an attention-grabbing role, we expect *PhotoPes* to dominate after adding pessimism embedded in text to the model.

Table 4 examines how *PhotoPes*, *TextPes*, and their interaction relates with market returns. As discussed earlier, we strictly use the *WSJ* sample for this analysis because we only have the photos from *Getty Images* without news articles associated with them. We run the following regression:

$$R_t = \beta_1 PhotoPes_t + \beta_2 TextPes_t + \beta_3 Interaction_t + \beta_4 L_s(PhotoPes_t) + \quad (5)$$

$$\beta_5 L_s(Text_t) + \beta_6 L_s(R_t) + \beta_7 L_s(R_t^2) + \beta_8 X_t + \varepsilon_t,$$

where $TextPes_t$ is average pessimism score from the CoreNLP model for all articles at time *t* (defined in the Section 1.4). All other variables are defined earlier.

After controlling for pessimism from news text in Table 4, we continue to observe $PhotoPes_t$ is negatively related to market returns for VWRETD, EWRETD, SPY, DIA, and IWM. The coefficients on $PhotoPes_t$ are significant at the 1% level for IWM and EWRETD, and at the 5% for VWRETD, SPY, and DIA. We note strong reversal for *PhotoPes* between lags three and five for all specifications at the 5% for VWRETD and EWRETD, and at the 1% level for SPY, DIA, and IWM.

The coefficients on $TextPes_t$ are negative, but not statistically significant at the 10% level except in fourth specification.[29] Adding $TextPes_t$ in the model results in larger coefficients on $PhotoPes_t$ in all our specifications. Given that *PhotoPes* dominates after controlling for *TextPes*, our results suggest that photos in news play an attention-grabbing role. Although coefficient on $TextPes_t$ is not significant, the magnitude of *TextPes* between *t* and *t-2* is negative and statistically significant in all five specifications suggesting that photos attract attention away from text and thus markets take more time to reflect pessimism embedded in text compared to pessimism embedded in photos. For example, in the first specification of Table 4, the coefficient on $TextPes_t$ is negative but not significant. However, the magnitude of *TextPes* between date *t* and *t-2* is -11.0 basis points and significant at the 5% level.

Next, we shift our focus to the interaction term of *PhotoPes* and *TextPes*. News media can use photos to reinforce pessimism embedded in text. If pessimism embedded in photos enhances pessimism embedded in news text, we expect to find that the coefficient on the interaction term, $Interaction_t$, is negative. On the other hand, news media can use photos to convey an alternative dimension of investor sentiment that is not already reflected in the text. If news media uses photos and text as substitutes, we expect to find that the coefficient on the interaction term, $Interaction_t$, is positive. In all five specifications, the coefficient on $Interaction_t$ is positive and significant supporting the substitution hypothesis. Because investor sentiment is

[29] Later in this section, we show results from time-series regression of market returns on *TextPes* and controls without controlling for *PhotoPes*. We confirm that *TextPes* (without controlling for *PhotoPes*) can predict return reversal as documented by earlier studies (Tetlock, 2007; Garcia, 2013).

multidimensional, this finding is important because it suggests that news photos contain a new dimension of investor sentiment that is not already captured in text (Zhou, 2018).

Although *PhotoPes* dominates once we control for *TextPes*, we corroborate Tetlcok (2007) and Garcia (2013) by showing that once we remove *PhotoPes* from our model, *TextPes* alone predicts return reversal. Table A6 in the appendix reports results from time-series regression of market returns (VWRETD, EWRETD, SPY, DIA, and IWM) on *TextPes* and controls (without controlling for *PhotoPes*). Because we do not require photos for this regression, we are able to expand our *WSJ* sample (from the *WSJ* online achieve) to start in December 1997 instead of September 2008. In all five specifications, $TextPes_t$ is negatively related to market returns. The initial decline is followed by a reversal between lags three and five. The magnitude of the reversal is between 9.9 and 17.0 basis points and statistically significant at the 1% level for DIA and IWM and at the 5% for VWRETD, EWRETD, and SPY.

Overall, we find that pessimism embedded in news photos subsumes pessimism embedded in text, suggesting photos detract investors away from text leading to delayed market reaction to pessimism embedded in text. Moreover, our evidence suggests that news media uses pessimism embedded in text and photos as substitutes and that pessimism embedded in photos contains a new dimension of investor sentiment that is not reflected in text.

2.1.4. Which information is more effectively transmitted by photos?

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

We proxy for periods of elevated fear in news 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 and emotions. TRMI uses proprietary dictionary to classify tone for different emotions and events in day *t* and quantify it into 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 photo relates to market returns by level of fear in news, we run the following regression:

$$R_t = (F_t)[\beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 TextPes_t + \beta_4 L_s(TextPes_t) + \beta_5 Interaction_t + \beta_6 L_s(R_t) + \beta_7 L_s(R_t^2)] + (1 - F_t)[\gamma_1 PhotoPes_t + \gamma_2 L_s(PhotoPes_t) + \gamma_3 TextPes_t + \gamma_4 L_s(TextPes_t) + \gamma_5 Interaction_t + \gamma_6 L_s(R_t) + \gamma_7 L_s(R_t^2)] + \beta_8 X_t + \varepsilon_t, \quad (6)$$

where T_t is a dummy variable that takes a value of one if day *t* is above median fear score (computed as the average TRMI score of the following topics: fear and gloom). All other variables are defined earlier.

Table 5 examines how the effect of *PhotoPes* and *TextPes* on market returns varies by fear in news. In all five specification, $PhotoPes_t$ is negatively related to market returns during both high and low fear periods. However, the magnitude is much larger during high fear periods—for example, the average impact of one standard deviation shift in *PhotoPes* on the VWRETD is 10.5 basis points during high fear period and only 3.5 basis points for the low fear period. The same cannot be said for *TextPes*. In all the specifications, once we control for *PhotoPes*, $TextPes_t$ is negative but not significant. In terms of magnitude, the average impact of one standard deviation shift in *TextPes* on the VWRETD is 2.8 basis points during high fear period and 2.1 basis points for the low fear period.

Overall, the coefficient on pessimism embedded in photos is three times larger during periods of elevated fear in news, while the coefficient on *TextPes* is only 1.3 times larger during periods of elevated fear in news. This evidence is consistent with photos being more effective at conveying emotionally charged news compared to text (Chemtob et al., 1999).

2.1.5. Influential photos

We explore how pessimisms embedded in photos relates to market returns in days when news photos are especially popular. Although we focus on the top 10 most popular photos in each day in our baseline results for the *Getty Images* sample, it is possible that we are including insignificant photos especially during uneventful days. To help narrow our sample to only days with influential photos, we select photos in the top 5% of the popularity rank distribution in each month. The approach of using a cutoff in the popularity rank distribution helps us avoid including insignificant photos during uneventful days.

Our sample of the top 5% photos contains 232,531 photos that occur over 10,577 days between 1926 and 2018. We expect to find that the relationship between *PhotoPes* and market returns is stronger in this sample compared to in the main sample we use in the paper. To test the relation between pessimism and market returns using the sample of influential photos, we run the following regressions:

$$R_i = \beta_1 PhotoPes_t + \beta_2 L_s(R_t) + \beta_3 L_s(R_t^2) + \beta_4 X_t + \varepsilon_t, \tag{7}$$

where $i = t$, (*t+1 to t+18*), *(t+19 to t+20), and (t to t+20)*. All other variables are defined earlier. Because we do not have photos daily in this sample, we do not include lags of pessimism embedded in photos in the models. Instead, we check for reversal by regressing contemporaneous and future market returns on proxies for pessimism embedded in photo.

In the first specification of Table 6, $PhotoPes_t$ is negatively related to VWRETD. The effect is significant at the 5% level. The magnitude of the effect is economically meaningful— the average impact of one standard deviation shift in *PhotoPes* on the VWRETD is 2.7 basis points (for comparison, the magnitude is about 40% larger compared to our baseline results). In the second specification, there is no significant reversal in the following 18 days. In the third specification, reversal is concentrated on days *t+19* and *t+20*. In comparison to our baseline results, the reversal in this set of photos takes longer (reversal occurs within the first three days in our baseline results). The reason behind the slow reversal might be because it takes longer for the impact of most influential photos to wind down. In the fourth specification, we show whether the reversal is complete by regressing cumulative returns between days *t* and *t+20* on *PhotoPes*. We find that $PhotoPes_t$ is not significantly related to cumulative VWRETD between day *t* and *t+20* suggesting that the initial effect is reversed in the following 20 days. In specifications (5) through (8), we replace VWRETD with EWRETD and reach a similar conclusion.

2.1.6. 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 bidask bounce or day-of-the-week effects, we use residuals from *recursively* regressing *PhotoPes* or *TextPes* on lagged returns and day-of-the-week dummies available at time *t-1*, denoted $PhotoPes^{\perp}$ or $TextPes^{\perp}$ respectively. The strategies require investors, each day, to either invest in SPY or the risk-free asset depending on lagged pessimism in news. The first strategy is based on news pessimism embedded in photos—following days in which $PhotoPes^{\perp}$ is above its historical mean (expanding), we invest in the SPY at market close of day *t+2* and sell on market close three days later (*t+5*).[30,31] 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 market close of day *t+2* and sell on market close three days later (*t+5*). The third strategy involves pessimism from both, text and photo—following days in which $PhotoPes^{\perp}$ and $TextPes^{\perp}$ are above their historical means, we invest in the SPY at market close of day *t+2* and sell on market close three days later (*t+5*).

In Panel A of Table 7, we report mean and standard deviation of daily excess returns (in percentages) and 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 2,386, 2,234, and 1,723 times, respectively. On average, the first strategy (based on $PhotoPes^{\perp}$) generates 5.3 basis points in daily excess returns, which is higher than the 4.7 basis points in excess returns for the buy-and-hold SPY strategy, and the 4.8 basis points in daily excess returns for the $TextPes^{\perp}$ strategy. In addition, the $PhotoPes^{\perp}$ strategy generates a 0.044 Sharpe ratio compared to the 0.036 Sharpe ratio from the buy-and-hold SPY strategy and the 0.040 Sharpe ratio from the $TextPes^{\perp}$ strategy. 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*), Carhart (1997) momentum factor (*MOM*) and Da, Liu, Schaumburg (2014) short-run reversal factor (*ST_Rev*). Results are reported in Panel B of Table 7. We find that the combined strategy generates a positive and significant 5-factor alpha of 5.19% per annum (2 basis points per day).

---

[30] We pick this investment rule because we find that *PhotoPes* and *TextPes* have the strongest predictability between lags three and five in our baseline results from the *WSJ* sample.

[31] Expanding means that we use expanding window and not fixed rolling window.

According to Figure 2, by the end of our sample, the first, second, and third strategies earned \$3.95, \$3.45, and \$5.04 in excess cumulative returns for a \$1 invested in the beginning of our sample. 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 in the beginning of our sample.

2.2. Validation of *PhotoPes*

2.2.1. Limits to arbitrage

Next, we focus on how limits to arbitrage affect the relation between *PhotoPes* and 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, Hirshelifer and Ma, 2020). De Long et al. (1990) suggest that investor sentiment should have the strongest effect on stocks that are hardest to arbitrage; while D'Avolio (2002) finds that arbitrage is riskier and costlier for riskier stocks than safer stocks. We focus on total volatility-sorted portfolios to test the prediction that *PhotoPes* should have the biggest impact on difficult-to-value or riskiest stocks. Wurgler and Zhuravskaya (2002) use total volatility of stocks to proxy for cost of arbitrage.

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

$$R_t = \beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2) + \beta_5 X_t + \varepsilon_t, \quad (8)$$

where R_t denotes the equal-weighted (Panel A) or value-weighted (Panel B) daily return on quintile portfolios sorted on monthly volatility (Sigma_t) or the difference (H-L) between the highest and lowest volatility sorted portfolios. Stocks in the CRSP universe are sorted into quintiles based on daily return volatility over the past month. All other variables are defined earlier.

The first five specifications of Panel A in Table 8 show that $PhotoPes_t$, constructed using the *Getty Images* sample, is negatively related to all equal-weighted portfolios sorted on volatility, but statistical significance for the highest volatility portfolio (1% level) is higher than the statistical significance for the lowest volatility portfolio (10% level). The strength of the relation between *PhotoPes* and portfolio returns is monotonically increasing with volatility. The magnitude of the effect is also larger for highest volatility portfolio compared to lowest volatility portfolio (over three times larger). The average impact of one standard deviation shift in *PhotoPes* on the highest and lowest equal-weighted volatility portfolio is 3.2 (significant at the 1% level) and 0.9 (significant at the 10% level) basis points, respectively. We examine the lags of *PhotoPes* to determine whether a reversal of the initial effect occurs in the following 5 days. There is clear and significant reversal pattern between lags one and three for all quintile portfolios. The magnitude of the reversal between lags one and three for the highest and lowest equal-weighted volatility portfolio is 3.4 (significant at the 10% level) and 1.8 (significant at the 5% level) basis points, respectively.

To test if the effect of sentiment on stock returns is stronger for the highest volatility stocks compared to lowest volatility stocks, we regress the difference between the returns on highest and lowest equal-weighted volatility quintile portfolios (H-L) on *PhotoPes* and controls. We show the results from this regression in the sixth specification of Panel A in Table 8. The coefficient on $PhotoPes_t$ is negative and significant at the 1% level.

Similar to Panel A, the first five specifications of Panel B in Table 8 show that $PhotoPes_t$ is negatively related to all value-weighted portfolios sorted on volatility. Compared to the equal-weighted portfolios sorted on volatility, the magnitude of the effect for all value-weighted volatility-sorted portfolios is larger. For example, the magnitude for the highest value-weighted volatility portfolio is 4.4 compared to 3.2 basis points for the highest equal-weighted volatility portfolio. In terms of reversal, there is significant reversal for specifications three to five. However, the statistical significance of the reversal for the highest two volatility sorted portfolio is weak. In the sixth specification of Panel B, we regress the difference between the returns on highest and lowest value-weighted volatility quintile portfolios (H-L) on *PhotoPes* and controls and find that the coefficient on $PhotoPes_t$ is -3.2 basis points and significant at the 5% level.

Table 9 reports results from analysis looking at how *PhotoPes* relates to stock returns differently depending on limits to arbitrage for the *WSJ* sample. We run the same regressions in Table 8, except *PhotoPes* is constructed using the *WSJ* sample.

The first specification of Panels A and B in Table 9 show that $PhotoPes_t$, constructed using the *WSJ* sample, is negatively related to equal- and value-weighted highest volatility portfolios at the 5% level. The magnitude of the effect is larger for the *WSJ* sample compared to the *Getty Images* sample—for example, the average impact of one standard deviation shift in *PhotoPes* on the highest value-weighted volatility portfolio is 10.2 basis points (compared to 4.4 basis points from the *Getty Images* sample). The fifth specification of Panels A and B in Table 9 show that $PhotoPes_t$ is not significantly related to the equal- and value-weighted lowest volatility portfolios. The sixth specification of Panels A and B in Table 9 show that $PhotoPes_t$ is significantly related to the equal- and value-weighted H-L volatility portfolios, which supports that the effect of sentiment on stock returns is stronger for the highest volatility stocks compared to lowest volatility stocks.

2.2.2. Uncertainty

We examine how the relation between pessimism embedded in photos varies depending on uncertainty in the market. Individuals process information differently depending on their emotions and some emotions can be associated with uncertainty (Smith and Ellsworth, 1985; Tiedens and Linton, 2001). Thus, we speculate that during periods of high market uncertainty, sentiment induced mispricing will be most severe. NVIX is a news implied volatility index constructed from textual analysis of front-page news articles in the *WSJ*. It is developed by Maneala and Moreira (2017) to capture aggregate uncertainty. NVIX is available in monthly frequency between 1889 and 2016 (covering most of our sample period).[32] We run the following regression using the *Getty Images* sample to differentiate the effect of *PhotoPes* on market returns by periods of high and low implied volatility:[33]

$$R_t = (N_t)[\beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2)] + \quad (9)$$

$$(1 - N_t)[\gamma_1 PhotoPes_t + \gamma_2 L_s(PhotoPes_t) + \gamma_3 L_s(R_t) + \gamma_4 L_s(R_t^2)] + \beta_5 X_t + \varepsilon_t,$$

where N_t is a dummy variable that takes a value of one if date *t* is in a month that has above median NVIX value, and zero otherwise. All other variables are defined earlier.

The first specification of Table 10 examines how the effect of *PhotoPes* on VWRETD varies during times of low and high market implied volatility. The β_1 coefficient measures the effect of $PhotoPes_t$ on market

[32] Source: http://apps.olin.wustl.edu/faculty/manela/data.html

[33] We perform this analysis using the *Getty Images* sample instead of the *WSJ* sample because we want to incorporate as many high implied volatility periods or recessions as possible in our analysis.

returns during periods of high market uncertainty. β_1 is negative and significant at the 5% level. The magnitude is larger than what is observed in Table 2 (3.5 basis points compared to 1.9 basis points in Table 2). The $\gamma_1$coefficient measures the effect of $PhotoPes_t$ on market returns during periods of low market uncertainty. The point estimate on γ_1 is only 0.1 basis points and statistically indistinguishable from zero. There is significant reversal between lags one and three during high NVIX periods. The magnitude of the reversal in specification one is 5.3 basis points and significant at the 5% level.

The second specification of Table 10 looks at how the relation between *PhotoPes* and EWRETD varies during times of low and high market implied volatility. Overall, we observe very similar results to the first specification.

2.3. Mechanisms of ***PhotoPes***

2.3.1. Non-linearity of ***PhotoPes***

Considering short selling constraints, we speculate that positive sentiment has a bigger effect on market returns compared to negative sentiment (Stambaugh, Yu, and Yuan, 2012). Miller (1977) argues that impediments to short selling limit the ability of investors to correct overpricing.[34] Table 11 Panel A examines whether relation between *PhotoPes* and market returns is symmetric between positive and negative sentiment for the *Getty Images* sample. To help answer this question, we run the following regression:

$$R_t = \beta_1 PhotoPes_{rank5_t} + \beta_2 PhotoPes_{rank4_t} + \beta_3 PhotoPes_{rank3_t} + \beta_4 PhotoPes_{rank2_t} + \beta_5 PhotoPes_{rank1_t} + \beta_6 L_s(R_t) + \beta_7 L_s(R_t^2) + \beta_8 X_t + \varepsilon_t, \tag{10}$$

where $PhotoPes_{rank5_t}$ is the highest quintile $PhotoPes_t$ (most photos are predicted to be negative) and $PhotoPes_{rank1_t}$ is the lowest quintile $PhotoPes_t$ (fewest photos are predicted to be negative). All other variables are defined earlier.

According to Table 11 Panel A, the relation between *PhotoPes* and market returns is primarily driven by days when most photos are positive (fewest photos are negative). The magnitude of the coefficient on

[34] Several studies empirically examine the relation between short selling constraints and overpricing, including Figlewski (1981), and Jones and Lamont (2002).

$PhotoPes_{rank1_t}$ is over four times the size of the coefficient on $PhotoPes_t$ reported in Table 2 and significant for both specifications (VWRETD and EWRETD). $PhotoPes_{rank3_t}$ is significant at 5% for VWRETD but marginally significant (at 10%) for EWRETD. $PhotoPes_{rank2_t}$ is marginally significant at 10% for VWRETD and insignificant for EWRETD. Taken together, $PhotoPes_{rank1_t}$is the most statistically significant. None of $PhotoPes_{rank4_t}$ and $PhotoPes_{rank5_t}$ are significant suggesting the market does not react strongly when most photos are negative. We conjecture this is due to short selling constraints on the day when most photos are negative.

Turning to the WSJ sample, Table 12 Panel A reports results from analysis looking at whether the relation between *PhotoPes* and market returns is symmetric between positive and negative sentiment. The magnitude of the coefficient on $PhotoPes_{rank1_t}$ is over two times the size of the coefficient on $PhotoPes_t$ reported for our baseline results using the *WSJ* sample in Table 3 and significant at the 10% level. None of the coefficients on $PhotoPes_{rank2_t}$ to $PhotoPes_{rank5_t}$ suggesting market react mostly to positive photos.

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

Another channel through which *PhotoPes* can impact market activity is trading volume. In Table 11 Panel B, we test 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 De Long et al. (1990) and Campbell, Grossman, and Wang (1993) predict that a shock to sentiment indicates that noise investors will want to buy or sell. As the market absorbs these orders, we should see that a change in *PhotoPes* is related to increase in trading volume.

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

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

where V_t denotes the log of 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 time trend. These are not the intended effects we attempt to explain using *PhotoPes*. Gallant, Rossi, and Tauchen (1992) and Garcia (2013) use a similar method to remove time trend and other effects in daily volume data. To test how *PhotoPes* is related to trading volume, we run the following models:

$$\bar{V}_t = \beta_1 PhotoPes_t^{high} + \beta_2 PhotoPes_t^{low} + \beta_3 L_s(PhotoPes_t^{high}) + \beta_4 L_s(PhotoPes_t^{low}) + \beta_5 L_s(R_t) + \beta_6 L_s(R_t^2) + \varepsilon_t, \quad (12)$$

where $\bar{V}_t$ is standardized (zero mean, unit variance) residual from Equation (11), $PhotoPes_t^{high}$is *PhotoPes* above 0.5, and $PhotoPes_t^{low}$ is *PhotoPes* equal or below 0.5.[35] All other variables are defined earlier. The coefficient on key variables of interests, $PhotoPes_t^{high}$ and $PhotoPes_t^{low}$, are positive and but not significant, suggesting that there is no significant positive relation between negative or positive sentiment from photos and contemporaneous abnormal trading volume. However, the coefficients on $PhotoPes_{t-1}^{high}$, $PhotoPes_{t-1}^{low}$, $PhotoPes_{t-4}^{high}$ and $PhotoPes_{t-5}^{high}$ are positive and significant suggesting that pessimism from photos is able to predict increase in abnormal trading volume. In terms of magnitude, a one standard deviation increase in *PhotoPes* is associated with moving future trading volume between 0.023 and 0.040 standard deviations. The relation between *PhotoPes* and future trading volume is primarily driven by negative sentiment.

Table 12 Panel B reports results from looking at how *PhotoPes* is related to trading volume for the *WSJ* sample. The coefficients on $PhotoPes_{t-1}^{high}$ and $PhotoPes_{t-1}^{low}$, are positive and significant suggesting that pessimism from photos is able to predict increase in abnormal trading volume.

2.4. Robustness

2.4.1. Variable construction

Table 13 reports results from the main regression for the *Getty Images* sample in 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 Panel A of Table 13, we adjust the cutoff for Neg_{it} from 50% to 55%. The regression specification is exactly like the one in specifications one and two of Table 2, except the cutoffs are adjusted. We also reduce the number of photos requirement from 15 to 10 because increasing the cutoff reduces the

[35] We split *PhotoPes* according to Tetlock (2007) who argues, if media pessimism presents investor sentiment, unusually low or high pessimism should increase trading volume.

number of eligible photos. Looking at the two specifications of Panel A, we note that the return reversal results documented earlier still hold with similar magnitudes. In Panel B of Table 13 we modify *PhotoPes* by not winsorizing the variable. We note that the results are similar.

Cautious readers might wonder why we do not use predicted likelihood of negative sentiment when calculating *PhotoPes* as it is a continuous measure and more refined than the binary variable. The reason is that our model is not trained to identify the intensity of pessimism in the photo. Instead, our model is trained to simply classify photos as either negative or positive. Thus, the predicted likelihood is the model's confidence about the prediction and not directly related to the intensity of the sentiment in the photo. However, our baseline results are not dependent on using dummy variable instead of predicted likelihood. In Panel C of Table 13, we modify *PhotoPes* by replacing the negative sentiment dummy variable, Neg_{it}, with predicted likelihood of negative sentiment and continue to observe similar results.

Table 14 repeats the analysis in Table 13 for the *WSJ* sample. We continue to find that our baseline results are robust to various modification to constructing *PhotoPes*.

In untabulated results, we find that our baseline results in Table 2 are not dependent on war photos. We identify war-related photos by searching photo descriptions following Verdickt (2020): at least one mention of "military" or "war" and at least one mention of these 12 words (risk, fear, concern, uncertain, threat, tension, invasion, army, start, beginning, battle and outbreak). Although we do not include photos of sporting events in our analysis, we confirm that including such photos does not impact our baseline results in Table 3. Edmans et al (2007) show how the outcome of soccer matches has asset pricing implications.

2.4.2. 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 seem more appropriate to avoid the in-sample over-fitting issue. Moreover, 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), and many others, we evaluate the out-of-sample explanatory 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 mean squared prediction error (MSPE) for the explanatory regression relative to the historical average benchmark:

$$R^2_{OOS} = 1 - \frac{\sum_{t=p}^{T-1}(R_t - \hat{R}_t)^2}{\sum_{t=p}^{T-1}(R_t - \bar{R}_t)^2}, \tag{13}$$

where $\hat{R}_t$ is the fitted values from a predictive regression of VWRETD or EWRETD on one-period lag of *PhotoPes* estimated recursively with information available at time *t-1*. $\bar{R}$ denotes the historical average benchmark estimated through period *t-1* from the constant expected return model ($R_t = \alpha + \varepsilon_{t+1}$). We only perform the out-of-sample analysis using the *Getty Images* sample because of the larger sample size compared to the *WSJ* sample. More specifically, we use the data from 1926 to 1969 as the initial estimation period so that the prediction evaluation period spans between 1970 and 2018. We also try different initial estimation windows ending in 1979, 1989, and 1999 and the results remain qualitatively similar. The length of the initial in-sample estimation period balances having enough observations for precisely estimating the initial parameters with 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 ($-\infty, 1$]. If $R^2_{OOS} > 0$, it means that the predicted $\hat{R}_t$ outperforms the historical average $\bar{R}_t$ in term 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 explanatory regression MSPE tests against the one-sided (upper-tail) alternative hypothesis that the historical average MSPE is greater than the explanatory regression MSPE, corresponding to H_0: $R^2_{OOS} \leq 0$ against H_A : $R^2_{OOS} > 0$. Clark and West (2007) show that the test has an asymptotically standard normal distribution when comparing forecasts from the nested models. Intuitively, under the null hypothesis that the constant expected return model generates the data, the predictive regression model produces a noisier forecast than the historical average benchmark because it estimates slope parameters with zero population values. We thus expect the benchmark model's MSPE to be smaller than the predictive regression model's MSPE under the null. 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 Panel A of Table A5 show R^2_{OOS} is positive and significant according to the CW-test. We try different starting points: 1970, 1980, 1990, 2000. Across all our specifications, the R^2_{OOS} ranges between 0.858% to 2.309% and they are all significant. We repeat the analysis using photos from the *WSJ* sample and report these results in Panel B of Table A5. Overall, we find similar results.

Curious readers might question whether our results are driven by the many transformations that are applied to construct our key pessimism variables since linear models are not well-suited for capturing dependencies in extreme variables. To alleviate these concerns, we examine the usefulness of *PhotoPes* at predicting market returns using a more flexible model like neural network regressions. Motivated by Gu et al. (2020), we use a shallow neural network with one hidden layer and 32 neurons. We use the standard least squares as the loss function and ReLU activation function at all nodes. We train the network using stochastic gradient descent in 100 epochs and 0.9 learning rate. The input layer contains five lags of *PhotoPes* (constructed using photos from *Getty Images*), five lags of market returns, five lags of squared market returns, day of the week dummies (except for Monday), and a recession dummy. The output is market returns ($VWRETD_t$ or $EWRETD_t$). We use the period between 1926 and 1969 to train the model, and the remainder of the sample to make predictions. The model is estimated recursively every year.

To evaluate the usefulness of *PhotoPes* at predicting market returns, we report R^2_{OOS} (out-of-sample R^2) in percent and associated p-values using the MSPE-Adjusted statistic in Clark and West (2007). In Panel C of Table A5, we show positive and significant R^2_{OOS} of 1.366% and 0.481% for VWRETD and EWRETD, respectively.[36] Our *WSJ* sample is not long enough to appropriately train the neural network for predicting returns and thus we only apply the flexible model on the *PhotoPes* constructed using *Getty Images*.

[36] We do not present estimated parameters from this analysis as a main result because our goal is to introduce *PhotoPes* and thus economic interpretation, which is a challenge for any machine learning technique, is a first order for our paper.

3. Conclusion

In this paper, we use a machine learning technique to extract information from large sample of photos in the press 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 reversal. This return reversal pattern is consistent with sentiment induced temporary mispricing (De Long et al., 1990; Campbell et al., 1993). This relation is strongest for stocks that are harder-to-arbitrage and accentuated during high uncertainty periods. Moreover, we show that *PhotoPes* can predict increase in market trading volume.

Second, we show that pessimism embedded in news photos serves an attention-grabbing role and subsumes pessimism embedded in news text. Our evidence shows that *PhotoPes* is especially useful for understanding and predicting the market risk premium during periods of elevated-fear in news.

Third, we demonstrate the benefit of using cutting-edge photo classification techniques to study how the information in a large sample of news photos is relevant in context of financial markets. Relying on surveys and crowd sourcing website (Amazon Mechanical Turk, MTurk) to evaluate photos has shortcomings that make it infeasible to study large amounts of photos. With the continued focus on artificial intelligence, the techniques for analyzing photos are bound to grow in popularity and improve. This 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 has implications in both corporate finance and asset pricing.

Future researchers should attempt to advance the machine learning technique for photos classification to help bridge the gap with text classification models to better capture sentiment and other important content embedded in news photos. We believe that applying more advanced photo classification models will yield more predictive power and provide further economic insights.

References

Ally, B. A., Gold, C. A., and Budson, A. E., 2009. The Picture Superiority Effect in Patients With Alzheimer's Disease and Mild Cognitive Impairment. *Neuropsychologia* 47(2), 595–598.

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

Bazley, W. J., Cronqvist, H., and Mormann, M. M., 2017. In the Red: The Effects of Color on Investment Behavior. Working paper.

Blankespoor, E., Hendricks, B. E., and Miller, G. S., 2017. Perceptions and Price: Evidence from CEO Presentations at IPO Roadshows. *Journal of Accounting Research* 55(2), 275-327.

Buehlmaier, M. and Whited, T., 2018. Are Financial Constraints Priced? Evidence from Textual Analysis. *Review of Financial Studies* 31(7), 2693-2728.

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

Campbell, J. Y., Grossman, S. J., and Wang, J., 1993. Trading Volume and Serial Correlation in Stock Returns. *Quarterly Journal of Economics* 108(4), 905-939.

Campbell, J. Y., and 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., and Giro-i-Nieto, X., 2017. From Pixels to Sentiment: Fine-tuning CNNs for Visual Sentiment Prediction. *Image and Vision Computing* 65, 15-22.

Carhart, MM., 1997. On persistence in mutual fund performance. *Journal of Finance* 52(1), 57-82.

Chemtob, C., Roitblat, H., Hamada, R., Muraoka, M., Carlson, J., and 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, 309-326.

Chen, H., De, P., Hu, Y., and 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., and Ma, L., 2020. The Causal Effect of Limits to Arbitrage on Asset Pricing Anomalies. *Journal of Finance*, forthcoming.

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

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

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

Connolly, R. and 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. and 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.

Curran, T. and Doyle, J., 2011. Picture superiority doubly dissociates the ERP correlates of recollection and familiarity. *Journal of Cognitive Neuroscience* 23(5), 1247-1262.

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

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

Da, Z., Liu, Q., and Schaumburg, E., 2014. A Closer Look at the Short-Term Return Reversal. *Management Science* 60(3), 658-674.

Defetyer, M. A., Russo, R., and McPartlin, P. L., 2009. The picture superiority effect in recognition memory: a developmental study using the response signal procedure. *Cognitive Development* 24(3), 265-273.

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

Duarte, J., Siegel, S., and 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., and 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.

Feng, G., Giglio, S., and Xiu, D., 2020. Taming the Factor Zoo: A Test of New Factors. *Journal of Finance* 75(3), 1327-1370.

Figlewski, S., 1981. The informational effects of restrictions on short sales: some empirical evidence. *Journal of Financial and Quantitative Analysis* 16(4), 463-476.

Fischhoff, B., Slovic, P., and 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., and 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.

Gentzkow, M. and Shapiro, J. M., 2010. What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica* 78(1), 35-71.

Goyal, A., and Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21(4), 1455-508.

Graham, J. R., Harvey, C. R., and Puri, M., 2017. A Corporate Beauty Contest. *Management Science* 63(9), 2773-3145.

Gu, S., Kelly, B., and 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. and 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., and Ermon, S., 2016. Combining Satellite Imagery and Machine Learning to Predict Poverty. *Science* 353(6301), 790-794.

Jiang, F., Lee, J., Martin, X., and Zhou, G., 2019. Manager Sentiment and Stock Returns. *Journal of Financial Economics* 132(1), 126-149

Kelly, B., and Pruitt, S., 2013. Market expectations in the cross-section of present values. *Journal of Finance* 68(5), 1721-1756.

Kozak, S., Nagel, S., and Santosh, S., 2018. Interpreting Factor Models. *Journal of Finance* 73(3), 1183-1223.

Krizhevsky, A., Sutskever, I., and Hinton, G. E., 2012. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 1, 1097-1105.

Loughran, T. and 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. and Moreira, A., 2017. News implied volatility and disaster concerns. *Journal of Financial Economics* 123(1), 137-162.

McBride, D. M. and Dosher, B.A., 2002. A comparison of conscious and automatic memory processes for picture and word stimuli: a process dissociation analysis. *Consciousness and Cognition* 11(3), 423-460.

Miller, E. M., 1977. Risk, uncertainty and divergence of opinion. *Journal of Finance* 32(4), 1151-1168.

Mullainathan, S. and Shleifer, A., 2005. The market for news. *American Economic Review* 95(4), 1031-1053.

Mullainathan, S. and Spiess, J., 2017. Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives,* 31(2), 87-106.

Newey, W. K. and West, K. D., 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55(3), 703-708.

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., and de Vreese, C. H., 2015. A Clearer Picture: The Contribution of Visuals and Text to Framing Effects. *Journal of Communication* 65(6), 997-1017.

Scoles, S., 2019. Researchers spy signs of slavery from space. *Science* 363(6429).

Shepard, R.N., 1967. Recognition memory for words, sentences, and pictures. *Journal of Learning and Verbal Behavior* 6, 156-163.

Shiller, R. J., 2005. Irrational exuberance. Princeton University Press.

Shleifer, A. and 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, Robert M., Dillman, Don A., Eltinge, John L., Little, Roderick J. A., 163–78. New York, NY: Wiley.
Smith, C. and Ellsworth, P., 1985. Patterns of Cognitive Appraisal in Emotion. *Journal of Personality and Social Psychology* 48(4), 813-838.
Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., and 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.
Stambaugh, R. F., Yu, J., and Yuan, Y., 2012. The Short of It: Investor Sentiment and Anomalies. *Journal of Financial Economics* 104(2), 288-302.
Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z., 2016. Rethinking the Inception Architecture for Computer Vision. *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.
Tiedens, L. and Linton, S., 2001. Judgment Under Emotional Certainty and Uncertainty: The Effects of Specific Emotions on Information Processing. *Journal of Personality and Social Psychology* 81(6), 973-988.
Todorov, A., Mandisodza, A. N., Goren, A., and Hall, C. C., 2005. Inferences of Competence from Faces Predict Election Outcomes. *Science* 308(5728), 1623-1626.
Tversky, A. and Kahneman, D., 1983. Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. *Psychological Review* 90(4), 293-315.
Verdickt, G., 2020. The Effect of War Risk on Managerial and Investor Behavior: Evidence from the Brussels Stock Exchange in the Pre-1914 Era. *The Journal of Economic History*, 80(3), 629-669.
Vissing-Jorgensen, A., 2003. Perspectives on Behavioral Finance: Does Irrationality Disappear with Wealth? Evidence from Expectations and Actions. *NBER Macroeconomics Annual* 18, 139-208.
Whitehouse, A. J., Maybery, M.T., Durkin, K., 2006. The development of the picture-superiority effect. *British Journal of Developmental Psychology* 24(4), 767-773.
Wurgler, J. and Zhuravskaya, E., 2002. Does arbitrage flatten demand curves for stocks*? Journal of Business* 75(4), 583-608.
Yang, L., Hanneke, S., and Carbonell, J., 2013. A theory of transfer learning with applications to active learning. *Machine Learning* 90, 161-189.
You, Q., Luo, J., Jin, H., and Yang, J., 2015. Robust Image Sentiment Analysis using Progressively Trained and Domain Transferred Deep Networks. *The Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI).*
Yu, J., Wang, Z., Majumdar, A., and Rajagopal, R., 2019. DeepSolar: A Machine Learning Framework to Efficiently Construct Solar Deployment Database in the United States. *Joule* 2(12), 2605-2617.
Yuille, J. C., 2014. Imagery, Memory and Cognition (PLE: Memory): Essays in Honor of Allan Paivio. Psychology Press.
Zhou, G., 2018. Measuring Investor Sentiment. *Annual Review of Financial Economics* 10, 239-259.

Table 1: Summary Statistics

Panel A reports summary statistics for Photo Pessimism (*PhotoPes*) and Text Pessimism (*TextPes*) and Panel B reports 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. For the *Getty Images* sample, we use the top 10 most popular photos each date, weight photos by the popularity ranking, and use the photo description for calculating *TextPes*. Non-weighted versions of the variables are denoted by " ′ ". The *Getty Images* sample period is between January 1926 and June 2018. For the *WSJ* sample, we use the headline and summary of each article for calculating *TextPes* and use news photos that belong in articles from the following sections: "Business", "Economy", "Markets", "Politics", and "Opinion". The *WSJ* sample period is between September 2008 and September 2020. *PhotoPes* and *TextPes* are winsorized at 1%. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Summary Statistics of Sentiment Variables

			Getty (1926 to 2018)				
Variable	N	Mean	Median	P25	P75	Std dev	AC(1)
PhotoPes	19,243	0.116	0.057	0.000	0.171	0.148	0.083***
PhotoPes′	19,243	0.309	0.250	0.000	0.500	0.300	0.106***

			WSJ (2008 to 2020)				
Variable	N	Mean	Median	P25	P75	Std dev	AC(1)
PhotoPes	3,066	0.228	0.222	0.179	0.270	0.082	0.108***
TextPes	3,066	-0.630	-0.639	-0.707	-0.556	0.126	0.530***

Panel B: Correlations between Sentiment Variables

	WSJ
	PhotoPes
TextPes	0.073*** <0.01

Table 2: The Impact of *PhotoPes* on Market Returns using *Getty Images* Sample

Panels A, B, C, and D of this table report β_1 and β_2 from the following time-series regression:

$$R_t = \beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2) + \beta_5 X_t + \varepsilon_t,$$

where $PhotoPes_t$ is calculated as the proportion of the 10 (15 or 20) most popular photos from *Getty Images* predicted to be negative on time *t*, L_s denotes an s-lag operator (s is set to 5), 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 1% and standardized to have zero mean and unit variance. In Panel A, R_t is log daily return on CRSP value-weight ($VWRETD_t$) index or CRSP equal-weight ($EWRETD_t$) index. In Panel B, R_t is normalized log-returns on the $VWRETD_t$ or $EWRETD_t$. We normalize returns by dividing them by the estimates volatility from the GARCH (1,1) model. In Panels C and D, we remove the 1% and 2% most extreme returns in our sample, respectively. The sample period is between January 1926 and June 2018. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Newey and West (1987) standard errors are applied to compute t-stat for all specifications.

Panel A: Log-Returns

	Top 10				Top 15				Top 20			
	(1)		(2)		(3)		(4)		(5)		(6)	
	$VWRETD_t$		$EWRETD_t$		$VWRETD_t$		$EWRETD_t$		$VWRETD_t$		$EWRETD_t$	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.019**	-2.181	-0.019**	-2.259	-0.019**	-2.313	-0.020**	-2.395	-0.016*	-1.951	-0.017**	-2.089
$PhotoPes_{t-1}$	0.010	1.156	0.017*	1.905	0.014*	1.645	0.017**	2.036	0.016**	1.921	0.021**	2.381
$PhotoPes_{t-2}$	-0.002	-0.214	-0.002	-0.258	-0.005	-0.564	-0.002	-0.286	-0.006	-0.704	-0.004	-0.443
$PhotoPes_{t-3}$	0.017**	2.034	0.015*	1.813	0.008	1.029	0.007	0.853	0.006	0.770	0.004	0.546
$PhotoPes_{t-4}$	-0.012	-1.392	-0.013	-1.509	-0.008	-0.918	-0.008	-0.992	-0.008	-0.895	-0.008	-0.972
$PhotoPes_{t-5}$	-0.010	-1.186	-0.002	-0.218	-0.011	-1.317	-0.003	-0.355	-0.011	-1.364	-0.004	-0.446
Sum t to t-5	-0.016		-0.004		-0.021		-0.009		-0.019		-0.008	
Sum t-1 to t-3	0.025*		0.030**		0.017		0.022*		0.016		0.021*	
	$\chi^2(1)$	*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 to t-5=0]	0.812	0.368	0.087	0.769	1.499	0.221	0.323	0.570	1.233	0.267	0.192	0.661
$\chi^2(1)$[Sum t-1 to t-3=0]	3.383*	0.066	4.622**	0.032	1.855	0.173	2.956*	0.086	1.697	0.193	3.001*	0.083
Adj R-sq	0.016		0.051		0.013		0.049		0.012		0.048	
N	16,430		16,430		18,513		18,513		19,213		19,213	

	Panel B: GARCH-Adjusted Returns				Panel C: Trim 1% Most Extreme Returns				Panel D: Trim 2% Most Extreme Returns			
	Top 10				Top 10				Top 10			
	(1)		(2)		(1)		(2)		(1)		(2)	
	$VWRETD_t$		$EWRETD_t$		$VWRETD_t$		$EWRETD_t$		$VWRETD_t$		$EWRETD_t$	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.016**	-2.051	-0.019***	-2.587	-0.015**	-2.074	-0.015**	-2.161	-0.013**	-1.974	-0.013**	-2.042
$PhotoPes_{t-1}$	0.014*	1.801	0.016**	2.117	0.006	0.861	0.005	0.778	0.004	0.667	0.008	1.266
$PhotoPes_{t-2}$	-0.002	-0.240	-0.006	-0.740	-0.004	-0.519	0.001	0.109	0.002	0.240	0.001	0.086
$PhotoPes_{t-3}$	0.012	1.613	0.011	1.501	0.009	1.272	0.010	1.451	0.009	1.397	0.008	1.261
$PhotoPes_{t-4}$	-0.014*	-1.859	-0.014*	-1.887	-0.010	-1.500	-0.014**	-2.181	-0.012*	-1.861	-0.009	-1.580

PhotoPes$_{t-5}$	-0.005	-0.719	-0.003	-0.380	-0.007	-0.971	-0.005	-0.688	-0.006	-0.964	-0.004	-0.590
Sum t to t-5	-0.011		-0.015		-0.021		-0.018		-0.016		-0.009	
Sum t-1 to t-3	0.024*		0.021*		0.011		0.016		0.015		0.017	
	$\chi^2(1)$	*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 to t-5=0]	0.489	0.484	0.891	0.345	1.866	0.172	1.586	0.208	1.336	0.248	0.558	0.455
$\chi^2(1)$[Sum t-1 to t-3=0]	3.817*	0.051	3.103*	0.078	1.023	0.312	2.201	0.138	2.061	0.151	2.611	0.106
Adj R-sq	0.023		0.081		0.011		0.049		0.011		0.051	
N	16,430		16,430		16,260		16,258		16,092		16,093	

Table 3: The Impact of *PhotoPes* on Market Returns using *WSJ* Sample

Panels A, B, C, and D of this table report β_1 and β_2 from the following time-series regression:

$$R_t = \beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2) + \beta_5 X_t + \varepsilon_t,$$

where $PhotoPes_t$ is calculated as the proportion of photos from the *WSJ* predicted to be negative on time *t*, L_s denotes an s-lag operator (s is set to 5), 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 1% and standardized to have zero mean and unit variance. We use news photos that belong in articles from the following sections: "Business", "Economy", "Markets", "Politics", and "Opinion". R_t is log daily return on CRSP value-weight ($VWRETD_t$) index, CRSP equal-weight ($EWRETD_t$) index, SPDR S&P 500 ETF Trust (SPY_t), SPDR Dow Jones Industrial Average ETF (DIA_t), or iShares Russell 2000 ETF (IWM_t). In Panel B, R_t is normalized log-returns. We normalize returns by dividing them by the estimates volatility from the GARCH (1,1) model. In Panels C and D, we remove the 1% and 2% most extreme returns in our sample, respectively. The sample period is between September 2008 and September 2020. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Newey and West (1987) standard errors are applied to compute t-stat for all specifications.

Panel A: Log-Returns

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		$EWRETD_t$		SPY_t		DIA_t		IWM_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.041*	-1.802	-0.039*	-1.789	-0.042*	-1.823	-0.046**	-2.165	-0.063**	-2.145
$PhotoPes_{t-1}$	0.061**	2.324	0.062**	2.542	0.045*	1.681	0.038	1.507	0.056*	1.702
$PhotoPes_{t-2}$	-0.057**	-2.356	-0.048**	-2.087	-0.034	-1.474	-0.030	-1.368	-0.027	-0.849
$PhotoPes_{t-3}$	0.032	1.389	0.025	1.145	0.025	1.112	0.033	1.515	0.024	0.795
$PhotoPes_{t-4}$	0.031	1.223	0.028	1.149	0.055**	2.059	0.054**	2.089	0.081**	2.472
$PhotoPes_{t-5}$	0.023	0.952	0.019	0.839	0.030	1.299	0.021	0.954	0.022	0.732
Sum t to t-5	0.049		0.047		0.079*		0.070		0.093	
Sum t-3 to t-5	0.086**		0.072**		0.110***		0.108***		0.127***	
	$\chi^2(1)$	*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 to t-5=0]	1.243	0.265	1.308	0.253	2.902*	0.088	2.534	0.111	2.664	0.103
$\chi^2(1)$[Sum t-3 to t-5=0]	5.283**	0.022	4.238**	0.040	8.33***	0.004	8.864***	0.003	7.323***	0.007
Adj R-sq	0.030		0.021		0.030		0.040		0.032	
N	2,854		2,854		3,043		3,043		3,043	

Panel B: GARCH-Adjusted Returns

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		$EWRETD_t$		SPY_t		DIA_t		IWM_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.036*	-1.867	-0.038**	-1.975	-0.034*	-1.858	-0.041**	-2.262	-0.039**	-2.070
$PhotoPes_{t-1}$	0.052***	2.624	0.056***	2.795	0.041**	2.168	0.044**	2.268	0.033*	1.752
$PhotoPes_{t-2}$	-0.050**	-2.530	-0.047**	-2.397	-0.045**	-2.387	-0.046**	-2.417	-0.033*	-1.759
$PhotoPes_{t-3}$	0.033*	1.733	0.031	1.623	0.027	1.501	0.034*	1.907	0.021	1.117
$PhotoPes_{t-4}$	0.028	1.425	0.033*	1.665	0.037*	1.930	0.031	1.596	0.052***	2.751
$PhotoPes_{t-5}$	-0.003	-0.151	-0.003	-0.142	-0.002	-0.117	-0.003	-0.167	-0.005	-0.237

Sum t to t-5	0.024		0.032		0.024		0.019		0.029	
Sum t-3 to t-5	0.058*		0.061*		0.062**		0.062**		0.068**	
	$\chi^2(1)$	*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 to t-5=0]	0.419	0.517	0.691	0.406	0.416	0.519	0.268	0.605	0.615	0.433
$\chi^2(1)$[Sum t-3 to t-5=0]	3.517*	0.061	3.743*	0.053	4.453**	0.035	4.58**	0.032	5.338**	0.021
Adj R-sq	0.012		0.015		0.010		0.012		0.011	
N	2,854		2,854		3,043		3,043		3,043	

Panel C: Trim 1% Most Extreme Returns

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		$EWRETD_t$		SPY_t		DIA_t		IWM_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.042**	-1.984	-0.041**	-2.078	-0.045**	-2.179	-0.042**	-2.169	-0.063**	-2.290
$PhotoPes_{t-1}$	0.051**	2.244	0.055**	2.545	0.036*	1.690	0.032	1.574	0.039	1.425
$PhotoPes_{t-2}$	-0.047**	-2.082	-0.041**	-1.997	-0.039*	-1.806	-0.039**	-1.963	-0.052*	-1.833
$PhotoPes_{t-3}$	0.044**	2.148	0.033*	1.675	0.042**	2.128	0.043**	2.361	0.040	1.474
$PhotoPes_{t-4}$	0.017	0.770	0.021	1.025	0.031	1.442	0.018	0.903	0.061**	2.152
$PhotoPes_{t-5}$	0.018	0.810	0.019	0.899	0.027	1.268	0.022	1.127	0.027	0.950
Sum t to t-5	0.041		0.046		0.052		0.034		0.052	
Sum t-3 to t-5	0.079**		0.073**		0.100***		0.083***		0.128***	
	$\chi^2(1)$	*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 to t-5=0]	0.940	0.332	1.569	0.210	1.639	0.200	0.776	0.378	1.043	0.307
$\chi^2(1)$[Sum t-3 to t-5=0]	5.125**	0.024	5.165**	0.023	8.822***	0.003	7.133***	0.008	8.753***	0.003
Adj R-sq	0.023		0.031		0.027		0.027		0.023	
N	2,824		2,824		3,011		3,011		3,011	

Panel D: Trim 2% Most Extreme Returns

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		$EWRETD_t$		SPY_t		DIA_t		IWM_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.036*	-1.898	-0.048***	-2.629	-0.034*	-1.756	-0.042**	-2.336	-0.064**	-2.469
$PhotoPes_{t-1}$	0.054***	2.595	0.051***	2.783	0.036*	1.823	0.042**	2.182	0.064**	2.427
$PhotoPes_{t-2}$	-0.041**	-2.079	-0.035*	-1.906	-0.044**	-2.251	-0.051***	-2.665	-0.048*	-1.864
$PhotoPes_{t-3}$	0.042**	2.205	0.034*	1.882	0.044**	2.348	0.045**	2.574	0.026	1.004
$PhotoPes_{t-4}$	0.013	0.603	0.027	1.384	0.024	1.221	0.020	1.095	0.054**	1.981
$PhotoPes_{t-5}$	0.016	0.756	0.004	0.199	0.020	1.035	0.012	0.658	-0.002	-0.067
Sum t to t-5	0.048		0.033		0.046		0.026		0.030	
Sum t-3 to t-5	0.071**		0.065**		0.088***		0.077***		0.078*	
	$\chi^2(1)$	*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 to t-5=0]	1.464	0.226	0.882	0.348	1.573	0.210	0.519	0.471	0.411	0.522
$\chi^2(1)$[Sum t-3 to t-5=0]	4.783**	0.029	4.730**	0.030	8.04***	0.005	7.095***	0.008	3.775*	0.052
Adj R-sq	0.016		0.019		0.018		0.017		0.017	
N	2,794		2,794		2,980		2,980		2,979	

Table 4: Pessimism from Photo and Text

This table reports $\beta_1, \beta_2, \beta_3, \beta_4$ and β_5 from the following time-series regression:

$$R_t = \beta_1 PhotoPes_t + \beta_2 TextPes_t + \beta_3 Interaction_t + \beta_4 L_s(PhotoPes_t) + \beta_5 L_s(TextPes_t) + \beta_6 L_s(R_t) + \beta_7 L_s(R_t^2) + \beta_8 X_t + \varepsilon_t,$$

where R_t is log daily return on CRSP value-weight ($VWRETD_t$) index, CRSP equal-weight ($EWRETD_t$) index, SPDR S&P 500 ETF Trust (SPY_t), SPDR Dow Jones Industrial Average ETF (DIA_t), or iShares Russell 2000 ETF (IWM_t). $PhotoPes_t$ is calculated as the proportion of photos from the *WSJ* predicted to be negative on 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. $Interaction_t$ is the interaction between $PhotoPes_t$ and $TextPes_t$. L_s denotes an s-lag operator (s is set to 5) 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 in articles from the following sections: "Business", "Economy", "Markets", "Politics", and "Opinion". *PhotoPes* and *TextPes* are winsorized at 1% and standardized to have zero mean and unit variance. The sample period is between September 2008 and September 2020. Newey and West (1987) standard errors are applied to compute t-stat. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		$EWRETD_t$		SPY_t		DIA_t		IWM_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.052**	-2.453	-0.053***	-2.635	-0.050**	-2.265	-0.055***	-2.589	-0.080***	-2.813
$TextPes_t$	-0.031	-0.925	-0.021	-0.663	-0.054	-1.588	-0.054*	-1.666	-0.038	-0.885
$Interaction_t$	0.038**	2.009	0.045**	2.440	0.035*	1.859	0.036**	2.017	0.062**	2.506
$PhotoPes_{t-1}$	0.059**	2.372	0.060***	2.606	0.047*	1.795	0.041	1.621	0.060*	1.850
$PhotoPes_{t-2}$	-0.050**	-2.107	-0.040*	-1.804	-0.030	-1.296	-0.026	-1.206	-0.021	-0.644
$PhotoPes_{t-3}$	0.037	1.610	0.031	1.416	0.029	1.267	0.036*	1.668	0.029	0.956
$PhotoPes_{t-4}$	0.025	1.025	0.023	0.986	0.048*	1.847	0.049*	1.891	0.075**	2.303
$PhotoPes_{t-5}$	0.017	0.708	0.013	0.596	0.026	1.139	0.018	0.818	0.017	0.565
$TextPes_{t-1}$	-0.048	-1.363	-0.042	-1.343	-0.049	-1.398	-0.048	-1.456	-0.082**	-1.978
$TextPes_{t-2}$	-0.031	-0.827	-0.037	-1.066	-0.023	-0.616	-0.011	-0.313	-0.034	-0.732
$TextPes_{t-3}$	-0.018	-0.558	-0.012	-0.389	-0.025	-0.763	-0.025	-0.800	-0.020	-0.488
$TextPes_{t-4}$	0.066*	1.832	0.048	1.428	0.089**	2.465	0.087***	2.605	0.100**	2.310
$TextPes_{t-5}$	0.060	1.623	0.069**	1.975	0.045	1.233	0.035	1.034	0.065	1.440
Sum t to t-5 *PhotoPes*	0.036		0.034		0.070		0.063		0.080	
Sum t to t-2 *PhotoPes*	-0.043		-0.033		-0.033		-0.040		-0.041	
Sum t-3 to t-5 *PhotoPes*	0.079**		0.067**		0.103***		0.103***		0.121***	
Sum t to t-5 *TextPes*	-0.002		0.005		-0.017		-0.016		-0.009	
Sum t to t-2 *TextPes*	-0.110**		-0.100**		-0.126***		-0.113***		-0.154***	
Sum t-3 to t-5 *TextPes*	0.108**		0.105**		0.109**		0.097**		0.145***	
	$\chi^2(1)$	*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 to t-5 *PhotoPes*=0]	0.724	0.395	0.751	0.386	2.382	0.123	2.053	0.152	1.988	0.159
$\chi^2(1)$[Sum t to t-2 *PhotoPes*=0]	1.486	0.223	1.007	0.316	0.771	0.380	1.224	0.269	0.713	0.399
$\chi^2(1)$[Sum t-3 to t-5 *PhotoPes*=0]	4.715**	0.030	3.918**	0.048	7.44***	0.006	8.052***	0.005	6.599**	0.010
$\chi^2(1)$[Sum t to t-5 *TextPes*=0]	0.006	0.940	0.020	0.886	0.326	0.568	0.321	0.571	0.046	0.830
$\chi^2(1)$[Sum t to t-2 *TextPes*=0]	6.016**	0.014	5.891**	0.015	8.319***	0.004	7.103***	0.008	7.638***	0.006

$\chi^2(1)$[Sum t-3 to t-5 *TextPes*=0]	5.632**	0.018	6.514**	0.011	5.294**	0.021	4.319**	0.038	6.797***	0.009
Adj R-sq	0.035		0.027		0.036		0.045		0.038	
N	2,854		2,854		3,043		3,043		3,043	

Table 5: The Impact of Pessimism Embedded in Photos and Text on Market Returns during Elevated Fear Periods

This table reports $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \gamma_1, \gamma_2, \gamma_3, \gamma_4$, and γ_5 from the following time-series regression:

$$R_t = (F_t)[\beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 TextPes_t + \beta_4 L_s(TextPes_t) + \beta_5 Interaction_t + \beta_6 L_s(R_t) + \beta_7 L_s(R_t^2)] + (1 - F_t)[\gamma_1 PhotoPes_t + \gamma_2 L_s(PhotoPes_t) + \gamma_3 TextPes_t + \gamma_4 L_s(TextPes_t) + \gamma_5 Interaction_t + \gamma_6 L_s(R_t) + \gamma_7 L_s(R_t^2)] + \beta_8 X_t + \varepsilon_t,$$

where F_t is an indicator variable for whether period *t* is above median fear score (computed as the average TRMI score of the following topics: fear and gloom), R_t is log daily return on CRSP value-weight ($VWRETD_t$) index, CRSP equal-weight ($EWRETD_t$) index, SPDR S&P 500 ETF Trust (SPY_t), SPDR Dow Jones Industrial Average ETF (DIA_t), or iShares Russell 2000 ETF (IWM_t). $PhotoPes_t$ is calculated as the proportion of photos from the *WSJ* predicted to be negative on 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. $Interaction_t$ is the interaction between $PhotoPes_t$ and $TextPes_t$. L_s denotes an s-lag operator (s is set to 5) 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 in articles from the following sections: "Business", "Economy", "Markets", "Politics", and "Opinion". *PhotoPes* and *TextPes* are winsorized at 1% and standardized to have zero mean and unit variance. The sample period is between September 2008 and September 2020. Newey and West (1987) standard errors are applied to compute t-stat. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$VWRETD_t$				$EWRETD_t$				SPY_t			
	(1)				(2)				(3)			
	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
$PhotoPes_t$	-0.105**	-2.195	-0.035*	-1.849	-0.106**	-2.405	-0.036**	-1.978	-0.090*	-1.759	-0.038**	-1.979
$TextPes_t$	-0.028	-0.427	-0.021	-0.781	-0.023	-0.373	-0.016	-0.624	-0.052	-0.812	-0.038	-1.387
$Interaction_t$	0.089**	2.200	0.009	0.559	0.097**	2.475	0.016	0.968	0.076*	1.879	0.005	0.289
$PhotoPes_{t-1}$	0.067	1.192	0.053**	2.434	0.073	1.397	0.050**	2.474	0.055	0.909	0.049**	2.259
$PhotoPes_{t-2}$	-0.031	-0.641	-0.061***	-2.722	-0.026	-0.575	-0.049**	-2.263	0.011	0.239	-0.059***	-2.706
$PhotoPes_{t-3}$	0.100**	2.018	0.006	0.283	0.097**	2.059	-0.001	-0.044	0.082*	1.688	0.003	0.123
$PhotoPes_{t-4}$	0.045	0.862	0.008	0.357	0.043	0.856	0.006	0.298	0.077	1.414	0.028	1.262
$PhotoPes_{t-5}$	0.049	1.018	-0.004	-0.179	0.039	0.850	-0.001	-0.047	0.062	1.347	0.000	0.020
$TextPes_{t-1}$	-0.044	-0.637	-0.052*	-1.855	-0.041	-0.673	-0.041	-1.576	-0.067	-0.982	-0.056**	-2.021
$TextPes_{t-2}$	-0.098	-1.316	0.008	0.309	-0.110	-1.544	0.005	0.188	-0.079	-1.090	0.018	0.692
$TextPes_{t-3}$	-0.043	-0.638	-0.005	-0.173	-0.024	-0.385	-0.006	-0.226	-0.073	-1.107	-0.001	-0.050
$TextPes_{t-4}$	0.098	1.360	0.039	1.368	0.066	0.972	0.033	1.264	0.130*	1.847	0.052*	1.848
$TextPes_{t-5}$	0.092	1.381	0.046	1.545	0.108*	1.742	0.050*	1.762	0.050	0.782	0.046	1.547
Sum t to t-5 *PhotoPes*	0.125		-0.033		0.120		-0.031		0.197		-0.017	
Sum t to t-2 *PhotoPes*	-0.069		-0.043		-0.059		-0.035		-0.024		-0.048	
Sum t-3 to t-5 *PhotoPes*	0.194**		0.010		0.179**		0.004		0.221***		0.031	
Sum t to t-5 *TextPes*	-0.023		0.015		-0.024		0.025		-0.091		0.021	
Sum t to t-2 *TextPes*	-0.170**		-0.065*		-0.174**		-0.052		-0.198**		-0.076**	
Sum t-3 to t-5 *TextPes*	0.147*		0.080**		0.150**		0.077**		0.107		0.097**	
	$\chi^2(1)$	*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 to t-5 *PhotoPes*=0]	1.504	0.220	0.628	0.428	1.618	0.203	0.610	0.435	3.421*	0.064	0.188	0.665

$\chi^2(1)$[Sum t to t-2 *PhotoPes*=0]	0.725	0.395	1.615	0.204	0.612	0.434	1.197	0.274	0.078	0.780	2.125	0.145
$\chi^2(1)$[Sum t-3 to t-5 *PhotoPes*=0]	6.178**	0.013	0.090	0.764	6.159**	0.013	0.016	0.898	7.522***	0.006	0.804	0.370
$\chi^2(1)$[Sum t to t-5 *TextPes*=0]	0.110	0.741	0.285	0.593	0.128	0.720	0.809	0.368	1.288	0.256	0.555	0.456
$\chi^2(1)$[Sum t to t-2 *TextPes*=0]	4.184**	0.041	2.951*	0.086	4.99**	0.026	2.228	0.136	6.176**	0.013	3.884**	0.049
$\chi^2(1)$[Sum t-3 to t-5 *TextPes*=0]	2.977*	0.085	4.039**	0.045	3.985**	0.046	4.235**	0.040	1.435	0.231	5.954**	0.015
Adj R-sq	0.045				0.035				0.049			
N	2,854				2,854				3,043			

	DIA_t				IWM_t			
	(4)				(5)			
	F_t=Fear Period				F_t=Fear Period			
Variables	β	t-stat	γ	t-stat	β	t-stat	γ	t-stat
$PhotoPes_t$	-0.099**	-2.032	-0.044**	-2.341	-0.139**	-2.202	-0.066**	-2.546
$TextPes_t$	-0.044	-0.738	-0.038	-1.449	-0.032	-0.399	-0.029	-0.801
Interaction	0.083**	2.134	0.003	0.162	0.133**	2.567	0.015	0.617
$PhotoPes_{t-1}$	0.039	0.682	0.052**	2.502	0.076	1.068	0.052*	1.874
$PhotoPes_{t-2}$	0.007	0.159	-0.053**	-2.573	0.043	0.692	-0.060**	-2.062
$PhotoPes_{t-3}$	0.081*	1.748	0.015	0.697	0.120*	1.939	-0.017	-0.567
$PhotoPes_{t-4}$	0.070	1.316	0.031	1.519	0.113*	1.706	0.047	1.582
$PhotoPes_{t-5}$	0.041	0.936	-0.002	-0.123	0.051	0.883	-0.004	-0.122
$TextPes_{t-1}$	-0.077	-1.196	-0.049*	-1.868	-0.122	-1.571	-0.073**	-2.050
$TextPes_{t-2}$	-0.059	-0.881	0.025	1.039	-0.125	-1.392	0.025	0.725
$TextPes_{t-3}$	-0.067	-1.075	-0.008	-0.326	-0.069	-0.846	0.007	0.199
$TextPes_{t-4}$	0.131**	2.062	0.046*	1.727	0.155*	1.795	0.068*	1.891
$TextPes_{t-5}$	0.036	0.591	0.038	1.368	0.093	1.164	0.054	1.379
Sum t to t-5 *PhotoPes*	0.139		-0.001		0.264**		-0.048	
Sum t to t-2 *PhotoPes*	-0.053		-0.045		-0.020		-0.074*	
Sum t-3 to t-5 *PhotoPes*	0.192**		0.044		0.284***		0.026	
Sum t to t-5 *TextPes*	-0.080		0.014		-0.100		0.052	
Sum t to t-2 *TextPes*	-0.180**		-0.062*		-0.279***		-0.077	
Sum t-3 to t-5 *TextPes*	0.100		0.076**		0.179*		0.129**	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t to t-5 *PhotoPes*=0]	1.899	0.168	0.004	0.952	4.051**	0.044	0.811	0.368
$\chi^2(1)$[Sum t to t-2 *PhotoPes*=0]	0.412	0.521	2.076	0.150	0.035	0.852	2.834*	0.092
$\chi^2(1)$[Sum t-3 to t-5 *PhotoPes*=0]	6.188**	0.013	1.783	0.182	8.64***	0.003	0.311	0.577
$\chi^2(1)$[Sum t to t-5 *TextPes*=0]	1.148	0.284	0.238	0.626	1.036	0.309	1.830	0.176
$\chi^2(1)$[Sum t to t-2 *TextPes*=0]	5.504**	0.019	2.842*	0.092	7.438***	0.006	2.381	0.123
$\chi^2(1)$[Sum t-3 to t-5 *TextPes*=0]	1.242	0.265	4.008**	0.045	3.004*	0.083	6.171**	0.013
Adj R-sq	0.064				0.050			
N	3,043				3,043			

Table 6: Pessimism in Influential Photos and Market Returns

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

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

where R_i is log daily return on CRSP value-weight ($VWRETD_t$) index or CRSP equal-weight ($EWRETD_t$) index where *i* = *t*, (*t+1*,*t+18*), (*t+19*, *t+20*), and (*t*,*t+20*). $PhotoPes_t$ is calculated as the proportion of the most influential photos from *Getty Images* predicted to be negative on time *t*, L_s denotes an s-lag operator (s is set to 5), and X_t contains a set of exogenous variables including a constant term, month indicators, and a recession dummy. The most influential photos are the photos in the top 5% of the popularity rank distribution in each month. *PhotoPes* is winsorized at 1% and standardized to have zero mean and unit variance. The sample period is between January 1926 and June 2018. Newey and West (1987) standard errors are applied to compute t-stat. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Top 5% Most Popular Photos Dataset							
	$VWRETD_t$		$VWRETD_{\{t+1,t+18\}}$		$VWRETD_{\{t+19,t+20\}}$		$VWRETD_{\{t,t+20\}}$	
	(1)		(2)		(3)		(4)	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat
PhotoPes$_t$	-0.027**	-2.522	-0.055	-0.956	0.027*	1.716	-0.029	-0.472
Adj R-sq	0.008		0.014		0.003		0.015	
N	10,577		10,577		10,558		10,577	
	$EWRETD_t$		$EWRETD_{\{t+1,t+18\}}$		$EWRETD_{\{t+19,t+20\}}$		$EWRETD_{\{t,t+20\}}$	
	(5)		(6)		(7)		(8)	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat
PhotoPes$_t$	-0.020**	-2.084	0.038	0.535	0.041**	2.364	0.079	1.037
Adj R-sq	0.051		0.041		0.009		0.036	
N	10,577		10,577		10,558		10,577	

Table 7: Performance of Trading Strategies based on *PhotoPes*

Panel A reports mean excess returns for various trading strategies involving pessimism measures in news. We report mean, standard deviation, and Sharpe ratio (SR) of excess daily returns (in percentages) for our trading strategies involving *PhotoPes*, *TextPes*, or a combination of the two calculated from the *WSJ* sample plus the SPY index. See Section 2.1.6 for details about the strategies. Panel B reports estimates from time series regressions of daily excess returns from the *PhotoPes* and combined strategies on Fama-French (1993) three factors (*Mkt_Rf*, *SMB*, and *HML*), Carhart (1997) momentum factor (*MOM*) and Da, Liu, Schaumburg (2014) short-run reversal factor (*ST_Rev*). Newey and West (1987) standard errors are applied to compute t-stat. The sample period is between September 2008 and September 2020. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: *PhotoPes* and *TextPes*					
Strategy	N	Mean	t-stat	Std dev	SR
PhotoPes	2,386	0.053***	2.820	1.207	0.044
TextPes	2,234	0.048**	2.558	1.222	0.040
Combined	1,723	0.060***	3.579	1.116	0.053
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	0.020*	1.764	0.005	0.506
Mkt_Rf$_t$	66.1***	19.898	79.4***	27.744
SMB$_t$	3.668	0.832	0.976	0.283
HML$_t$	-13.4***	-3.919	-9.669***	-3.317
UMD$_t$	-4.311*	-1.773	-3.134	-1.494
ST_Rev$_t$	6.724**	2.544	4.699**	2.212
Adj R-sq	0.703		0.819	
N	3,034		3,034	

Table 8: Validation Test using *Getty Images* (*PhotoPes* and Volatility)

This table reports results of β_1 and β_2 from the following time-series regression:

$$R_t = \beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2) + \beta_5 X_t + \varepsilon_t,$$

where R_t is log equal-weighted (Panel A) or value-weighted (Panel B) daily return on quintile portfolios sorted on monthly volatility ($Sigma_t$) or the difference between highest and lowest volatility portfolio return (*H-L*), $PhotoPes_t$ is calculated as the proportion of the 10 most popular photos from *Getty Images* predicted to be negative on time *t*, L_s denotes an s-lag operator (s is set to 5), 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. Stocks in the CRSP universe are sorted into quintiles based on daily return volatility over the past month. *PhotoPes* is winsorized at 1% and standardized to have zero mean and unit variance. The sample period is between January 1926 and June 2018. Newey and West (1987) standard errors are applied to compute t-stat. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Panel A: EW $Sigma_t$											
	(1)		(2)		(3)		(4)		(5)		(6)	
	High		4		3		2		Low		H-L	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.032***	-2.848	-0.024**	-2.239	-0.019**	-1.960	-0.014*	-1.801	-0.009*	-1.646	-0.023***	-3.035
$PhotoPes_{t-1}$	0.024**	2.117	0.019*	1.785	0.016	1.615	0.012	1.565	0.009*	1.823	0.016**	2.024
$PhotoPes_{t-2}$	-0.005	-0.410	-0.002	-0.221	-0.001	-0.129	-0.001	-0.157	-0.001	-0.192	-0.003	-0.343
$PhotoPes_{t-3}$	0.015	1.333	0.015	1.471	0.018**	1.970	0.015**	2.025	0.010**	2.126	0.006	0.741
$PhotoPes_{t-4}$	-0.013	-1.186	-0.017	-1.513	-0.015	-1.519	-0.012	-1.426	-0.009*	-1.831	-0.003	-0.366
$PhotoPes_{t-5}$	0.005	0.407	-0.003	-0.297	-0.005	-0.504	-0.005	-0.633	-0.003	-0.724	0.010	1.130
Sum t to t-5	-0.006		-0.012		-0.006		-0.005		-0.003		0.003	
Sum t-1 to t-3	0.034		0.032*		0.033**		0.026**		0.018**		0.019	
	$\chi^2(1)$	*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 to t-5=0]	0.071	0.790	0.303	0.582	0.073	0.788	0.067	0.795	0.093	0.760	0.037	0.848
$\chi^2(1)$[Sum t-1 to t-3=0]	3.63*	0.057	3.448*	0.063	4.567**	0.033	4.516**	0.034	5.12**	0.024	2.315	0.128
Adj R-sq	0.086		0.045		0.034		0.032		0.055		0.101	
N	16,429		16,429		16,429		16,429		16,429		16,429	

	Panel B: VW $Sigma_t$											
	(1)		(2)		(3)		(4)		(5)		(6)	
	High		4		3		2		Low		H-L	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.044**	-2.564	-0.031**	-2.229	-0.026**	-2.342	-0.019**	-2.062	-0.012*	-1.903	-0.032**	-2.306
$PhotoPes_{t-1}$	0.024	1.474	0.022	1.574	0.016	1.366	0.012	1.245	0.009	1.337	0.016	1.281
$PhotoPes_{t-2}$	-0.024	-1.465	-0.012	-0.833	-0.004	-0.354	0.000	0.030	0.001	0.106	-0.023*	-1.800
$PhotoPes_{t-3}$	0.017	1.067	0.018	1.368	0.020*	1.793	0.018**	1.982	0.013**	2.080	0.005	0.434
$PhotoPes_{t-4}$	-0.002	-0.141	-0.017	-1.263	-0.017	-1.470	-0.016*	-1.663	-0.013**	-1.988	0.013	1.007
$PhotoPes_{t-5}$	-0.008	-0.496	-0.012	-0.862	-0.013	-1.194	-0.010	-1.089	-0.008	-1.304	0.000	0.011
Sum t to t-5	-0.037		-0.032		-0.024		-0.015		-0.010		-0.021	
Sum t-1 to t-3	0.017		0.028		0.032*		0.030**		0.023**		-0.002	

	$\chi^2(1)$	*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 to t-5=0]	1.363	0.243	1.241	0.265	1.114	0.291	0.609	0.435	0.710	0.400	0.712	0.399
$\chi^2(1)$[Sum t-1 to t-3=0]	0.393	0.531	1.654	0.198	2.952*	0.086	3.917**	0.048	4.791**	0.029	0.007	0.931
Adj R-sq	0.024		0.014		0.011		0.012		0.018		0.025	
N	16,429		16,429		16,429		16,429		16,429		16,429	

Table 9: Validation Test using *WSJ* (*PhotoPes* and Volatility)

This table reports results of β_1 and β_2 from the following time-series regression:

$$R_t = \beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2) + \beta_5 X_t + \varepsilon_t,$$

where R_t is log equal-weighted (Panel A) or value-weighted (Panel B) daily return on quintile portfolios sorted on monthly volatility ($Sigma_t$) or the difference between highest and lowest volatility portfolio return (*H-L*), $PhotoPes_t$ is calculated as the proportion of photos from the *WSJ* predicted to be negative on time *t*, L_s denotes an s-lag operator (s is set to 5), 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 1% and standardized to have zero mean and unit variance. We use news photos that belong in articles from the following sections: "Business", "Economy", "Markets", "Politics", and "Opinion". The sample period is between September 2008 and September 2020. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Newey and West (1987) standard errors are applied to compute t-stat for all specifications.

	Panel A: EW $Sigma_t$											
	(1)		(2)		(3)		(4)		(5)		(6)	
	High		4		3		2		Low		H-L	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.058**	-1.980	-0.052*	-1.873	-0.040*	-1.686	-0.032*	-1.650	-0.014	-1.302	-0.047**	-2.300
$PhotoPes_{t-1}$	0.081**	2.524	0.075**	2.451	0.064**	2.416	0.054**	2.431	0.031**	2.473	0.050**	2.307
$PhotoPes_{t-2}$	-0.048	-1.622	-0.060**	-2.045	-0.053**	-2.107	-0.048**	-2.278	-0.029***	-2.581	-0.017	-0.820
$PhotoPes_{t-3}$	0.023	0.802	0.030	1.076	0.027	1.134	0.026	1.283	0.015	1.305	0.007	0.344
$PhotoPes_{t-4}$	0.027	0.864	0.032	1.072	0.033	1.249	0.031	1.376	0.016	1.179	0.012	0.596
$PhotoPes_{t-5}$	0.021	0.701	0.025	0.875	0.017	0.699	0.020	0.994	0.013	1.093	0.009	0.460
Sum t to t-5	0.046		0.050		0.048		0.051		0.032		0.014	
Sum t-3 to t-5	0.071		0.087**		0.077**		0.077**		0.044**		0.028	
	$\chi^2(1)$	*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 to t-5=0]	0.669	0.413	0.916	0.339	1.132	0.287	1.801	0.180	2.524	0.112	0.141	0.707
$\chi^2(1)$[Sum t-3 to t-5=0]	2.412	0.120	3.877**	0.049	4.079**	0.043	5.764**	0.016	6.064**	0.014	0.847	0.357
Adj R-sq	0.039		0.020		0.019		0.019		0.024		0.054	
N	2,854		2,854		2,854		2,854		2,854		2,854	

	Panel B: VW $Sigma_t$											
	(1)		(2)		(3)		(4)		(5)		(6)	
	High		4		3		2		Low		H-L	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.102**	-2.263	-0.066*	-1.903	-0.045	-1.634	-0.031	-1.444	-0.015	-1.038	-0.083**	-2.345
$PhotoPes_{t-1}$	0.098**	2.253	0.088**	2.366	0.080**	2.510	0.057**	2.242	0.031*	1.855	0.063**	2.051
$PhotoPes_{t-2}$	-0.064	-1.481	-0.074**	-2.075	-0.069**	-2.421	-0.053**	-2.375	-0.032**	-2.063	-0.023	-0.690
$PhotoPes_{t-3}$	0.038	0.891	0.037	1.109	0.034	1.216	0.030	1.379	0.025*	1.656	0.008	0.260
$PhotoPes_{t-4}$	0.046	0.986	0.048	1.216	0.038	1.245	0.029	1.207	0.020	1.223	0.026	0.775
$PhotoPes_{t-5}$	0.044	0.955	0.034	0.954	0.030	1.039	0.016	0.704	0.016	1.060	0.032	0.899
Sum t to t-5	0.060		0.067		0.068		0.048		0.045		0.023	
Sum t-3 to t-5	0.128*		0.119**		0.102**		0.075**		0.061**		0.066	

	$\chi^2(1)$	*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 to t-5=0]	0.571	0.450	0.993	0.319	1.594	0.207	1.292	0.256	2.514	0.113	0.157	0.692
$\chi^2(1)$[Sum t-3 to t-5=0]	3.514*	0.061	4.541**	0.033	5.081**	0.024	4.560**	0.033	6.362**	0.012	1.662	0.197
Adj R-sq	0.033		0.029		0.025		0.033		0.033		0.030	
N	2,854		2,854		2,854		2,854		2,854		2,854	

Table 10: Validation Test: *PhotoPes* and Aggregate Uncertainty

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

$$R_t = (N_t)[\beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2)] + (1 - N_t)[\gamma_1 PhotoPes_t + \gamma_2 L_s(PhotoPes_t) + \gamma_3 L_s(R_t) + \gamma_4 L_s(R_t^2)] + \beta_5 X_t + \varepsilon_t,$$

where N_t is an indicator variable for whether period *t* is during a high volatility month as measured by NVIX and zero otherwise, R_t is log daily return on CRSP value-weight ($VWRETD_t$) index or CRSP equal-weight ($EWRETD_t$) index, $PhotoPes_t$ is calculated as the proportion of the 10 most popular photos from *Getty Images* predicted to be negative on time *t*, L_s denotes an s-lag operator (s is set to 5), 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. NVIX is a news implied volatility index provided by Manela and Moreira (2017). High (low) NVIX is during the period when NVIX is above (below) the median. *PhotoPes* is winsorized at 1% and standardized to have zero mean and unit variance. The sample period is between January 1926 and June 2016. Newey and West (1987) standard errors are applied to compute t-stat. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	$VWRETD_t$					$EWRETD_t$			
	(1)					(2)			
	High NVIX		Low NVIX			High NVIX		Low NVIX	
Variables	β	t-stat	γ	t-stat		β	t-stat	γ	t-stat
$PhotoPes_t$	-0.035**	-2.327	0.001	0.087		-0.033**	-2.116	-0.003	-0.469
$PhotoPes_{t-1}$	0.020	1.256	0.000	-0.059		0.028*	1.729	0.007	1.071
$PhotoPes_{t-2}$	0.006	0.402	-0.009	-1.017		0.006	0.402	-0.009	-1.248
$PhotoPes_{t-3}$	0.027*	1.775	0.013	1.545		0.023	1.502	0.011	1.505
$PhotoPes_{t-4}$	-0.019	-1.157	-0.004	-0.536		-0.024	-1.428	-0.001	-0.090
$PhotoPes_{t-5}$	-0.015	-1.007	-0.007	-0.854		0.001	0.090	-0.007	-0.890
Sum t to t-5	0.060		0.067		0.068	0.048		0.045	
Sum t-1 to t-3	0.128*		0.119**		0.102**	0.075**		0.061**	
	$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value		$\chi^2(1)$	*p*-value	$\chi^2(1)$	*p*-value
$\chi^2(1)$[Sum t to t-5 =0]	0.282	0.595	0.141	0.707		0.004	0.951	0.017	0.896
$\chi^2(1)$[Sum t-1 to t-3 =0]	4.847**	0.028	0.064	0.801		5.389**	0.02	0.504	0.478
Adj R-sq	0.020					0.059			
N	15,923					15,923			

Table 11: Non-linearity of *PhotoPes*, Market Returns, and Trading Volume for *Getty Images*

Panel A of this table reports β_1 to β_5 from the following time-series regression:

$$R_t = \beta_1 PhotoPes_{rank5_t} + \beta_2 PhotoPes_{rank4_t} + \beta_3 PhotoPes_{rank3_t} + \beta_4 PhotoPes_{rank2_t} + \beta_5 PhotoPes_{rank1_t} + \beta_6 L_s(PhotoPes_t) + \beta_7 L_s(R_t) + \beta_8 L_s(R_t^2) + \beta_9 X_t + \varepsilon_t,$$

where R_t denotes daily log-returns on VWRETD or EWRETD, $PhotoPes_{rank5_t}$ is the highest quintile $PhotoPes_t$ (days when most photos from *Getty Images* are predicted to be negative) and $PhotoPes_{rank1_t}$ is the lowest quintile $PhotoPes_t$ (days with fewest photos from *Getty Images* are predicted to be negative), and X_t contains a set of exogenous variables including a constant term, month indicators, and a recession dummy.

Panel B of this table reports β_1 to β_4 from the following time-series regression:

$$\bar{V}_t = \beta_1 PhotoPes_t^{high} + \beta_2 PhotoPes_t^{low} + \beta_3 L_s\left(PhotoPes_t^{high}\right) + \beta_4 L_s(PhotoPes_t^{low}) + \beta_5 L_s(R_t) + \beta_6 L_s(R_t^2) + \varepsilon_t,$$

where $\bar{V}_t$ is standardized (zero mean, unit variance) residual from regressing log of aggregate NYSE trading volume on its own lags (s=5) and month and day-of-the-week dummies, R_t is log daily return on CRSP value-weight index, $PhotoPes^{high}$ is *PhotoPes* above 0.5, $PhotoPes^{low}$ is *PhotoPes* below 0.5. In both panels, L_s denotes an s-lag operator (s is set to 5). *PhotoPes* is winsorized at 1% and standardized to have zero mean and unit variance. The sample period is between January 1926 and June 2018. Newey and West (1987) standard errors are applied to compute t-stat. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Non-linearity in Relation between *PhotoPes* and Market Returns

	$VWRETD_t$		$EWRETD_t$	
	(1)		(2)	
Variables	β	t-stat	β	t-stat
$PhotoPes_{rank5}$	0.005	0.397	-0.002	-0.130
$PhotoPes_{rank4}$	0.106	1.311	0.059	0.719
$PhotoPes_{rank3}$	-0.143**	-2.134	-0.108*	-1.673
$PhotoPes_{rank2}$	-0.110*	-1.684	-0.086	-1.340
$PhotoPes_{rank1}$	-0.085***	-2.947	-0.067**	-2.396
Adj R-sq	0.057		0.051	
N	16,430		16,430	

Panel B: *PhotoPes* and Abnormal NYSE Trading Volume

	NYSE Volume	
Variables	β	t-stat
$PhotoPes_t^{high}$	0.015	0.988
$PhotoPes_t^{low}$	0.006	0.187
$PhotoPes_{t-1}^{high}$	0.033**	2.027
$PhotoPes_{t-1}^{low}$	0.033**	2.027
$PhotoPes_{t-2}^{high}$	0.012	0.828
$PhotoPes_{t-2}^{low}$	0.004	0.125
$PhotoPes_{t-3}^{high}$	0.007	0.464
$PhotoPes_{t-3}^{low}$	0.005	0.188
$PhotoPes_{t-4}^{high}$	0.040**	2.213
$PhotoPes_{t-4}^{low}$	-0.010	-0.314
$PhotoPes_{t-5}^{high}$	0.023*	1.672
$PhotoPes_{t-5}^{low}$	-0.026	-0.887
Adj R-sq	0.004	
N	14,488	

Table 12: Non-linearity of *PhotoPes*, Market Returns, and Trading Volume for *WSJ*

Panel A of this table reports β_1 to β_5 from the following time-series regression:

$$R_t = \beta_1 PhotoPes_{rank5_t} + \beta_2 PhotoPes_{rank4_t} + \beta_3 PhotoPes_{rank3_t} + \beta_4 PhotoPes_{rank2_t} + \beta_5 PhotoPes_{rank1_t} + \beta_6 L_s(PhotoPes_t) + \beta_7 L_s(R_t) + \beta_8 L_s(R_t^2) + \beta_9 X_t + \varepsilon_t,$$

where R_t denotes daily log-returns on VWRETD or EWRETD, $PhotoPes_{rank5_t}$ is the highest quintile $PhotoPes_t$ (days when most photos from the *WSJ* are predicted to be negative) and $PhotoPes_{rank1_t}$ is the lowest quintile $PhotoPes_t$ (days with fewest photos from the *WSJ* are predicted to be negative), and X_t contains a set of exogenous variables including a constant term, month indicators, and a recession dummy.

Panel B of this table reports β_1 to β_4 from the following time-series regression:

$$\bar{V}_t = \beta_1 PhotoPes_t^{high} + \beta_2 PhotoPes_t^{low} + \beta_3 L_s\left(PhotoPes_t^{high}\right) + \beta_4 L_s(PhotoPes_t^{low}) + \beta_5 L_s(R_t) + \beta_6 L_s(R_t^2) + \varepsilon_t,$$

where $\bar{V}_t$ is standardized (zero mean, unit variance) residual from regressing log of aggregate NYSE trading volume on its own lags (s=5) and month and day-of-the-week dummies, R_t is log daily return on CRSP value-weight index, $PhotoPes^{high}$ is *PhotoPes* above 0.5, $PhotoPes^{low}$ is *PhotoPes* below 0.5. In both panels, L_s denotes an s-lag operator (s is set to 5). *PhotoPes* is winsorized at 1% and standardized to have zero mean and unit variance. We use news photos that belong in articles from the following sections: "Business", "Economy", "Markets", "Politics", and "Opinion". The sample period is between September 2008 and December 2019 for Panel A (June 2018 for Panel B). Newey and West (1987) standard errors are applied to compute t-stat. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Non-linearity in Relation between *PhotoPes* and Market Returns				
	VWRETD$_t$		EWRETD$_t$	
	(1)		(2)	
Variables	β	t-stat	β	t-stat
$PhotoPes_{rank5}$	-0.032	-0.784	-0.022	-0.560
$PhotoPes_{rank4}$	0.040	0.288	0.080	0.617
$PhotoPes_{rank3}$	-0.260	-0.858	-0.283	-1.015
$PhotoPes_{rank2}$	0.090	0.675	0.093	0.742
$PhotoPes_{rank1}$	-0.080*	-1.723	-0.094**	-2.176
Adj R-sq	0.031		0.023	
N	2,854		2,854	

Panel B: *PhotoPes* and Abnormal NYSE Trading Volume		
	NYSE Volume	
Variables	β	t-stat
$PhotoPes_t^{high}$	0.001	0.033
$PhotoPes_t^{low}$	0.019	0.428
$PhotoPes_{t-1}^{high}$	0.080**	2.227
$PhotoPes_{t-1}^{low}$	0.080**	2.227
$PhotoPes_{t-2}^{high}$	-0.027	-0.866
$PhotoPes_{t-2}^{low}$	0.029	0.784
$PhotoPes_{t-3}^{high}$	0.044	1.229
$PhotoPes_{t-3}^{low}$	-0.019	-0.402
$PhotoPes_{t-4}^{high}$	-0.022	-0.670
$PhotoPes_{t-4}^{low}$	-0.014	-0.323
$PhotoPes_{t-5}^{high}$	0.007	0.199
$PhotoPes_{t-5}^{low}$	0.014	0.320
Adj R-sq	0.013	
N	2,291	

Table 13: Robustness Test (*Getty Images*)

This table reports results of β_1 and β_2 from the following time-series regression:

$$R_t = \beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2) + \beta_5 X_t + \varepsilon_t,$$

where R_t is log daily return on CRSP value-weighted or equal-weight index, $PhotoPes_t$ is calculated as the proportion of the 10 most popular photos from *Getty Images* predicted to be negative on time *t* (modified in several ways as discussed below), L_s denotes an s-lag operator (s is set to 5), 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. Panel C presents the results when *PhotoPes* is constructed using predicted likelihood of negative sentiment in the photo instead of negative sentiment dummy indicators. *PhotoPes* in all but Panel B is winsorized at 1%. *PhotoPes* is standardized to have zero mean and unit variance. The sample period is between January 1926 and June 2018. Newey and West (1987) standard errors are applied to compute t-stat. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Panel A: Cutoff={0.55,0.45}				Panel B: No Winsorization				Panel C: Predicted likelihood			
	(1)		(2)		(1)		(2)		(1)		(2)	
	$VWRETD_t$		$EWRETD_t$		$VWRETD_t$		$EWRETD_t$		$VWRETD_t$		$EWRETD_t$	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.018**	-2.052	-0.023***	-2.691	-0.019**	-2.227	-0.021**	-2.304	-0.018**	-2.115	-0.019**	-2.192
$PhotoPes_{t-1}$	-0.002	-0.258	0.008	0.892	0.011	1.204	0.018**	1.992	0.012	1.353	0.020**	2.215
$PhotoPes_{t-2}$	-0.010	-1.168	-0.008	-0.864	-0.002	-0.185	-0.002	-0.243	-0.005	-0.615	-0.007	-0.830
$PhotoPes_{t-3}$	0.017*	1.907	0.014*	1.678	0.017**	2.018	0.015*	1.864	0.016*	1.913	0.015*	1.840
$PhotoPes_{t-4}$	-0.006	-0.688	-0.009	-1.035	-0.013	-1.418	-0.014	-1.554	-0.014	-1.609	-0.013	-1.438
$PhotoPes_{t-5}$	-0.002	-0.269	0.002	0.194	-0.010	-1.187	-0.002	-0.186	-0.010	-1.174	-0.002	-0.218
Sum t to t-5	-0.021		-0.016		-0.016		-0.006		-0.019		-0.006	
Sum t-3 to t-5	0.009		0.007		-0.006*		-0.001**		-0.008		0.000**	
	$\chi^2(1)$	*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 to t-5=0]	1.651	0.199	0.754	0.385	0.861	0.353	0.089	0.765	1.218	0.270	0.096	0.757
$\chi^2(1)$[Sum t-3 to t-5=0]	0.082	0.774	1.142	0.285	3.481*	0.062	5.041**	0.025	2.660	0.103	3.986**	0.046
Adj R-sq	0.014		0.052		0.016		0.051		0.016		0.051	
N	17,069		17,069		16,430		16,430		16,430		16,430	

Table 14: Robustness Test (*WSJ*)

This table reports results of β_1 and β_2 from the following time-series regression:

$$R_t = \beta_1 PhotoPes_t + \beta_2 L_s(PhotoPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2) + \beta_5 X_t + \varepsilon_t,$$

where R_t is log daily return on CRSP value-weight ($VWRETD_t$) index, CRSP equal-weight ($EWRETD_t$) index, SPDR S&P 500 ETF Trust (SPY_t), SPDR Dow Jones Industrial Average ETF (DIA_t), or iShares Russell 2000 ETF (IWM_t), $PhotoPes_t$ is calculated as the proportion of photos from the *WSJ* predicted to be negative on time *t* (modified in several ways as discussed below), L_s denotes an s-lag operator (s is set to 5), 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. Panel C presents the results when *PhotoPes* is constructed using predicted likelihood of negative sentiment in the photo instead of negative sentiment dummy indicators. *PhotoPes* in all but Panel B is winsorized at 1%. We use news photos that belong in articles from the following sections: "Business", "Economy", "Markets", "Politics", and "Opinion". *PhotoPes* is standardized to have zero mean and unit variance. The sample period is between September 2008 and September 2020. Newey and West (1987) standard errors are applied to compute t-stat. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Cutoff={0.55,0.45}										
	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		$EWRETD_t$		SPY_t		DIA_t		IWM_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.039*	-1.728	-0.037*	-1.718	-0.060**	-2.011	-0.045**	-2.092	-0.060**	-2.011
$PhotoPes_{t-1}$	0.050**	2.101	0.052**	2.351	0.047	1.464	0.030	1.268	0.047	1.464
$PhotoPes_{t-2}$	-0.048**	-2.021	-0.039*	-1.683	-0.017	-0.537	-0.025	-1.152	-0.017	-0.537
$PhotoPes_{t-3}$	0.037	1.577	0.030	1.315	0.036	1.157	0.044**	2.006	0.036	1.157
$PhotoPes_{t-4}$	0.026	1.021	0.023	0.965	0.075**	2.268	0.046*	1.802	0.075**	2.268
$PhotoPes_{t-5}$	0.017	0.716	0.012	0.549	0.021	0.695	0.022	1.016	0.021	0.695
Sum t to t-5	0.043		0.041		0.102*		0.072		0.102*	
Sum t-3 to t-5	0.080**		0.065*		0.132***		0.112***		0.132***	
	$\chi^2(1)$	*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 to t-5=0]	0.944	0.331	1.042	0.307	3.066*	0.080	2.512	0.113	3.066*	0.080
$\chi^2(1)$[Sum t-3 to t-5=0]	4.556**	0.033	3.512*	0.061	7.659***	0.006	8.975***	0.003	7.659***	0.006
Adj R-sq	0.029		0.020		0.031		0.040		0.031	
N	2,854		2,854		3,043		3,043		3,043	

Panel B: No Winsorization										
	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		$EWRETD_t$		SPY_t		DIA_t		IWM_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$PhotoPes_t$	-0.042*	-1.805	-0.040*	-1.802	-0.042*	-1.773	-0.047**	-2.118	-0.065**	-2.115
$PhotoPes_{t-1}$	0.066**	2.451	0.067***	2.644	0.053*	1.881	0.046*	1.718	0.064*	1.833
$PhotoPes_{t-2}$	-0.058**	-2.291	-0.048**	-2.035	-0.035	-1.411	-0.030	-1.301	-0.029	-0.841
$PhotoPes_{t-3}$	0.032	1.329	0.024	1.070	0.022	0.958	0.028	1.270	0.019	0.625
$PhotoPes_{t-4}$	0.029	1.104	0.027	1.060	0.052*	1.909	0.052**	1.966	0.082**	2.435

PhotoPes$_{t-5}$	0.026	1.053	0.023	0.954	0.035	1.498	0.026	1.180	0.031	0.991
Sum t to t-5	0.053		0.053		0.085*		0.075		0.102*	
Sum t-3 to t-5	0.087**		0.074**		0.109***		0.106***		0.132***	
	$\chi^2(1)$	*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 to t-5=0]	1.402	0.236	1.464	0.226	3.004*	0.083	2.619	0.106	2.822*	0.093
$\chi^2(1)$[Sum t-3 to t-5=0]	5.163**	0.023	4.198**	0.040	7.611***	0.006	8.025***	0.005	7.31***	0.007
Adj R-sq	0.030		0.021		0.030		0.040		0.032	
N	2,854		2,854		3,043		3,043		3,043	

Panel C: Predicted likelihood

	(1)		(2)		(3)		(4)		(5)	
	VWRETD$_t$		EWRETD$_t$		SPY$_t$		DIA$_t$		IWM$_t$	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
PhotoPes$_t$	-0.050**	-2.222	-0.047**	-2.227	-0.053**	-2.280	-0.057**	-2.554	-0.079***	-2.735
PhotoPes$_{t-1}$	0.040	1.614	0.044*	1.908	0.027	1.085	0.024	1.036	0.038	1.169
PhotoPes$_{t-2}$	-0.038	-1.569	-0.031	-1.354	-0.021	-0.883	-0.015	-0.671	-0.011	-0.340
PhotoPes$_{t-3}$	0.016	0.661	0.015	0.637	0.016	0.647	0.027	1.146	0.013	0.428
PhotoPes$_{t-4}$	0.043*	1.703	0.039	1.630	0.062**	2.377	0.056**	2.210	0.087***	2.672
PhotoPes$_{t-5}$	0.023	1.014	0.022	0.990	0.030	1.322	0.021	0.944	0.032	1.042
Sum t to t-5	0.034		0.042		0.061		0.056		0.080	
Sum t-3 to t-5	0.082**		0.076**		0.108***		0.104***		0.132***	
	$\chi^2(1)$	*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 to t-5=0]	0.787	0.375	1.298	0.255	2.183	0.140	1.943	0.163	2.318	0.128
$\chi^2(1)$[Sum t-3 to t-5=0]	5.369**	0.020	5.110**	0.024	8.227***	0.004	8.079***	0.004	8.001***	0.005
Adj R-sq	0.028		0.020		0.030		0.040		0.032	
N	2,854		2,854		3,043		3,043		3,043	

Figure 1: Model Accuracy

This figure shows the training and validation accuracy for the first 500 training steps using Google Inception v3 model and the DeepSent training dataset with clean labels. 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 is learning steps or the number of times we pass all our training set through our 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 that are 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. We have 10% of the photos for the validation set, 80% for the training set, and 10% for the testing set.

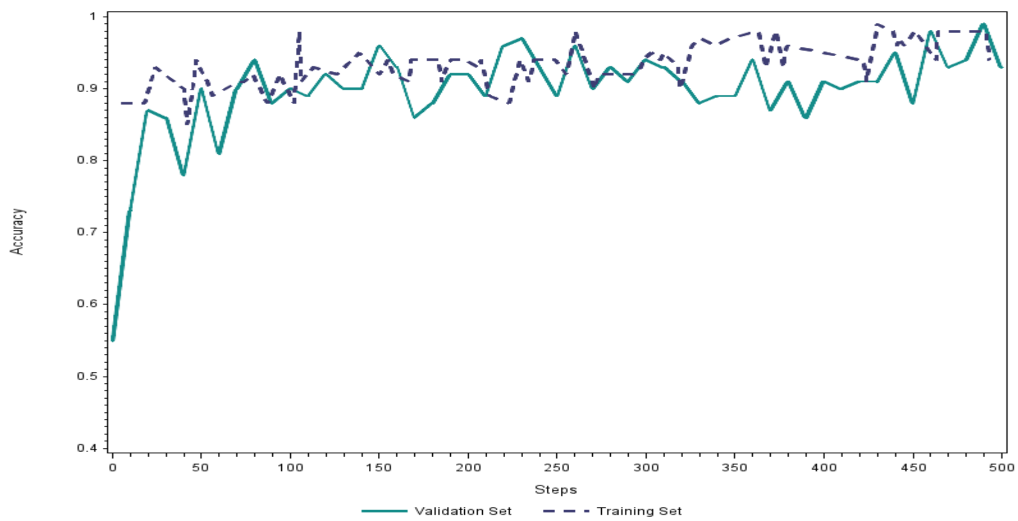
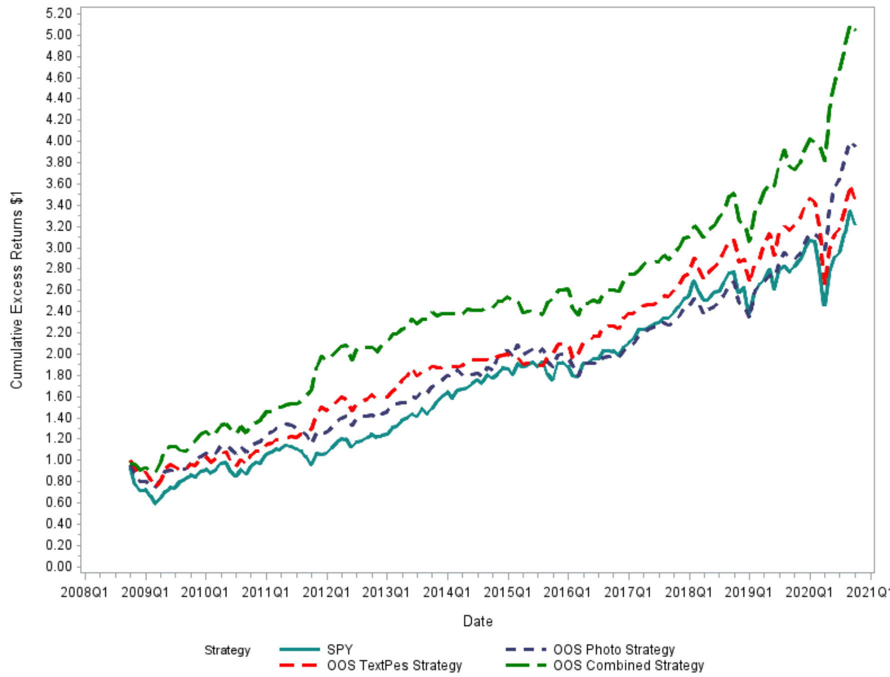


Figure 2: Trading Strategy

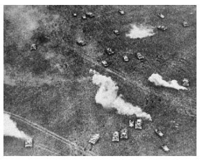
This graph depict the accumulated dollar excess return from the out-of-sample trading strategies based on pessimism from news compared to passive strategy of holding the SPY. See Section 2.1.6 for the details of trading strategies.



Appendix

Table A1: Photos with Highest Probability of Negative and Positive Sentiment

This table presents twenty photos that have highest *PhotoNeg* or the highest probability to have negative sentiment (top) and twenty photos that have lowest *PhotoNeg* or the highest probability to have positive sentiment (bottom) from *Getty* sample.[37] These photos can be accessed by using the following link (https://www.gettyimages.com/detail/news-photo/news-photo/[Photo ID]).


52594147


78950824


124131318


513288040


517664989


522516562


525582218


542345406


551552711


591735262


645906848


699744474


830711908


949069266


951036618


961107126


1053619406


1080938748


1081539576


1096123462


50613925


55949473


98263816


108120820


134367376


156250153


514188266


522613560


525547380


535494744


542622952


607809433


612556268


612557514


669382238

[37] *PhotoNeg* presents the probability *each* photo has negative sentiment.


830306150


892927502


909061342


918594174


928894700

Photo ID	*PhotoNeg*	Photo ID	*PhotoNeg*
50613925	0	52594147	1
55949473	0	78950824	1
98263816	0	124131318	1
108120820	0	513288040	1
134367376	0	517664989	1
156250153	0	522516562	1
514188266	0	525582218	1
522613560	0	542345406	1
525547380	0	551552711	1
535494744	0	591735262	1
542622952	0	645906848	1
607809433	0	699744474	1
612556268	0	830711908	1
612557514	0	949069266	1
669382238	0	951036618	1
830306150	0	961107126	1
892927502	0	1053619406	1
909061342	0	1080938748	1
918594174	0	1081539576	1
928894700	0	1096123462	1

Table A2: News Photos from *WSJ* with Highest Probability of Negative and Positive Sentiment

This table presents twenty photos that have highest *PhotoNeg* or the highest probability to have negative sentiment (top) and twenty photos that have lowest *PhotoNeg* or the highest probability to have positive sentiment (bottom) from the *WSJ* sample. *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. The news photos are taken from the articles appearing in the following sections: "Business", "Economy", "Markets", "Politics", and "Opinion".


4e201c6f3e148ae90f9bdd8c41a57e73


08c36aef1ee30d62b9029893cf40f225


8c2fd22616c8844807534f2597ec9790


9aeac583e1b513d681044d9a24db1ea8


9fd393829624389126e25bd881b2321d


579a29c006cb95ee1b698164ea9d9afd


772c5701dbd098be15ea5c493cfebd70


950a7b33ffc87217bdfdaf071ec13785


4950cd76498e0a52259019a43f25da23


5030cd537044c8ef790237eeb191a905


9654c566def2ae0777c05367b013eb48


921097eeb44bf9977cebe1b169554e13


a2d087f194d3125c0997e7ac26b360a9


a12bd74200607bdd61216957450f038d


b2eeb3c4d2ffb0c9516d229a5fa02057


b03de67c18c6a9f23faed00ccd7a35bc


bbc52eb2de9f7ec56947ae1796e95a0d


be11b1f893dda53e459e5475a3d9277b


e62cb9633cbe96890feabc1fe62e5a9b


fd2db287dd03b19b405ecf7a352dc9f3


0d5366c32b5068c0809e6d2e07ea10c7


2d82d14ac7c5a1d81ed80b1837d6e513

3a3dcb6becc0de67a26fc8ae4f1b5c11


3c16e5d737914f883206a9d046f2bb63


6c0bed71c09bee6612354501da7abc63

7bc2510f073c9bdcde05f9a9397c541e

11b4fff04ae7186c710de515b9ea8b58


17a7af07615a0443b213e01f5a01d03b


31abc1d9e89f58106c9275297aebb46d


574f2d774ceaaed6c435552925fba451


3787d96c0aa80e47f97998414abe64d8


30384050b194f263670fd89290ef18e6

a237ea60796e94f1d8ce6fc33c939bf5


aadf33324144f9494de0b3861828019f


b1ed808a11ce86dd832d1ba9cb848e7c

b26fd8e4befebf1685dc41b0e39b225e

bd5ca94a10cb9a1771b489604f61fc51

c6f6d3bcab85972ab4be94ba2a3ee338

d262fb8f242f558996e057456f6669b5

fb3b3a11e51ed42b5e35c78fa05bd09b

Photo ID	*PhotoNeg*	*TextNeg*	Photo ID	*PhotoNeg*	*TextNeg*
a237ea60796e94f1d8ce6fc33c939bf5	0	-1.5	8c2fd22616c8844807534f2597ec9790	1	0
aadf33324144f9494de0b3861828019f	0	-0.5	b03de67c18c6a9f23faed00ccd7a35bc	1	0
d262fb8f242f558996e057456f6669b5	0	-0.75	bbc52eb2de9f7ec56947ae1796e95a0d	1	-1
3787d96c0aa80e47f97998414abe64d8	0	-0.75	9654c566def2ae0777c05367b013eb48	1	-0.5
3c16e5d737914f883206a9d046f2bb63	0	-0.5	b2eeb3c4d2ffb0c9516d229a5fa02057	1	0
17a7af07615a0443b213e01f5a01d03b	0	-0.5	a2d087f194d3125c0997e7ac26b360a9	1	0
c6f6d3bcab85972ab4be94ba2a3ee338	0	-0.5	9aeac583e1b513d681044d9a24db1ea8	1	0
3a3dcb6becc0de67a26fc8ae4f1b5c11	0	-1	e62cb9633cbe96890feabc1fe62e5a9b	1	0
30384050b194f263670fd89290ef18e6	0	-1	4950cd76498e0a52259019a43f25da23	1	0
11b4fff04ae7186c710de515b9ea8b58	0	-1	9fd393829624389126e25bd881b2321d	1	-0.5
bd5ca94a10cb9a1771b489604f61fc51	0	-1.5	fd2db287dd03b19b405ecf7a352dc9f3	1	0
fb3b3a11e51ed42b5e35c78fa05bd09b	0	-1	be11b1f893dda53e459e5475a3d9277b	1	-0.3333
2d82d14ac7c5a1d81ed80b1837d6e513	0	-1	5030cd537044c8ef790237eeb191a905	1	-1
b1ed808a11ce86dd832d1ba9cb848e7c	0	-0.25	08c36aef1ee30d62b9029893cf40f225	1	0
574f2d774ceaaed6c435552925fba451	0	-0.5	950a7b33ffc87217bdfdaf071ec13785	1	-0.8333
0d5366c32b5068c0809e6d2e07ea10c7	0	-1	772c5701dbd098be15ea5c493cfebd70	1	0
b26fd8e4befebf1685dc41b0e39b225e	0	-1.5	a12bd74200607bdd61216957450f038d	1	-1
6c0bed71c09bee6612354501da7abc63	0	-1	4e201c6f3e148ae90f9bdd8c41a57e73	1	-1
31abc1d9e89f58106c9275297aebb46d	0	-1	921097eeb44bf9977cebe1b169554e13	1	0
7bc2510f073c9bdcde05f9a9397c541e	0	-0.75	579a29c006cb95ee1b698164ea9d9afd	1	-1.5

Table A3: The Most Influential Photos with Highest Probability of Negative and Positive Sentiment

This table presents twenty influential photos that have highest *PhotoNeg* or the highest probability to have negative sentiment (top) and twenty photos that have lowest *PhotoNeg* or the highest probability to have positive sentiment (bottom) from *Getty* sample. The most influential photos are photos in the top 5% of the popularity rank distribution in a given month. These photos can be accessed by using the following link (https://www.gettyimages.com/detail/news-photo/news-photo/[Photo ID]).


51896187


51909903


52494062


81258023


82921542


85327803


95892358


103938287


121136144


187854864


187967864


451195036


495852752


504956752


514133374


625020084


683750210


828105222


840141664


967002442


3012886


52945917


53127470


71183785


71355772


74898934


83868221


114989427

451635860


480200034


480887939


543414756


586214376


587482390


612557514


936467030


948016810


1037223756


1041156822


1125254610

Table A4: Descriptive Statistic of Market Returns

This table reports sample statistics for the daily returns on CRSP value-weight (VWRETD) and equal-weight (EWRETD) indices, SPDR S&P 500 ETF Trust (SPY_t), SPDR Dow Jones Industrial Average ETF (DIA_t), and iShares Russell 2000 ETF (IWM_t). Numbers reported are in percentages.

Panel A: Market Returns Descriptive Statistics for *Getty Images* Sample						
R_t	N	Mean	P50	P25	P75	Std dev
VWRETD	23,230	0.036	0.071	-0.394	0.502	1.063
EWRETD	23,230	0.071	0.123	-0.298	0.492	1.044

Panel B: Market Returns Descriptive Statistics for *WSJ* Sample						
R_t	N	Mean	P50	P25	P75	Std dev
VWRETD	2,854	0.0439	0.0755	-0.3806	0.5656	1.2309
EWRETD	2,854	0.0517	0.1077	-0.4114	0.5636	1.1415
SPY	3,043	0.0480	0.0700	-0.3700	0.5800	1.3266
DIA	3,043	0.0464	0.0700	-0.3700	0.5500	1.2955
IWM	3,043	0.0422	0.1100	-0.6300	0.8000	1.6273

Table A5: Out-of-Sample Analysis and Flexible Model

Panel A reports the R^2_{OOS} (out-of-sample R^2) in percent and associated p-values using the MSPE-Adjusted statistic in Clark and West (2007) for recursively estimated predictive regression of VWRETD or EWRETD on one-period lag of *PhotoPes* using sample of photos from *Getty Images*. We include different initial estimation windows: 1926 to 1969, 1926 to 1979, 1926 to 1989, and 1926 to 1999. Panel B repeats the analysis for Panel A, except *PhotoPes* is constructed using photos from the *WSJ* sample. Panel C reports the R^2_{OOS} (out-of-sample R^2) in percent and associated p-values using the MSPE-Adjusted statistic in Clark and West (2007) for estimated neural network predictive regression of VWRETD or EWRETD on five lags of *PhotoPes* (constructed using photos from the *Getty Images* sample), five lags of market returns, five lags of squared market returns, day of the week dummies (except for Monday), and a recession dummy with 1 hidden layer and 32 neurons. The loss function we use is standard least squares. We use ReLU activation function at all nodes. We train the network using stochastic gradient descent in 100 epochs. We refit the model every 100 days. Each time we refit the model, we increase the training sample to include all available information at date *t-1*. *Start* refers to the start year for the model evaluation. The sample period is between January 1926 and June 2018. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Linear (Predictive) using *Getty Images*			
Return	Start	R^2_{OOS} (%)	MSPE-Adj
VWRETD	1970	1.995*	0.051
EWRETD	1970	1.979***	0.000
VWRETD	1980	2.309*	0.051
EWRETD	1980	2.230***	0.000
VWRETD	1990	0.958**	0.021
EWRETD	1990	1.661***	0.000
VWRETD	2000	0.858*	0.064
EWRETD	2000	1.128***	0.002

Panel B: Linear (Predictive) using *WSJ*			
Return	Start	R^2_{OOS} (%)	MSPE-Adj
VWRETD	2010	0.879**	0.016
EWRETD	2010	1.292***	0.001

Panel C: Neural Network (Predictive) using *Getty Images*			
Return	Start	R^2_{OOS} (%)	MSPE-Adj
VWRETD	1970	1.366**	0.031
EWRETD	1970	0.481*	0.067

Table A6: *TextPes* from *WSJ*

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

$$R_t = \beta_1 TextPes_t + \beta_2 L_s(TextPes_t) + \beta_3 L_s(R_t) + \beta_4 L_s(R_t^2) + \beta_5 X_t + \varepsilon_t,$$

where $TextPes_t$ is calculated as the average pessimism score for headline and summary of each article generated from the sentiment tool in Stanford's CoreNLP software on time *t*, L_s denotes an s-lag operator (s is set to 5), 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 1% and standardized to have zero mean and unit variance. We use news articles that belong in articles from the following sections: "Business", "Economy", "Markets", "Politics", and "Opinion". R_t is log daily return on CRSP value-weight ($VWRETD_t$) index, CRSP equal-weight ($EWRETD_t$) index, SPDR S&P 500 ETF Trust (SPY_t), SPDR Dow Jones Industrial Average ETF (DIA_t), or iShares Russell 2000 ETF (IWM_t). The sample period is between January 1998 and September 2020. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Newey and West (1987) standard errors are applied to compute t-stat for all specifications.

	(1)		(2)		(3)		(4)		(5)	
	$VWRETD_t$		$EWRETD_t$		SPY_t		DIA_t		IWM_t	
Variables	β	t-stat	β	t-stat	β	t-stat	β	t-stat	β	t-stat
$TextPes_t$	-0.076**	-2.088	-0.068**	-2.122	-0.098**	-2.470	-0.097**	-2.492	-0.079*	-1.647
$TextPes_{t-1}$	-0.032	-0.929	0.001	0.018	-0.044	-1.229	-0.041	-1.153	-0.065	-1.452
$TextPes_{t-2}$	-0.021	-0.595	-0.027	-0.839	-0.013	-0.341	-0.016	-0.434	-0.038	-0.781
$TextPes_{t-3}$	0.012	0.362	0.005	0.151	0.013	0.381	0.010	0.295	0.000	0.003
$TextPes_{t-4}$	0.019	0.527	0.008	0.245	0.067	1.592	0.093**	2.323	0.076	1.584
$TextPes_{t-5}$	0.074*	1.959	0.086**	2.546	0.051	1.311	0.036	0.958	0.094*	1.898
Sum t to t-5	-0.024		0.005		-0.024		-0.015		-0.012	
Sum t-3 to t-5	0.105**		0.099**		0.131**		0.139***		0.170***	
	$\chi^2(1)$	*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 to t-5=0]	1.792	0.181	0.122	0.727	1.673	0.196	0.699	0.403	0.248	0.619
$\chi^2(1)$[Sum t-3 to t-5=0]	4.970**	0.026	5.69**	0.017	6.594**	0.010	7.664***	0.006	7.513***	0.006
Adj R-sq	0.014		0.020		0.018		0.021		0.019	
N	5,518		5,518		5,707		5,707		5,113	