Measurement Error in Dependent Variables: An Illustration within the Accounting Literature

Ed deHaan
edehaan@uw.edu
University of Washington

Alastair Lawrence
alastairlawrence@london.edu
London Business School

Robin Litjens
h.j.r.litjens@uvt.nl
Tilburg University

February 2019

ABSTRACT: This paper illustrates how measurement error ("ME") in dependent variables can bias inferences. Specifically, we show that ME in dependent variables does not only reduce power but, under common conditions in accounting studies, can lead to statistical biases and erroneous inferences even when the ME is uncorrelated with the explanatory variables. These confounds exists because ME in accounting measures is typically nonadditive, which violates the simple assumptions discussed in most econometrics texts. Using simulation analyses of commonly-used accounting variables as well as a case-study analysis of Google ticker searches as a measure of investor attention, our analyses find that nonadditive ME drives nontrivial biases and premature conclusions in accounting research. Our findings indicate that researchers should carefully consider the extent and form of ME in dependent variables, and how such ME may bias inferences.

KEYWORDS: Dependent variables; measurement error; bias; Google search; SVI.

JEL CLASSIFICATION: C13, C15, M41.

ACKNOWLEDGMENTS: We are especially grateful for comments and guidance from Mike Drake and Jake Thornock. Their willingness to assist us in improving upon their Google SVI research is a testament to their commitment to moving the literature forward. Thanks also to David Burgstahler, Peter Easton, Zachary Kaplan, Miguel Minutti-Meza, Sönke Sievers, Dan Taylor, David Veenman, and workshop participants at the Tilburg Winter Camp and Paderborn University.

1. Introduction

Researchers often hold the view that measurement error (ME) in dependent variables reduces the efficiency of statistical analyses and generally biases against finding statistical relations.[1,2] For example, it is often stated that while ME in dependent variables affects the efficiency of OLS coefficient estimates, the coefficients remain unbiased and consistent. Accordingly, little thought is given to ME in dependent variables if statistical relations are documented. While under basic assumptions this foregoing statement is true, it does not hold for *nonadditive* ME commonly observed in accounting and finance studies.

Additive ME takes the form of $y = y' + v$ (or equivalently, $y' = y - v$), where y' is the relevant theoretical construct and y is the observable proxy measured with error v. Nonadditive ME takes the form of $y \times v$, y/v, or any number of other functions. This study shows that nonadditive ME in dependent variables, under common conditions, not only decreases the efficiency of statistical analyses but also leads to biased coefficient estimates. Biased coefficients can under- or over-state true statistical relations, confound interpretations of both statistical and economic significance, and are especially problematic in cross-sectional analyses.

Nonadditive ME in dependent variables is likely common in accounting studies for several reasons. First, it is well-accepted that accounting values are typically measured with error relative to the underlying economic constructs, due to both GAAP and subjective estimations (e.g., Holthausen 1990; Barth and Landsman 1995; Fields et al. 2001; Schipper and Vincent

[1] Measurement error (ME) is simply defined as "the difference between an observed variable and the variable that belongs in a multiple regression equation" (Wooldridge 2012, p852). ME frequently has a non-zero mean and can be systematically related either to the true variable or to other variables within a regression (Cameron & Trivedi 2005). ME should not be confused with "random noise."

[2] The attenuation bias from ME in independent variables is fairly well known and has been the subject of prior accounting papers (Brown et al. 1987; Easton and Zmijewski 1989). Our study focuses on ME in dependent variables, the effects of which are not well known and are different from ME in independent variables.

2003). Accounting ME is typically thought to be proportional (not additive) to the true economic values. For example, conservatism biases in assets and earnings are modelled as percentages of the unbiased amounts; e.g., firms' book assets are *v* percent lower than economic assets, not *v* dollars lower (e.g., McNichols et al. 2014; Basu 1997; LaFond and Watts 2008). Second, dependent accounting variables are frequently scaled, so ME in the numerator or denominator becomes inherently nonadditive. For example, any (additive or nonadditive) ME in book values and assets becomes nonadditive in variables such as return on equity, return on assets, and Tobin's Q (e.g., in ROA = earnings / (assets – *v*), the effect of *v* is nonadditive).[3] Moreover, incomplete data or inadequate proxies (e.g., Easton and Zmijewski 1989; Watts and Zimmerman 1990; Subramanyam 1996; McNichols 2000; Givoly et al. 2007; Lawrence et al. 2013) often end up being a function of scaled variables and in turn, further result in nonadditive ME. Third, variables that are calculated as residuals from first-stage regressions have nonadditive ME from crude coefficient estimates that are multiplied by observational values (e.g., Gerakos 2012).[4] Thus, nonadditive ME is prevalent in a wide array of other variables such as abnormal investments, accruals, report readability, disclosure quality, effective tax rates, and financial reporting quality. Accordingly, understanding the effects of nonadditive ME in dependent variables is likely important for most accounting studies.

This paper first reviews the specific conditions where ME in the dependent variable, while decreasing the efficiency of the analyses, does not result in biased coefficients. In short, noisy dependent variables produce consistent and unbiased OLS coefficient estimates only when

[3] One may argue that ROA has little ME if a researcher is literally interested in modeling return on assets as produced by GAAP. However, studies of economically-motivated hypotheses are typically interested in modeling a firm's economic performance, for which ROA is employed as an observable proxy. ROA is certainly measured with error relative to economic performance.

[4] Chen et al. (2017), who examine incorrect inferences when using residuals as dependent variables, analyze attenuation bias in two-stage regressions.

the ME is additive (i.e., $y' = y - v$), uncorrelated with the explanatory variables X, and uncorrelated with the residual μ. If so, an OLS regression of $[(y\text{-}v) = \beta X + \mu]$ reorganizes to $[y = \beta X + (\mu+v)]$, which reduces efficiency without affecting β. We then show how nonadditive ME such as $y' = y\times v$ can lead to downward or upward biased coefficient estimates even where v is uncorrelated with X and μ. Intuitively, a regression of $[(y\times v) = \beta X + \mu]$ reorganizes to $[y = \beta(X/v) + (\mu/v)]$, so both the coefficient estimates and standard errors are affected. The sign and magnitude of the bias in β depends on the values of v and X. For example, for an average 10% measurement error of $v = 1.1$, then $(X/v) < X$ and the coefficient estimate β is biased away from zero and towards finding larger results. We next show that nonadditive ME is especially problematic in cross-sectional analyses when ME is correlated with the partitioning variable, which is again likely commonplace in accounting and finance studies. For example, ME in accounting book values likely varies cross-sectionally with partitioning variables such as size, industry, time, and stage of growth. Thus, tests interacting X with partitioning variable Z can produce upwards or downwards biased interaction coefficients. Section 2.2 further discusses these conditions in an intuitive, accessible manner.

Providing multiple case studies of the effects of nonadditive ME in published studies is challenging because ME is typically unobservable. Instead, we use both simulation analyses and one case study analysis to demonstrate the effects of nonadditive ME. First, in Section 3 we use simple simulations to illustrate the potential biases caused by nonadditive ME in common dependent variables used in accounting research. We begin using univariate regressions of *ROA* to show: i) ME in accounting assets can positively bias regression coefficients even when ME is uncorrelated with X; and ii) attempting to remove ME with log transformations can exacerbate incorrect inferences.[5] Second, we use simulations to show that correlation between ME and a partitioning variable can generate type 1 errors in cross-sectional tests and, more importantly, to quantify how big or small ME must be to generate type 1 errors in a typical accounting sample. We find that regressions with *ROA, Tobins Q,* and *Investments* as dependent variables all produce unacceptably high type 1 errors in cross-sectional tests if groups of firms have differences in ME of just 0.5% of assets. That such a small difference in ME can cause type 1 errors is perhaps surprising given that 0.5% of assets is a common threshold for audit immateriality (Eilifsen and Messier 2015).

In Section 4, we provide an extensive case study analysis clearly demonstrating the effects of nonadditive ME in one commonly used variable: Google ticker search volume index (SVI), which is used as a measure of retail investor attention (Da et al. 2011; Drake et al. 2012) in more than 60 published papers (see Appendix B). Researchers frequently use Google SVI as a dependent variable in examining investor attention to information events.[6] Google SVI is a noisy measure of investor attention because searches for tickers such as "CAT" are conducted by investors searching for Caterpillar Inc. as well as by internet users searching for felines. Section 4.1 shows that, because of the way SVI is constructed, non-investor search introduces non-additive ME in raw SVI as well as common measures of "abnormal" SVI.

Prior literature readily acknowledges the likelihood for ME in Google search and did their best to exclude presumably noisy tickers from analyses. We directly estimate ME in Google SVI using a unique dataset that provides the websites visited after Google ticker searches and,

---

[5] Negative biases can also occur depending on the form of the nonadditive ME (see Section 2.1).

[6] Published papers using Google SVI are listed in Appendix B. The ME in SVI that we document will also confound tests using SVI as an independent variable. The effects of ME in independent variables are better understood and therefore not a prime focus of this paper.

therefore, allows us to identify non-investor searches.[7] Our analyses include searches for S&P 500 firms' tickers for 2016 and 2017. We find that 69% of all ticker searches result in visits to websites that do not contain investing information, indicating that Google ticker searches are a highly noisy proxy for investor attention. Moreover, our results find that the distribution of ME is highly skewed towards 100% as 135 firms have noise levels exceeding 95%. We also find that ME in SVI is highly correlated with variables such as firm size and analyst following that are commonly used as partitions in cross-sectional tests. Thus, Google SVI is a strong candidate to empirically document the effects of nonadditive ME in dependent variables.

Section 4.5 uses SVI to investigate biased inferences caused by nonadditive ME in dependent variables. We do so using analyses akin to the those in Drake et al. (2012, hereafter "DRT"), which uses abnormal SVI (*ASVI*) to examine changes in investor attention around earnings announcements. The functional form *ASVI* indicates that nonadditive ME proportionally reduces *ASVI* relative to true investor search and, therefore, should bias regression estimates towards zero and inflate standard errors (see Section 4.1). Pooled regressions find an increase in *ASVI* around earnings announcements of 47%. As expected though, regressions find that firms with the least noise search have increases in *ASVI* of 186% (t = 6.19) while firms with the most noise in SVI have insignificant increases in *ASVI* of 0.001% (t = 0.04). Further, the coefficient estimates and t-statistics descend almost monotonically across deciles of noise search. Simulations using induced increases in SVI confirm that the differences in coefficients across firms are not driven by unobservable differences in true investor search. Additionally, simulations find that regressions of *ASVI* including all S&P 500 firms are unable to reliably identify induced increases in SVI of less than 25%, and for the noisiest tercile of firms are unable

[7] Naturally, our "noise search" estimates contain ME of their own. See Section 4.4 for discussion.

to identify induced increases in SVI of up to 500%. In sum, these analyses demonstrate that nonadditive ME in SVI can bias inferences when using SVI in dependent variables.

Section 4 also demonstrates how ME in SVI can seriously perturb inferences from cross-sectional tests. Similar to DRT, we find that increases in *ASVI* around earnings announcements are greater for firms that are larger, have more analysts, and have high bid-ask spreads, indicating that investor attention differs across firms. Our data reveal that noise in SVI is significantly decreasing with all three of these firm characteristics, and simulation tests demonstrate that the correlation between noise and these firm characteristics produces unacceptably high type 1 errors in cross-sectional tests even for modest increases in true investor search. In fact, simply controlling for noise search reduces some cross-sectional test coefficients to insignificant levels. Hence, in the presence of nonadditive ME, researchers should exercise caution in interpreting cross-sectional variation in SVI, and more generally, for many dependent variables in accounting research.

A final set of tests in our case study analysis demonstrate that researchers should be cautious in drawing inferences from relative comparisons between SVI and other measures of investor attention (e.g., EDGAR downloads or Bloomberg searches). For example, Ben-Rephael et al. (2017) compare SVI with Bloomberg activity to conclude that institutional investor attention responds more strongly to news events than does retail attention. However, such inferences are directly influenced by the extent of ME in the two measures, and SVI is likely noisier than Bloomberg activity. We show that tests like those in Ben-Rephael et al. (2017) find opposite results for the higher versus lower deciles of SVI noise. While the broader inferences from Ben-Rephael et al. (2017) are likely unaffected, we caution researchers of making such relative comparisons.

This paper's primary contribution is to inform researchers about how nonadditive ME in dependent variables – which is likely commonplace in accounting research – does not simply reduce test efficiency and "work against finding results." Instead, nonadditive ME can produce positively or negatively biased coefficient estimates, confound interpretations of statistical and economic significance, and lead to both types I and II errors. While we highlight such problems using simulations of key accounting variables and a case study of a common investor attention measure, the documented issues pertain to many commonly used dependent variables used in accounting and finance. Our paper is in a similar vein to other methodological papers in the accounting literature (Gow et al. 2010; Armstrong et al. 2010; Lawrence et al. 2011; Lennox et al. 2012) that are aimed at improving research designs. Unfortunately, in most cases there is no easy fix for unobservable measurement error. Rather, we encourage researchers to carefully consider the extent of ME in dependent variables and, to the extent possible, avoid research designs in which nonadditive ME is a likely confound.

A second contribution of our paper is to the literature using Google SVI as a measure of investor attention. Our intention is certainly not to criticize those past papers; we believe that Da et al. (2011), Drake et al. (2012), and Ben-Rephael et al. (2017) make a substantial contribution by permitting a new wave of research into investor attention as a market friction. Rather, our intention for using SVI as a case study is that our analyses will serve as a guide on how to improve this important proxy for investor attention in future research. To that end, Appendix C lists our noise estimates for each S&P 500 ticker, and we suggest that researchers use this list to refine samples to less noisy firms. Section 4.9 discusses a new SVI measure provided by Google that is intended to more closely capture search for finance and investing topics: Finance-Investor SVI (FISVI). While we are unable to directly estimate ME in FISVI, initial analyses indicate that it is substantially better specified than traditional SVI and should likely be used in future research.

2. The Effect of ME in Dependent Variables

2.1 Modelling the Effects of ME in Dependent Variables – Univariate Case

Consider a population model below where *y'* is regressed on *X* using OLS estimation:[8]

$$y' = \alpha + \beta X + \mu \qquad \textbf{(1a)}$$

The dependent variable *y'* is unobservable so, in empirical estimation, *y* is employed as an observable proxy for *y'*. Variable *y* contains true *y'* plus ME *v*:

$$y = y' + v \qquad \textbf{(2)}$$

where *v* has a mean zero and is uncorrelated with both *X* and μ. Substituting *y'* for (*y* – *v*) and rearranging *v* to the right-hand side yields the following:

$$y - v = \alpha + \beta X + \mu \qquad \textbf{(1b)}$$
$$y = \alpha + \beta X + (\mu + v) \qquad \textbf{(1c)}$$

(1c) is the estimable model. When *v* is additive and uncorrelated with *X* and μ, OLS yields unbiased coefficient estimates, standard errors are larger given the increase in the error variance (i.e., $(\sigma_v+\sigma_\mu)>\sigma_\mu$), and ME decreases the power of tests of β. The intercept α is biased if *v* has a non-zero mean, but the intercept is seldom of interest. Roberts and Whited (2013) highlight that most financial economics studies rely on these basic assumptions to draw inferences from OLS regressions.

Many econometrics texts stop short of discussing the problem arising when the ME, *v*, is nonadditive (but still uncorrelated with α, *X,* and μ). For example:

---

[8] Our example uses a univariate model for simplicity, but *X* could also be a vector of regressors.

$$y = y' / v \qquad \Rightarrow y' = y \times v \tag{3}$$

In this case, the estimable model is now in the following form:

$$y \times v = \alpha + \beta X + \mu \tag{4a}$$

$$y = (\alpha/v) + \beta(X/v) + (\mu/v) \tag{4b}$$

Intuitively, if v is between 0 and 1, then y will be inflated relative to the true value y' and the estimated β is biased away from zero by an amount equal to *average(1/v)*.[9] Assuming v is continuous and uniformly distributed over (a,b) and $a > 0$, the bias in β is equal to:

$$bias_{\beta} = \frac{1}{b-a}\int_{a}^{b}\frac{1}{v}dv$$

While we can derive the bias in this particular case, a general closed form solution is unavailable because bias depends on the specific nonadditive function and the distributional properties of v (e.g., discrete versus continuous, or spanning zero). The extent of efficiency loss typically increases with the variation in v and X and decreases with variation in μ.[10]

2.2 Modelling the Effects of ME in Dependent Variables – Cross-Sectional Tests

Additional confounds are introduced when nonadditive ME is correlated with a partitioning variable in cross-sectional tests. Consider the following model where *Part* is an indicator for a subsample of firms (e.g., large firms, high analyst following, etc.):

For *Part* = 0: $y' = \alpha_1 + \beta_1 X + \mu_1$ **(5a)**

---

[9] Two common misunderstandings are worth addressing. First, the bias in β is not equal to *1/average(v)*, but is rather *average(1/v)*. Second, nonadditive ME does not necessarily mean that the dependent variable itself is directionally biased. For example, for $y' \sim N(0,1)$ and $v = 1.1$, then $E[y' \times v] = E[y'] \times E[v] = 0$, and the bias in y is $E[y'] - E[y' \times v] = 0$. Instead, nonadditive ME increases the variance of y and biases β.

[10] In the simple examples given here and assuming $y, v > 0$, logging the dependent variable can convert nonadditive ME to additive ME and eliminate bias in estimating model (4). As shown below, log transformations cannot eliminate bias in common accounting cases. Of course, log transformations also change the regression interpretation and may not be appropriate in many settings (Brown et al. 1987; Gelman 2008).

For *Part* = 1: $y' = \alpha_2 + \beta_2 X + \mu_2$ **(5b)**

Or equivalently:

Pooled: $y' = \alpha + \phi_1 X + \phi_2 Part + \phi_3 X \times Part + \mu$ **(5c)**

Cross-sectional models are commonly used in accounting studies for two purposes. The first is as a comparative static validation test; e.g., "if the results of a main test are driven by hypothesis H, then we expect the relation between y' and X to be greater in firms with *Part* = 1 (e.g., large firms)." The second use is as a direct hypothesis test; e.g., "hypothesis H is that the relation between y' and X is an increasing function of *Part*." Tests using (5a) and (5b) would predict that $\beta_2 > \beta_1$. Or equivalently, using (5c) the test is $\phi_3 > 0$. While we use a binary *Part* for simplicity, our discussion generalizes to continuous partitions.

Biased cross-sectional tests arise when nonadditive ME in y' differs in the subsample of firms where *Part = 1*. For example, consider the extreme case where $y' = y \times v$, v has a non-zero mean, and v is only present in firms where *Part* = 1. That is, X and v remain uncorrelated, but Corr$(v, Part) \neq 0$:

For *Part* = 0: $y = \alpha_1 + \beta_1 X + \mu_1$ **(6a)**

For *Part* = 1: $y = (\alpha_2/v) + \beta_2(X/v) + (\mu_2/v)$ **(6b)**

Or equivalently:

$$y = \alpha_1 + \phi_1 X + \phi_2(Part/v) + \phi_3(X/v)(Part/v) + (\mu_1 + \mu_2/v) \quad \textbf{(6c)}$$

Model (6a) is unaffected by v while (6b) generates a biased β_2 coefficient in the same manner as in model (4b). Thus, even if the true relation between X and y does not differ with *Part*, tests of $\phi_3 > 0$ can yield spuriously significant results due to the relation between v and *Part*.

3. Illustrating the Effects of ME using A Typical Accounting Sample

3.1 Simulations when ME is uncorrelated with X

Nonadditive ME in dependent variables is likely common in accounting studies because scalars such as assets or equity are measured with error. For example, studies frequently use *ROA* as a proxy for economic performance when investigating hypotheses that firm performance is higher among firms with characteristics such as better governance (e.g. Gompers et al. 2003; Malmendier and Tate 2009), certain accounting policies (e.g. Aboody et al. 1999; Hsu et al. 2017), or CEO traits (e.g. Bertrand and Schoar 2003; Ham et al. 2018). Typical regressions are in the form of:

$$ROA = \alpha + \beta_1 Characteristic + \beta_k Controls_k + \varepsilon \qquad \textbf{(7)}$$

ROA is calculated as earnings (*E*) divided by assets (*A*). Accounting rules introduce ME to both the numerator and denominator in *ROA* as a proxy for performance. ME in earnings can be positive or negative as accruals reverse and earnings converge on cash flows, and therefore is hard to sign in a given period. ME in assets is likely persistent due to rules such as historic cost accounting and expensing of R&D, and is likely proportional to true asset values; e.g., *A_ME = [A(1-v)],* where *v* is between (0,1). Thus, ROA with measurement error, *ROA_ME,* is inflated relative to true *ROA*, and β_1 is biased away from zero by *[average(1/(1-v))]* even if *v* and *Characteristic* are uncorrelated.

As ME is typically unobservable, we use simulations to illustrate how bias in an important accounting-based proxy for economic performance can alter inferences in a generic sample of Compustat firm-quarters (see the Table header for sample details). Table 1, Panel A begins by running a simple univariate regression of performance, proxied by *ROA*, on firm age.

$$ROA = \alpha + \beta_1 log(Age) + \varepsilon \qquad \textbf{(8)}$$

We choose firm age for demonstration purposes only because it is strongly associated with *ROA*.[11] Results in column (i) of Panel A show a strong statistical relation between *ROA* and *log(Age)*: β_1 = 1.12, t = 39.351.

We next randomly assign ME, which we label v, in *Assets* uniformly distributed from 0% to 9%, which should on average introduce a bias of *average(1/(1-v))*= 4.8%.[12] *ROA_ME* is calculated as $[E/(A(1-v))]$. As expected, the average β_1 coefficient is inflated relative to column (i) (β_1 = 1.17, t = 39.349) and the t-statistic is slightly smaller. These effects are amplified in column (iii) with v of [0%,45%], and in column (iv) with v of [0%,90%].[13]

Results in Panels B and C of Table 1 demonstrate the complications induced by attempting to remove nonadditive ME by logging the dependent variable. Analytically, logging *ROA_ME* does the following:

$$log(ROA) = \alpha + \beta_1 Log(Age) + \varepsilon$$

$$log(E) - log(A\_ME) = \alpha + \beta_1 Log(Age) + \varepsilon$$

$$log(E) = \alpha + \beta_1 Log(Age) + \varepsilon + log(A\_ME) \quad \textbf{(9)}$$

A_ME becomes part of the residual, which removes ME from the dependent variable at the cost of reduced efficiency. However, logging introduces two serious problems. First, 23% of the sample has negative *ROA* so undefined *log(ROA)*, which introduces sample selection biases. In this sample, the relation between *ROA* and age is much stronger for loss firms than profit firms (untabulated), so excluding losses likely biases β_1 towards zero. Second, if *A* is correlated with

[11] We are not hypothesizing a reason for statistical association, or assuming that the association would remain in different model specifications. The purpose of this analysis is simply to use simulations in order to illustrate the analytical predictions from Section 2.
[12] Calculated as -*ln*(1-*v*)*(1/0.09).
[13] ME up to 90% is intentionally large for demonstrative purposes. Section 3.2 shows that ME of just 0.5% can cause serious biases in common regressions.

both *Age* and *E*, then *A_ME* in the residual creates a correlated omitted variable. In this particular sample *A* is positively correlated with both *Age* and *E*, so the omitted variable likely biases β_1 away from zero. The net effect of these two biases is unclear.

Column (i) of Table 1, Panel B shows that the relation between *ROA* and *Age* is statistically weaker (β_1 = 9.94, t = 11.852) when *ROA* is logged, even in the absence of ME. The net direction of bias in β_1 is unknown, complicating interpretation. However, columns (ii) through (iv) show that increasing v now has little effect on the coefficient magnitudes, and only a small negative impact on the t-statistics.[14] Thus, logging the dependent variable mitigates ME while introducing new problems.

Panel C shows that the common method of using *log(ROA_ME+1)* again mitigates one problem while creating another. Sample sizes are preserved except for a few extreme outliers where *ROA_ME* < -1, but v remains in the dependent variable because *log(E/A_ME)+1)* expands to [*log(E/A_ME) + log(1+(1/(E/A_ME))*].[15] Thus, β_1 remains biased, as can be seen across columns (ii) through (iv) in Panel C by the increasing β_1 coefficients.[16] Together, the performance simulations in Table 1 illustrate that a small amount of nonadditive ME in the dependent variable can significantly bias the coefficients on the independent variable of interest even when it is uncorrelated with the ME in the dependent variable. Moreover, while logging the dependent variable does mitigate ME in the dependent variable, it imposes sample size restrictions and introduces correlated omitted variables problems.

3.2 Simulations when ME is correlated with X

---

[14] The coefficients are not perfectly equal across the columns due to random correlation generated in the simulations.
[15] Proof: log(a+b) = log(a*(1+b/a)) = log(a)+log(1+b/a).
[16] We must use unwinsorized variables in order for the coefficients and t-statistics to be comparable across columns. However, untabulated tests that winsorize each simulation continue to produce biases that are consistent with our analytical predictions. Robust regressions (Leone et al. 2018 also produce similar results).

In accounting studies, ME in dependent variables likely often correlates with *Characteristic* variables of interest, in which case ME can generate type 1 errors even if no economic relation exists. For example, expensing of R&D and advertising causes ME in asset values, and this ME likely varies systematically with characteristics such as accounting conservatism, industry, firm age, growth, and extent of prior acquisitions. Or, ME in 10-K readability (e.g., FOG) is systematically increasing for firms with debt, investments, subsidiaries, and R&D (e.g., Loughran and McDonald 2011). Hence, cross-sectional analyses in accounting studies are likely confounded by ME in dependent variables.

How big must ME be to generate type I errors in cross-sectional tests? The answer depends on the variables and sample. However, simulations in Table 2 give us some insight using *ROA* and other common accounting variables. These simulations use a binary cross-sectional variable, *Z*, which is randomly equal to one for half of the sample. We run the following regression:

$$ROA\_ME = \alpha + \beta_1 log(Age) + \beta_2 Z + \beta_3 log(Age) \times Z + \varepsilon \qquad \textbf{(10)}$$

Observations with *Z* equal to zero have no ME in *ROA* and firms with *Z* equal to one of ME = v. As before, $ROA\_ME = [E/(A(1\text{-}v))]$. β_1 estimates the slope for firms with zero ME and, therefore, should be roughly the same as the average β estimated in Table 1, Panel A, column (i). In the absence of any ME, we would expect $\beta_3 = 0$ because *Z* is randomly assigned. However, as ME increases and *ROA_ME* inflates, we expect to observe an upwards bias in β_3. If the bias is sufficiently large, we may observe β_3 is statistically significant; i.e., a type 1 error.

We estimate (10) a total of 1,000 times for each levels v from 0% to 4.5%. Panel A of Table 2 plots the results. The horizontal axis plots results for levels of v. The flat curve (left axis) shows that the average β_1 coefficient is 1.12, consistent with Table 1. The hashed line (left axis) plots the sum of ($\beta_1 + \beta_3$) and, as expected, increases proportionally with *v*. The dotted line (right axis) shows the number of trials finding ($\beta_3 - \beta_1$) $\neq 0$ at a 5% level of confidence; i.e., type 1 errors generated by *v*. In second group with *v*=0.5%, 5.6% of trials generate type 1 errors, which is above the 5% threshold. Untabulated analyses using a 10% level of confidence find an excessive type 1 error rate of 11.9%. Put simply, these results show that as small as a 0.5% difference in *v* between $Z = 0$ versus $Z = 1$ can generate spurious results in a typical accounting sample. Differences in *v* of 4.5% (rightmost group) produce a type 1 error rate of 81% using a 5% level of confidence. Panel B performs a similar analysis for *Tobin's Q.* The unconditional slope of regressing *Tobin's Q* on *log(Age)* is -0.26 and unacceptably high type 1 error rates are observed for *v* = 1% and higher. Panel C finds that regressions of *Investments* generate an unacceptable type 1 error rate with just *v*=0.5%. Hence, the above analyses illustrate that ME in key accounting variables as small as 0.5% percent could lead to excessive type 1 errors. This is small considering that an error of 0.5% of assets would often be considered immaterial in an audit (Eilifsen and Messier 2015).

Are ME in assets of 1% likely to occur in accounting research? Almost certainly yes considering flexibility in accounting standards and conservative nature of accounting net book values. For example, firms' brands and intellectual properties often make up large percentages of market values, and they are for the most part, not included in net book values. Moreover, prior research indicates that measurement error can be very large. For example, McNichols et al. (2014) analytically estimate that conservatism in expensing of advertising and R&D overstates market-to-book ratios by an average of roughly 187%, with the ME being much larger in certain industries. Or, Kothari et al. (2005) report that the 25^{th} (75^{th}) percentiles of estimated discretionary accruals range from -4.6% to -7.5% of assets (4.4% to 7.4% of assets). Given that the typical interquartile range of ROA spans roughly 1% to 9%, the discretionary accruals in Kothari et al. (2005) seem extraordinarily large. Lastly, in our study of Google SVI below, we find ME frequently approaches 10,000% of the actual investor search values. Hence, ME > 1% is likely commonplace in accounting variables.

4. Google SVI Case Study

4.1 ME in Google SVI as a Measure of Investor Attention

To demonstrate the effects of nonadditive ME, we use Google ticker search volume, a commonly used empirical measure of retail investor attention introduced by Da et al. (2011) and Drake et al. (2012) as a case study to further support our simulation analyses in Section 3. This measure, labelled Google Search Volume Index (SVI), has become a standard proxy in the rapidly growing research examining investor attention. Google SVI for firm *i* in period *t* in geographic location *g* is calculated as follows:

$$SVI'_{i,t} = \left(\frac{Keyword\_Search_{i,t} / Geo\_Search_{g,t}}{max_w \left[Keyword\_Search_{i,t} / Geo\_Search_{g,t} \right]_w} \right) * 100 \qquad \textbf{(11)}$$

where *Keyword_Search* in accounting and finance studies is typically firms' tickers. The scalar $Geo\_Search_{g,t}$ is the total Google searches in geography *g* during the period *t*. Google scales by $Geo\_Search_{g,t}$ to facilitate comparisons of "relative popularity" of given keyword across geographies within a given period, "otherwise places with the most search volume would always be ranked highest."[17] The geographic region is typically set to the United States in studies of U.S. firms. The denominator is the maximum scaled search for firm *i* observed over the time

[17] https://support.google.com/trends/answer/4365533?hl=en&ref_topic=6248052. Accessed March 2018. Examining relative keyword popularity across geographies within a given period is likely useful for firms' marketing decisions, which is the primary function for which Google Trends was originally developed.

window *w*, such that SVI represents the within-firm relative keyword search popularity on a scale of 0 to 100.[18] A common usage of SVI in accounting research is to investigate variation in investor attention around events, using a model such as $SVI' = \alpha + \beta_1 Event + \mu$.

Ambiguous tickers introduce an unknown amount of ME into SVI as a proxy for investor attention. Thus, *Keyword_Search* includes search by investors (*Investor_Search*) as well as noise search by non-investors (*Noise_Search*). Together, equation (11) can be expressed as the observable SVI as follows:

$$SVI_{i,t} = \left(\frac{(Investor\_Search_{i,t} + Noise\_Search_{i,t}) / Geo\_Search_{g,t}}{max_w \left[(Investor\_Search_{i,t} + Noise\_Search_{i,t}) / Geo\_Search_{g,t} \right]_w} \right) * 100 \qquad \textbf{(12)}$$

Noise_Search inflates both the numerator and denominator of (12), introducing nonadditive ME to SVI as a measure of investor attention.[19] While it is not completely clear whether *Noise_Search* over- or under-states *SVI* relative to *SVI'*, the maximum operator in the denominator likely means that average *SVI* is understated and coefficient estimates are biased towards zero.

Researchers often attempt to mitigate the effects of ME by creating an abnormal measure of SVI. One common abnormal SVI measure, *ASVI*, is calculated by subtracting the average SVI observed over a pre-event control window, and then the difference is scaled by the pre-event average. Like Drake et al. (2012), our pre-event control window is defined as the firm's SVI on the same weekday over the trailing 10 weeks, which helps to eliminate systematic variation in

---

[18] By construction, SVI is a within-firm measure and cannot be directly used to draw inferences about relative magnitudes in search volume across firms.

[19] Variation in *Geo_Search* is also a source of noise; e.g., even for identical *Investor_Search* on two days, variation in *Geo_Search* can cause variation in *SVI*. We disregard *Geo_Search* in our discussion because it does not vary across tickers within the U.S. and, thus, likely generates fewer systematic biases than does *Noise_Search*.

Geo_Search. This approach assumes that variation in *Noise_Search* and *Geo_Search* do not materially correlate with the event being investigated, which is not an assumption we necessarily dispute.[20] However, even with generous assumptions that $Noise\_Serch_{i,t} = Noise\_Search_{i,t-j}$ and that $Geo\_Search_{g,t} = Geo\_Search_{g,t-j}$, *ASVI* still contains nonadditive ME:

$$ASVI = \left(\frac{Investor\_Search_{i,t} - \overline{Investor\_Search}_{i,t-j}}{max_w\left[(Investor\_Search_{i,t} + Noise\_Search_{i,t}\right]_w}\right) \quad \textbf{(13)}$$

Noise_Search remains in the denominator, and in turn, average *ASVI* is understated relative to *ASVI',* and regression coefficients should be biased towards zero. Da et al. (2011) use a slightly different abnormal measure based on the difference between logged event-window SVI minus logged trailing average (*ASVI2)*, but again *Noise_Search* remains in the denominator.

In sum, nonadditive ME in SVI remains even in abnormal specifications and accordingly, SVI as a measure of investor attention is an appropriate candidate to illustrate the effects of nonadditive ME in dependent variables. The ME in SVI should result in biased regression coefficients when any of *SVI, ASVI*, or *ASVI2* are used as dependent variables. It is unclear *ex ante* how the ME in these variables correlate with cross-sectional partitions such as firm size; for example, large firms have shorter tickers that may have higher *Noise_Search*, or large firms may attract greater investor attention around events so *Noise_Search* will have a smaller effect.

4.2 Sample Selection

Table 3, Panel A details our sample selection for the Google search analysis. Our sample includes S&P 500 firms as of January 1st, 2016. We include tickers for all share classes, yielding 511 tickers. Our sample spans 2016 through 2017. We download SVI data from Google for each

[20] However, one could easily imagine a situation where *Noise_Search* does correlate with the events being examined. For example, a product launch might generate investor search as well as search by customers.

ticker, and construct a daily dataset using the procedures described in Appendix A. We drop two tickers for which SVI is unavailable. For consistency, we drop 19 firms with ticker changes during our sample period. Lastly, we require each firm to have necessary variables in Compustat, CRSP, I/B/E/S, and FactSet. Our final sample includes 481 firms, 490 tickers, and 245,015 trading days. Summary statistics are provided in Panel B of Table 3 and variable definitions are defined in Appendix A.

4.3 Estimating Investor_Search versus Noise_Search

We assess ME in SVI by assessing whether ticker searches are made by investors searching for current information about the ticker in question. We make this determination using proprietary data on the websites visited following ticker searches, which we label ticker "click-throughs." The data include click-throughs for each website as a fraction of total click-throughs during the month. The data were obtained by a marketing firm from a variety of sources including internet service providers, browser trackers, and data sharing agreements with major websites. The marketing firm sells web traffic data for commercial purposes and report an accuracy rate of over 99%. Ticker searches that result in click-throughs to websites that contain investment-related information are designated as *Investor_Search*. Ticker searches that result in click-throughs to other websites are designated as *Noise_Search*.[21]

We use data from the same marketing firm to make an initial assessment of whether each click-through website has investment-related content. As shown in Table 4, 35.3% of all click-throughs go to websites categorized as "Shopping." The next highest categories are "Unknown" at 17.0% and "Finance" at 9.6%. Based on these categories alone, it appears that many ticker

[21] Our method of identifying *Investor_Search* versus *Noise_Search* contains measurement error from several sources. See Section 4.4 for further discussion.

searches are likely *Noise_Search*. Rather than relying solely on the marketing firm's categorizations, we manually review websites to determine whether they contain investor-related information. This determination requires subjectivity and we applied the coding rules below. Incorrect classifications of *Investor_Search* means introduces ME, the effects of which we discuss in Section 4.4. However, except for the first rule, we use the same website classifications for all firms (e.g., wsj.com is designated as investor-related for all tickers), which reduces the risk that ME from misclassifications varies systematically across firms.

1) Firms' investment-specific domains are *Investor_Search* (e.g., investor.fb.com). Firms' commercial homepages are *Noise_Search* (e.g., facebook.com). While visits to commercial webpages could be performed by investors gauging the company's products or services, the volume of visits indicates that most visits to commercial websites are not by investors (e.g., 97% of all ticker searches for "CVS" go to cvs.com).[22] Still, reperforming our analyses in Table 6 while including commercial homepages visits as *Investor_Related* produces unchanged inferences.
2) News and Media websites are *Investor_Search* if they contain primarily financial news (e.g., marketwatch.com and thestreet.com). News and media websites primarily containing general-interest news are *Noise_Search* (e.g., people.com and espn.com).
3) Trading websites such as wfadvisors.com or fidelity.com are *Investor_Search*. Visits to retail bank websites such as wellsfargo.com are *Noise_Search*.

Reviewing every click-through website is costly, so instead we take a sampling approach. We start by reviewing the top ten click-through websites for each ticker-month. If the top ten

[22] Untabulated robustness tests find insignificant differences in firm characteristics between those that have a separate investor relations domain versus those that do not (e.g., investor.company.com versus company.com/investor).

websites do not comprise at least 70% of the total traffic, we review additional websites until at least 70% of traffic is covered. To ensure that we have good coverage across website categories, we also audit a minimum of 70% of traffic within each website category. As shown in Table 4 Panel A, following these procedures means that we audit 94% of all website traffic. For unaudited websites, we use the category's average *Investor_Search* to estimate investor-related search. For example, Table 4 shows that 64.8% of the "Finance" category audited web traffic is designated at *Investor_Search*, but that we did not audit 1.4% of "Finance" website traffic. So, we assume that 64.8% of the Finance category's unaudited website traffic is *Investor_Search*. Panel B of Table 4 lists the top 20 website domains across all categories that are designated as *Investor_Search*.

Table 3, Panel B shows that our sample average *Investor_Search* is 0.311, indicating that 31.1% of Google ticker searches are performed by investors. The remaining 68.9% of ticker searches are designated as *Noise_Search*. Figure 1 provides a histogram of *Noise_Search* for each ticker and shows that it is highly skewed, with 125 tickers having *Noise_Search* of over 90%. Columns (ii) and (iii) of Appendix C list *Investor_Search* and *Noise_Search* for each of the 490 tickers included in our sample.

Table 5 analyses *Investor_Search* by observable firm characteristics. Da et al. (2011) and Drake et al. (2012) note that tickers with ambiguous meanings (e.g., BABY and CAT) and tickers that are also brands (ABC and UPS) likely have high non-investor search. Tickers that are shorter (one or two letters) are also likely to be noisier. Panel A of Table 5 shows that these intuitions are correct. Ambiguous, one-letter tickers, and two-letter tickers have 15.1%, 6.6%, and 14.4% *Investor_Search*. Tickers of three- through five-letters have 27.7%, 53.0%, and 61.3% *Investor_Search*, indicating that more unique tickers have less ME. However, Appendix C shows that there are many deviations from these trends. For example, of the 30 tickers with less than 1% *Investor_Search*, 28 tickers have three or more letters. Moreover, VZ has 68% *Investor_Search* despite being only two letters. Panel B of Table 5 shows that there is substantial variation in *Investor_Search* across the Fama-French 12 industries. Panel C of Table 5 shows that *Investor_Search* varies systematically across a variety of firm characteristics, many of which are commonly used as partitioning variables in cross-sectional tests.

4.4 ME in our estimate of Investor-Search versus Noise-Search

Our estimates of *Noise_Search* suffer from their own ME. First, our data only allow us to reliably estimate each firm's *Noise_Search* over the pooled two-year period, while actual *Noise_Search* likely varies over time and could possible correlate with earnings announcements. Second, our classifications of websites as *Investor_Search* versus *Noise_Search* are imperfect. Third, we cannot observe ticker searches that did not result in a website click-through; e.g., if an investor learns solely from the stock information boxes that Google returns for some ticker searches.[23] Fourth, we cannot observe variation in *Geo_Search*. These sources of ME mean that our assignments of observations to *Noise_Search* deciles below are noisy unto themselves, but we have no reason to believe that ME in *Noise_Search* systematically confounds our inferences. Further, ME in *Noise_Search* is unlikely to materially affect our inferences because we are not testing hypotheses, but rather simply aim to illustrate how noisy dependent variables can confound OLS regressions. Still, the extent and effects of ME are unobservable, so it is possible that they cause unanticipated confounds.

4.5 Measurement Tests of Investor Search Increases Around Earnings Announcements

[23] In untabulated tests we exclude all tickers for which a Google search produced a stock information box as of August 2018. The inferences from our main analyses in Table 6 are unchanged.

This section investigates the effects of ME in SVI using analyses like those in Drake et al. (2012, "DRT"), which investigates investor search around earnings announcements.[24]

$$Search_{it} = \beta_0 + \beta_1 EA_{it} + \beta_{2 \ldots n} Controls_{it} + \varepsilon \qquad \textbf{(14)}$$

$Search_{it}$ is one of $SVI_{it,}$ $ASVI_{it}$, or $ASVI2_{it}$. SVI_{it} is the level of Google SVI. $ASVI_{it}$ is abnormal SVI based on percentage change (similar to DRT) and $ASVI2_{it}$ is the logged SVI minus the logged trailing average (similar to Da et al. 2011) (see Section 4.1 and Appendix A). EA_{it} is an indicator variable equal to one on earnings announcement days. $Controls_{it}$ are similar to DRT and include: *News Articles*, *Abs Return, MVE, Analysts, Trading Volume, Spread, Fourth Qtr, Total EAs, Inst Own, BTM,* and *year-week fixed effects*. Standard errors are clustered by firm.

Panel A of Table 6 provides results of equation (14) excluding controls and fixed effects. The leftmost column presents results for the pooled sample. The upper rows display results for *SVI,* middle rows for *ASVI*, and lower rows for *ASVI2*. All three measures find highly significant increases in search around earnings announcements (note: the coefficient magnitudes cannot be compared across *SVI, ASVI,* and *ASVI2* due to different functional forms). Columns (ii) through (xi) rerun equation (14) by decile of *Noise_Search* calculated at the ticker level. The results clearly show that both statistical and economic significance of β_1 decrease in deciles of *Noise_Search*, becoming insignificant by the highest decile. The adjusted r-squared also declines across deciles. The exception is that *Noise_Search* decile 1 tends to have a smaller adjusted r-squared than decile 2, which is likely due to imperfect measurement of *Noise_Search*. The trends for *SVI, ASVI,* and *ASVI2* are highly similar, indicating that the abnormal transformations in *ASVI* and *ASVI2* are ineffective in eliminating ME. All three measures perform worse in Panel B

[24] DRT also examine search around other announcements. We focus on earnings announcements for simplicity and because DRT find that they elicit the largest increases in ticker search. The same econometric issues would apply to search around any event.

of Table 6 once controls and fixed effects are added. Figure 2, Panel A graphically illustrates the findings for *ASVI* of Table 6, Panel B. In sum, the results in Panels A and B of Table 4 and Panel A of Figure 2 strongly indicate that ME in SVI as a dependent variable produces downward biased coefficient estimates.

Panel C of Table 6 evaluates the sensitivity of the Panel B results to following some prior papers' methods of dropping tickers that are assumed to be noisy. Da et al. (2011) note that guessing at noisy tickers introduces subjectivity in sample construction and, therefore, report results including all tickers but note that untabulated tests excluding ambiguous tickers are similar. DRT exclude likely ambiguous tickers in all tests. Ben-Rephael et al. (2017) exclude ambiguous tickers in certain tests but keep all tickers in tests around earnings announcements. In addition to introducing subjectivity, another problem with guessing at noisy tickers is that papers often to not report the tickers that are excluded (e.g., deHaan et al. 2015), which complicates replication. Our analyses in Panel C of Table 6 drop the ambiguous tickers identified by DRT and are listed in the Table 5 header. We do not re-form the deciles of *Noise_Search* and comparing the samples sizes from Panels B and C shows that the dropped tickers tend to be concentrated in the upper deciles. However, many firms remain and the attenuated regression coefficients are still clearly evident across deciles. Panel D further drops all one- and two-letter tickers, but again attenuation is clearly visible across deciles of *Noise_Search*. Thus, ad hoc approaches to dropping noisy tickers do not resolve the problem.

4.6 Google-Search Simulation Tests

A concern with the analyses in Table 6 is that it is possible that *Investor_Search* around earnings announcements ("EAs") is lower for firms that have higher *Noise_Search*, in which case it is impossible to isolate the effects of ME. This section addresses this concern using simulations in which we induce specified increase in *Investor_Search* around random dates. For brevity these tests use only *ASVI*. Our procedures are as follows:

1) Drop all EA days and replace each with a randomly selected non-EA day (*Random_Day*).
2) Induce a specific amount of *Investor_Search* on each *Random_Day*. For example, the ticker UNM has *Noise_Search* of 99.2%, so inducing a 100% increase in *Investor_Search* increases SVI by 0.8.[25] For a ticker with 0% *Noise_Search,* inducing a 100% increase in *Investor_Search* increases SVI by 100. Calculate *ASVI* using the updated data.
3) Estimate model (10) where *Random_Day* replaces *EA* to see whether the model rejects the null that the *Random_Day* coefficient is equal to zero. Use a 5% level of confidence.
4) Repeat this process 1,000 times selecting *Random_Day* with replacement.

Table 7 summarizes the simulation results. Controls and fixed effect estimates are untabulated. The uppermost rows have results for a 0% increase in *Investor_Search*. As expected, the average coefficient estimates are all close to 0 and 0.1% of trials in the pooled sample (column (i)) reject the null hypothesis at a five percent level of confidence. Thus, there is no evidence of ME in SVI generating Type I errors.

The lower rows test induced increases in SVI ranging from 5% to 500%. Starting with 5%, column (i) finds an average pooled coefficient of 0.018. This finding is roughly as expected given that the sample average *Investor_Search* is roughly 31%; i.e., 5% inducement times 31% *Investor_Search* = 0.015. Just 2.8% of trials reject the null hypothesis that there is no increase in SVI, indicating that a pooled sample is unlikely to identify a 5% increase in *Investor_Search*. Further, the coefficient estimates in columns (ii) through (xi) for each decile of ME are also

[25] To facilitate interpretation of our regression coefficients, we do not rescale SVI from 0 to 100 after inducing *Investor_Search*. However, rescaling SVI after the inducement has minimal impact on our results.

insignificant, indicating that small increases in investor ticker searches are likely unidentifiable even among the least noisy SVI deciles.

As expected, the models' rejection rates improve as the induced increase grows. At 20% the pooled model identifies an increase at the 5 percent level in SVI in 86.7% of trials. Moreover, the deciles continue to perform poorly, especially those higher in noise search. An inducement of 25% in the pooled model rejects the null in 96.5% of trials, a 50% inducement rejects the null in 99.9% of trials, and 100% to 500% inducements identify increases in 100% of trials (pooled). However, the highest three noise deciles continue to perform poorly. Figure 2, Panel B visually illustrates the results of Table 7.

The main takeaway from Table 7 and Figure 2, Panel B is to confirm that *Noise_Search* causes attenuated coefficient estimates. Another key takeaway is that, at least within our sample, increases in investor search smaller than 25% are not reliably identifiable. This issue is unlikely to be problematic in studying events such as earnings announcements that likely generate large increases in investor search. However, researchers should exercise caution in trying to examine increases in investor search around less significant events.

4.7 Erroneous Inferences in Cross-Sectional Tests Using Google Search

As illustrated in Sections 2.2 and 2.4, problems arise in cross-sectional tests if *Noise_Search* correlates with the cross-sectional variable of interest:

$$Search_{it} = \beta_0 + \beta_1 EA_{it} + \beta_2 Partition_{it} + \beta_3 EA_{it} Partition_{it} + \beta_{4 \ldots n} Controls_{it} + \varepsilon \qquad \textbf{(15)}$$

If *Partition* is correlated with *Noise_Seach*, then the β_3 coefficient will be biased. DRT perform cross-sectional analyses and find that increases in search around earnings announcements differ for firms in the highest quintile of market capitalization, analyst following, and bid-ask spread.[26] Table 5, Panel C and Figure 3 show that *Noise_Search* is decreasing with size, analyst following, and spread, indicating that ME may bias results. We start by running cross-sectional tests where the partitioning variables are indicators: *Large_Firms*, *High_Following*, and *Large_Spread*.

As seen in Panel A of Table 8, the interactions in columns (i), (iii), and (v) on *EA×Large_Firms*, *EA×High_Following*, and *EA×Large_Spread* are significantly different from zero, indicating that search around earnings announcements differs across firms. Columns (ii), (iv), and (vi) control for *Noise_Search* and *EA×Noise_Search*, which should reduce the bias caused by ME in SVI. However, because our measure of *Noise_Search* is itself noisy, these results provide a lower bound on the effects of eliminating ME in SVI. In columns (ii) and (vi), the interaction terms on *EA×Large_Firms* and *EA×Large_Spread* become insignificant at the 10% level after controlling for ME. The coefficient on *EA×High_Following* in column (iv) remains statistically significant but is attenuated. Thus, these results indicate that cross-sectional variation in *Noise_Search* very likely biases, or even generates, common cross-sectional results.

Panel B of Table 8 re-runs our previous simulations while including interaction terms with the three foregoing cross-sectional variables. The objective of these simulations is to evaluate how frequently cross-sectional tests reject the null hypothesis that β_3 is equal to zero even though the actual increase in *Investor_Search* is equal across firms (i.e., produce Type 1 errors). The upper rows are for interactions of *Random_EA×Large_Firms*. 0% of trials reject the null with a 25% inducement in *Investor_Search*. However, with a 50% increase of *Investor_Search*, which is quite plausible around corporate announcements, 28% of trials reject

[26] DRT also do cross-sectional tests based on idiosyncratic volatility. We exclude volatility for brevity and because DRT do not provide a variable definition.

the null. This type I error rate is above most studies' acceptable thresholds of a five or ten-percent level of confidence. The middle and lower rows of Panel B show that cross-sectional tests of *High-Following* and *Large_Spread* perform slightly better, but both exceed a five percent Type I error rate when *Investor_Search* is at least 50%. Together, the results in Panels A and B demonstrate that ME in dependent variables can seriously confound cross-sectional analyses.

4.8 Erroneous Inferences in Comparing Different Dependent Search Variables

Nonadditive ME can also cause erroneous inferences when comparing different dependent variables across models. For example, Ben-Rephael et al. (2017, hereafter "BRDI") compare retail investor attention to institutional investor attention around corporate news events, using SVI to measure retail attention and Bloomberg terminal activity to measure institutional investor attention. BRDI's first primary result in Table 3 finds that news events explain 5.14% of the variation in institutional investor attention but only 0.15% of retail investor attention. A concern is that these results are confounded by differential ME in the two measures of attention. While SVI is noisy, Bloomberg terminal activity is a relatively low-ME measure of institutional investor attention because non-investors do not use Bloomberg terminals and because Bloomberg has unambiguous company identifiers.

Column (i) of Table 9 performs analyses similar to those in Table 3 of BRDI. The dependent variable in Panel A, *AIA*, is an indicator variable for high institutional investor attention (see Appendix A). We regress *AIA* on an indicator for an earnings announcement, *EA*, and find a pseudo r-squared of 0.06. Panel B of Table 9 uses the dependent variable *DADSVI*, which is an indicator variable for high retail investor attention as in BRDI (see Appendix A). Regressing *DADSVI* on *EA* in column (i) of Panel B produced a pseudo r-squared of 0.02. Panel C tests the difference in explanatory power using a Vuong test in which *EA* is regressed on each of *AIA* and *DADSVI* using OLS, and finds that the difference in explanatory powers is highly significant.[27] In sum, our results are qualitatively the same as in BRDI: value-relevant news events are larger drivers of institutional investor attention than of retail investor attention.

Columns (ii) through (xi) of Table 9 repeat the analyses in column (i) but by decile of *Noise_Search*. In Panel A the coefficient estimates and pseudo r-squareds for *AIA* are similar across deciles, potentially with a small upward trend in the highest deciles. The upward trend in the higher deciles is either because increases in *AIA* around earnings announcements are larger for the types of firms with more ME in SVI, or because ME in *AIA* is lower in those firms. As expected, in Panel B the coefficient estimates and pseudo r-squared for *DADSVI* tend to decline across the noise deciles. Columns (ii) through (iv) find that the pseudo r-squareds are actually larger for *DADSVI* than *AIA* when ME is minimized. Thus, the inferences from BRDI are eliminated or reversed for firms with the least noisy tickers.

In sum, these results demonstrate that ME in SVI can cause biased inferences when comparing across different dependent variables. A similar concern is relevant to comparisons of SVI to other investor attention dependent variables (e.g., Drake et al 2016; Ben-Rephael et al. 2017). Differential ME is also a concern when using SVI as an independent variable in comparison to other measures of attention (e.g., deHaan et al. 2015; Drake et al. 2016).

4.9 Google SVI – Guidance for Future Research

[27] In OLS univariate regressions, regressing Y on X produces an identical r-squared as adjusting X on Y. Thus, reversing the EA and attention variables in our Vuong tests is valid. Also, we note that our pseudo r-squareds for both *AIA* and *DADSVI* are higher than in BRDI. A likely but unverifiable explanation is that our sample of S&P 500 firms generate larger spikes in attention around earnings announcements than do the Russell 3000 firms used in BRDI. Finally, sample size reduction in Table 9 is due to requiring Bloomberg data.

The purpose of this paper is to illustrate the effects of nonadditive ME in dependent variables. In no way do we intend to suggest that SVI should no longer be used as a measure of investor attention. Quite the opposite, SVI as a measure of investor attention allows researchers to examine questions that were previously inaccessible. DRT, BRDI, Da et al. (2011), and the other papers mentioned herein have provided an important and lasting contribution to the accounting literature. This section is intended to facilitate future research by providing guidance on how to most effectively use Google search as a proxy for investor attention.

First, researchers should be sure to consider biases in using SVI in pooled analyses, and in particular the potential for spurious results in cross-sectional tests. Appendix C lists *Noise_Search* for each S&P 500 ticker, so researchers can either restrict their samples to low-noise tickers or perform sensitivity tests excluding these firms.[28] This data-driven approach is preferable to ad hoc approaches used in prior papers. Alternatively, Table 5 illustrates that tickers with four or more letters have the more than 53.0% percent investor search; and accordingly, robustness analyses could be performed with such tickers where there is substantially less noise.

We also consider a measure of SVI based on refined sub-category "Finance – Investor" searches provided by Google (Category 107, hereafter "FISVI"). Appendix D provides a guide on how to select this category when downloading SVI data. We use this FISVI to create a set of variables that are analogous to *SVI*, *AVSI*, and *ASVI2*: *FISVI*, *AFIVSI*, and *AFISVI2*. Our data do not allow us to observe which websites Google classifies under FISVI so we cannot directly estimate noise search in this new measure. However, we are able to assess how well this alternate measure performs across deciles of SVI *Noise_Search*.

[28] As resources permit, we aim to also provide a website with *Noise_Search* for the Russell 3000 firms.

Table 10 repeats the analyses of Panel B of Table 6 but using *FISVI*, *AFIVSI*, and *AFISVI2*. We include controls and fixed effects to best mimic how these variables would be used in future studies' regressions. In the pooled regressions in column (i) of Table 10, the coefficients for *AFISVI* and *AFISVI2* are considerably larger than the comparable models in Table 6 Panel B, and the t-statistics are higher in Table 10 for all three variables. Moreover, *FISVI, AFISVI*, and *AFISVI2* in Table 10 tend to demonstrate less attenuation across the *Noise_Search* deciles than do the results in Table 6 Panel B, although the results for the highest deciles of *Noise_Search* in Table 10 are still insignificant. Taken together, it appears that abnormal measures based on the FISVI performs better than the corresponding SVI-based measures, indicating that FISVI likely has less ME. While we cannot directly assess the ME in FISVI, it appears that using a FISVI-based measures likely produces better specified tests in examining investor attention.

5. Conclusion

This study illustrates the importance of carefully considering the extent of measurement error (ME) in dependent variables and how it can significantly bias inferences in accounting studies. While *additive* ME in dependent variables is often benign, *nonadditive* ME, which is likely commonplace in accounting research, leads to upward or downward biased coefficient estimates and can generate both type 1 and 2 errors. We demonstrate the potential confounding effects of nonadditive ME in dependent variables using simulations and a case study of Google ticker search. We provide suggestions for mitigating the effects ME in Google ticker search, and caution researchers that thoughtful consideration is needed to adequately address the effects of nonadditive ME in dependent variables in accounting research.

REFERENCES

Aboody, D., Barth, M.E. and Kasznik, R., 1999. Revaluations of fixed assets and future firm performance: Evidence from the UK. *Journal of Accounting and Economics*, 26(1-3), pp.149-178.

Armstrong, C., A. Jagolinzer, and D. Larcker. 2010. Chief executive officer equity incentives and accounting irregularities. *Journal of Accounting Research* 48 (2): 225-271.

Barth, M., and W. Landsman. 1995. Fundamental issues related to using fair value accounting for financial reporting. *Accounting Horizons* 9 (4): 97.

Basu, S., 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24 (1): 3-37.

Ben-Rephael, A., Z. Da, and R. Israelsen. 2017. It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies* 30 (9): 3009-3047.

Bhagat, S. and B. Bolton. 2008. Corporate governance and firm performance. *Journal of corporate finance* 14(3): 257-273.

Brown, L., G. Richardson, and S. Schwager. 1987. An information interpretation of financial analyst superiority in forecasting earnings. *Journal of Accounting Research* 25 (1): 49-67.

Cameron, C., and Trivedi, P. 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press.

Chen, W., P. Hribar, and S. Melessa. 2017. Incorrect inferences when using residuals as dependent variables. *Journal of Accounting Research* 56 (3): 751-796.

Da, Z., J. Engelberg, and P. Gao. 2011. In search of attention. *The Journal of Finance* 66 (5): 1461-1499.

deHaan, E., T. Shevlin, and J. Thornock. 2015. Market (in) attention and the strategic scheduling and timing of earnings announcements. *Journal of Accounting and Economics* 60 (1): 36-55.

Drake, M., D. Roulstone, and J. Thornock. 2012. Investor information demand: Evidence from Google searches around earnings announcements. *Journal of Accounting Research* 50 (4): 1001-1040.

Drake, M., J. Jennings, D. Roulstone, and J. Thornock. 2016. The comovement of investor attention. *Management Science* 63 (9): 2847-2867.

Easton, P., and M. Zmijewski. 1989. Cross-sectional variation in the stock market response to accounting earnings announcements. *Journal of Accounting and Economics* 11 (2-3): 117-141.

Eilifsen, A. and W. Messier. 2015. Materiality guidance of the major public accounting firms. *Auditing: A Journal of Practice and Theory* 34: 3–26.

Fields, T., T. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31 (1-3): 255-307.

Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* 27 (15): 2865-2873.

Gerakos, J. 2012. Discussion of detecting earnings management: A new approach. *Journal of Accounting Research*, 50 (2): 335-347.

Givoly, D., C. Hayn, and A. Natarajan. 2007. Measuring reporting conservatism. *The Accounting Review* 82 (1): 65-106.

Gompers, P., J Ishii, and A Metrick, 2003. Corporate governance and equity prices. *Quarterly Journal of Economics* 118, 107-155.

Gow, I., G. Ormazabal, and D. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (2): 483-512.

Ham, C., Seybert, N., & Wang, S. 2018. Narcissism is a bad sign: CEO signature size, investment, and performance. *Review of Accounting Studies*, 23(1): 234-264.

Holthausen, R. 1990. Accounting method choice: Opportunistic behavior, efficient contracting, and information perspectives. *Journal of Accounting and Economics* 12 (1-3): 207-218.

Hsu, C., Novoselov, K.E. and Wang, R., 2017. Does Accounting Conservatism Mitigate the Shortcomings of CEO Overconfidence?. *The Accounting Review*, 92(6), pp.77-101.

Kothari, S.P., A. Leone, and C. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163-197.

LaFond, R., and R. Watts. 2008. The information role of conservatism. *The Accounting Review* 83(2): 447-478.

Lawrence, A., M. Minutti-Meza, and P. Zhang. 2011. Can Big 4 versus non-Big 4 differences in audit-quality proxies be attributed to client characteristics? *The Accounting Review* 86 (1): 259-286.

Lawrence, A., R. Sloan, and Y. Sun. 2013. Non-discretionary conservatism: Evidence and implications. *Journal of Accounting and Economics* 56 (2-3): 112-133.

Lawrence, A., J. Ryans, E. Sun, and N. Laptev. 2018. Earnings announcement promotions: A Yahoo Finance field experiment. Forthcoming *Journal of Accounting & Economics*.

Leone, A., Minutti-Meza, M., and Wasley, C. 2018. Influential Observations and Inference in Accounting Research. Working paper available on *SSRN*.

Lennox, C., J. Francis, and Z. Wang. 2012. Selection models in accounting research. *The Accounting Review* 87 (2): 589-616.

Loughran, T., and B. McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance* 66(1): 35-65.

Malmendier, U. and G. Tate. 2009. Superstar CEOs. *The Quarterly Journal of Economics.* 124(4):1593–1638.

McNichols, M. 2000. Research design issues in earnings management studies. *Journal of Accounting and Public Policy* 19 (4-5): 313-345.

McNichols, M., M. Rajan, and S. Reichelstein. 2014. Conservatism correction for the market-to-book ratio and Tobin's Q. *Review of Accounting Studies* 19 (4): 1393-1435.

Roberts, M. and T. Whited. 2013. Endogeneity in empirical corporate finance. In *Handbook of the Economics of Finance* (Vol. 2, pp. 493-572). Elsevier.

Schipper, K., and L. Vincent. 2003. Earnings quality. *Accounting Horizons* 17 (Supplement): 97-110.

Subramanyam, K.R., and J. Wild. 1996. Going-concern status, earnings persistence, and informativeness of earnings. *Contemporary Accounting Research* 13(1): 251-273.

Vlastakis, N., and R. Markellos. 2012. Information demand and stock market volatility. *Journal of Banking & Finance* 36 (6): 1808-1821.

Watts, R., and J. Zimmerman. 1990. Positive accounting theory: A ten year perspective. *Accounting Review* 65 (1): 131-156.

Wooldridge, Jeffrey. 2012. *Introductory Econometrics: A Modern Approach*. South-Western, Cengage Learning.

Figure 1 – Histogram of *Noise Search* Across Tickers

This figure shows the distribution of the variable *Noise_Search* for the 490 tickers in our final sample. The Y-axis is the number of observations (i.e., tickers) and the X-axis is *Noise_Search* variable ranging from 0% to 100%. The reference line represents the mean of *Noise_Search* (at 69%).

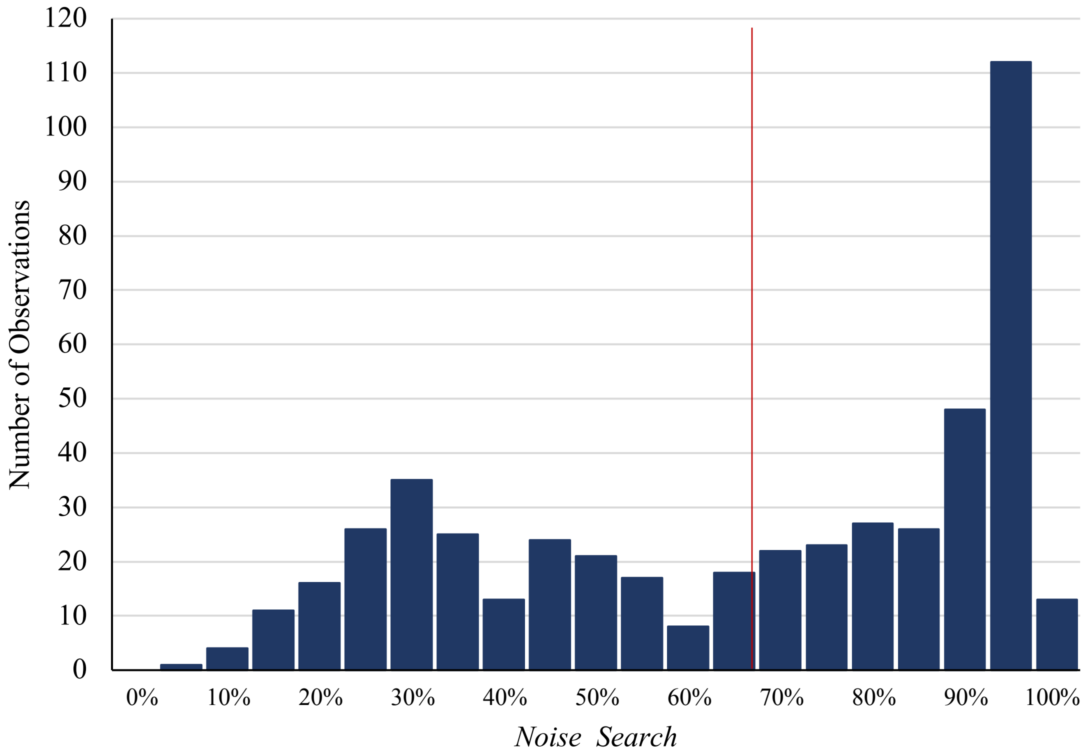
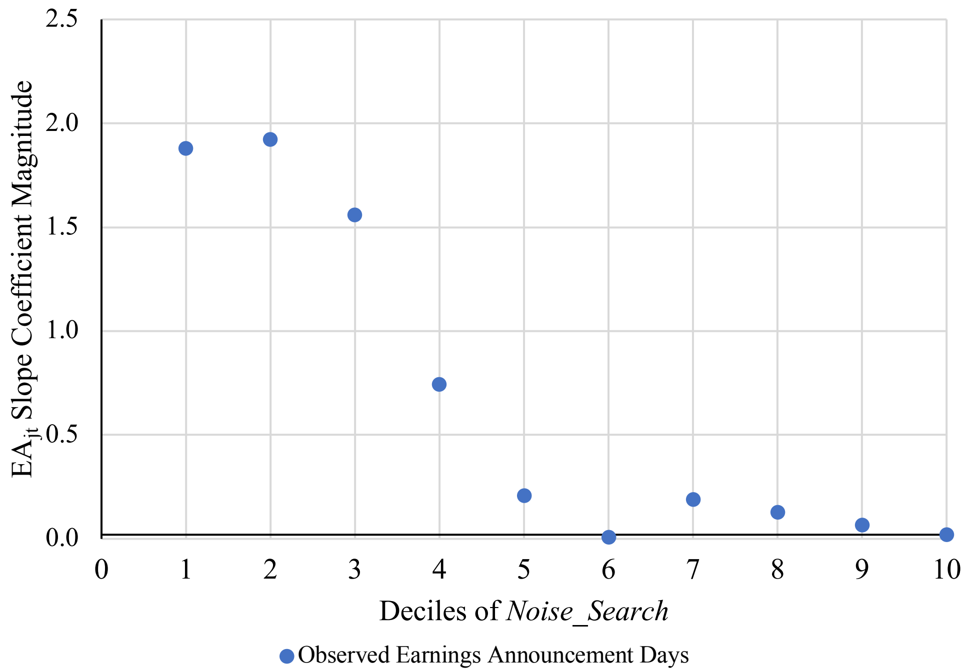


Figure 2 – Abnormal Google Search

Panel A presents abnormal google search (ASVI) on earnings announcement days, i.e. the estimated coefficient of EA from Equation (8) by decile of *Noise_Search*. The fitted values are plotted using quadratic prediction. For more details, please refer to Table 6. Panel B presents induced abnormal search (50%, 100%, and 200%) on random "earnings announcement" days using Equation (8) (Monte Carlo simulations) by decile of *Noise_Search*. For more details, please refer to Table 7. Again, the fitted values are plotted using quadratic prediction.

Panel A: Abnormal Google Search (ASVI) on earnings announcement (EA) days by deciles of Noise_Search



Panel B: Induced Abnormal Google Search on random days by deciles of Noise_Search

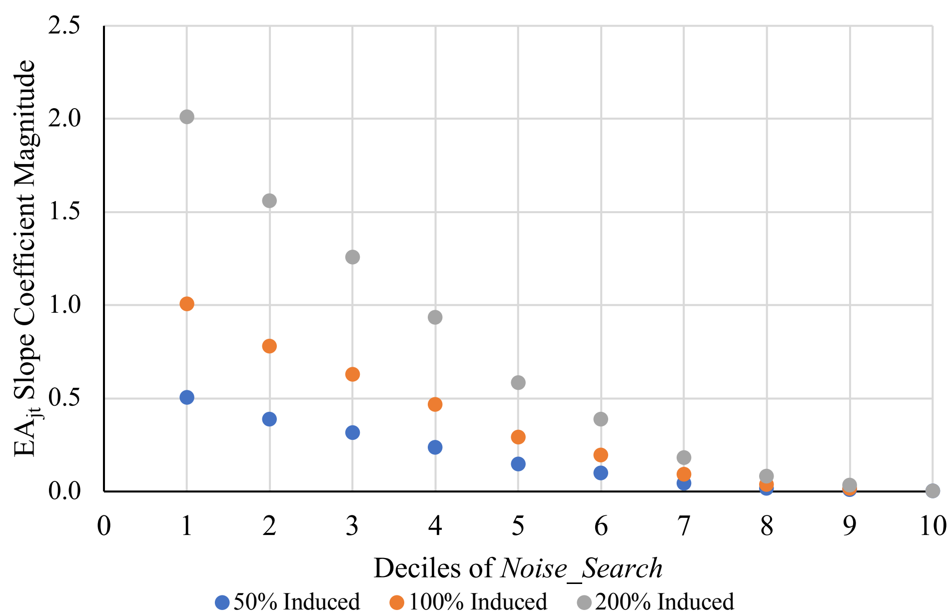


Figure 3 – *Noise_Search* by Quintile of Firm Characteristics

Figure 3 presents *Noise_Search* by quintiles (low to high) of the following key firm characteristics: market value of equity (MVE), book-to-market ratio (BTM), analyst following, and return volatility. Variable definitions are provided in Appendix A.

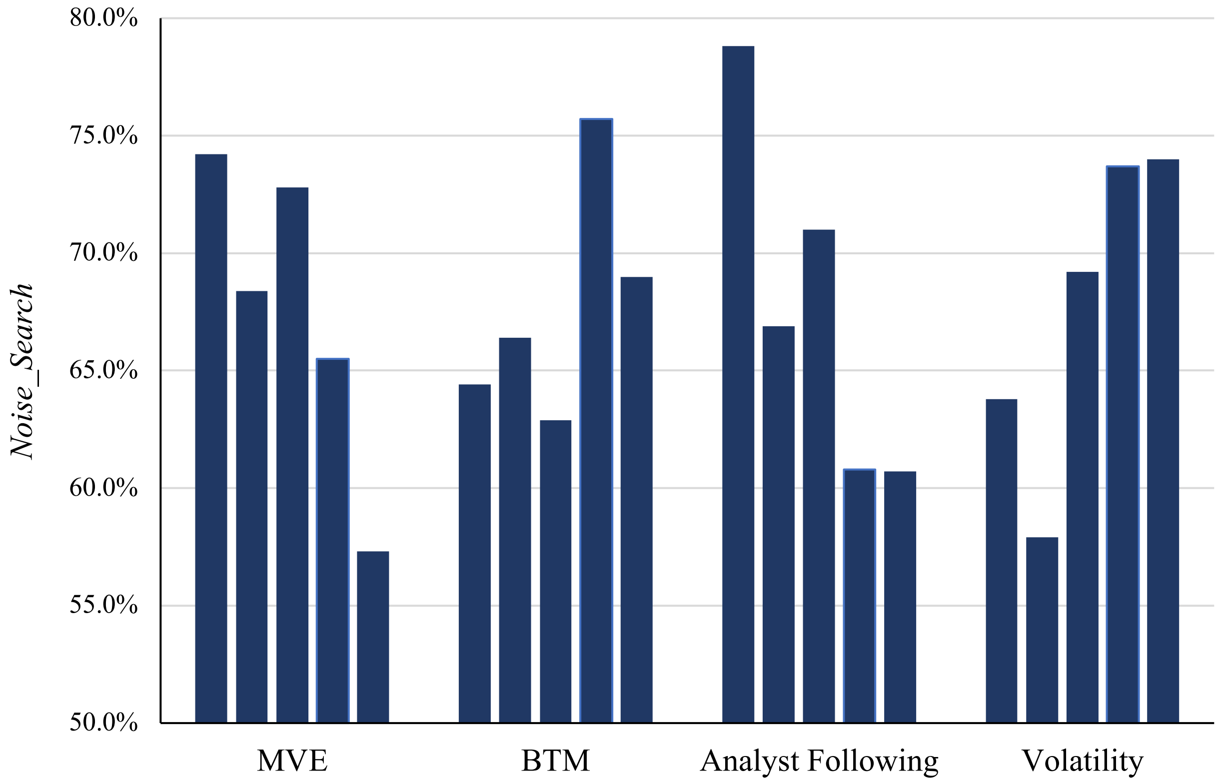


Table 1 – Simulation Analyses: When Measurement Error is Uncorrelated with X

This table presents simulation analyses of univariate regressions of performance, proxied by *ROA*, on firm age as follows: $ROA = \alpha + \beta_1 log(Age) + \varepsilon$.

These analyses use a sample of CRSP/Compustat firm-quarters from 2000 through 2018. *ROA* is defined as [operating income after depreciation scaled by assets]×100. *Age* is defined as the difference between the fiscal quarter-end and the first date the firm appears in Compustat. *Tobin's Q* is defined as market value of common equity plus long-term debt, scaled by assets. *Investments* is defined as [capital expenditures plus acquisitions plus R&D expense less capital sales, scaled by assets] × 100 (numerator variables are set to zero if missing). We drop observations without necessary data for *ROA*, *Q*, and *investments*, with market value under $10 million, or with price under $2. Observations are truncated at the 1st and 99th percentiles of *ROA*, *Q*, *investments*, and total assets. Regression standard errors are clustered by firm. Panel A uses unlogged *ROA* as the dependent variable, Panel B uses *Log(ROA)* as the dependent variable, and Panel C uses *Log(ROA+1)* as the dependent variable. Column (i) of each panel represents the simple regressions without introducing any ME, *v*, into the dependent variables. In columns (ii) through (iv), we utilize the variable, *ROA_ME*, where nonadditive ME of *v* is introduced into assets, the denominator of *ROA*, as follows: $A\_ME = [A(1-v)]$. In column (ii), we randomly assign ME, in *Assets* ranging from 0% to 9%. We repeat this process 1,000 times and column (ii) presents the average β_1 and t-statistic. The extent of measurement error is amplified in column (iii) with a randomly assigned *v* between [0%,45%], and in column (iv) with a randomly assigned *v* between [0%,90%].

Panel A: ROA simulations

	ROA	***ROA_ME***	***ROA_ME***	***ROA_ME***
Measurement Error (*v*)	n/a	[0%,9%]	[0%,45%]	[0%,90%]
Avg. Beta	1.11	1.17	1.48	2.85
Avg. T	[39.351]***	[39.349]***	[39.284]***	[38.237]***
Avg N	361,926	361,926	361,926	361,926

Panel B: Log(ROA) simulations

	ROA	***ROA_ME***	***ROA_ME***	***ROA_ME***
Measurement Error (*v*)	n/a	[0%,9%]	[0%,45%]	[0%,90%]
Avg. Beta	9.94	9.94	9.94	9.94
Avg. T	[11.857]***	[11.857]***	[11.843]***	[11.709]***
Avg N	279,861	279,861	279,861	279,861

Panel C: Log(ROA+1) simulations

	ROA	***ROA_ME***	***ROA_ME***	***ROA_ME***
Measurement Error (*v*)	n/a	[0%,9%]	[0%,45%]	[0%,90%]
Avg. Beta	1.17	1.23	1.60	3.17
Avg. T	[39.579]***	[39.583]***	[39.514]***	[38.159]***
Avg N	361,926	361,926	361,926	360,984

Table 2 – Simulation Analyses: When Measurement Error is Correlated with X

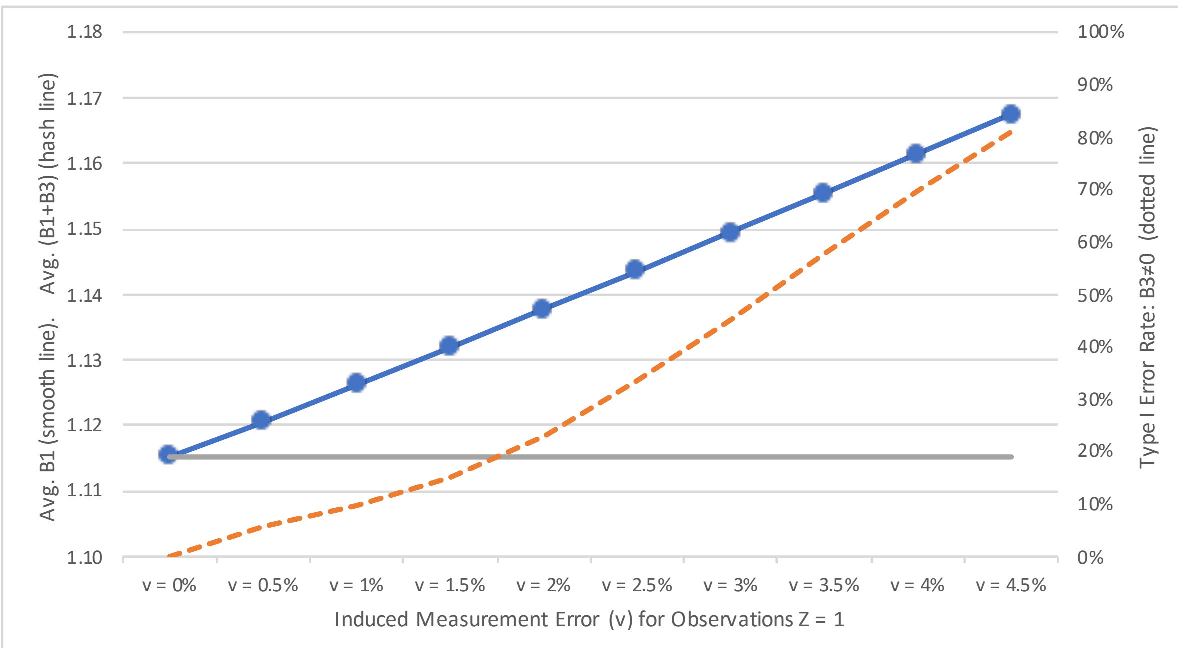
This table presents simulation analyses and figures of simulations using the following model:

$$ROA\_ME = \alpha + \beta_1 log(Age) + \beta_2 Z + \beta_3 log(Age) \times Z + \varepsilon$$

The sample and variables are described in the header to Table 1. *ROA_ME* is *ROA* measured with error *v* which is introduced into assets as follows: [E/(A(1-v))]. *Z* is a binary variable randomly assigned to observations. Observations with *Z=0* are assigned *v=0*. Observations with *Z=1* are assigned *v=v*, where v varies from 0 to 4.5% of assets. We repeat this process 1,000 times and report average β_1, average (β_1+β_3), and the number of trials rejecting the null that (β_3 - β_1) ≠ 0 at a 5% level of confidence. This process is then repeated separately for ten levels of *v* shown at the bottom of each figure.

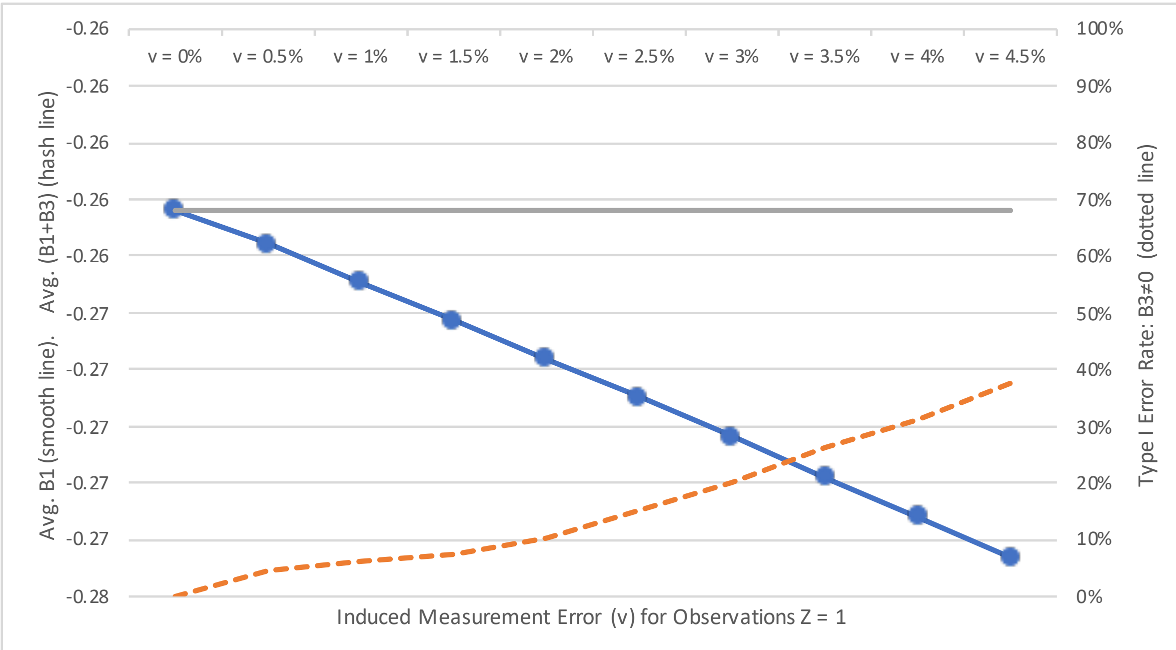
Panel A: ROA simulation

Smooth line = Average β_1 across 1,000 trials. Hashed line = Average (β_1+β_3). Dotted line = the number of trials rejecting the null that (β_3 - β_1) ≠ 0 at a 5% level of confidence.



Panel A: Tobin's Q simulation

Smooth line = Average β_1 across 1,000 trials. Hashed line = Average (β_1+β_3). Dotted line = the number of trials rejecting the null that (β_3 - β_1) ≠ 0 at a 5% level of confidence.



Panel C: Investments simulation

Smooth line = Average β_1 across 1,000 trials. Hashed line = Average (β_1+β_3). Dotted line = the number of trials rejecting the null that (β_3 - β_1) ≠ 0 at a 5% level of confidence.

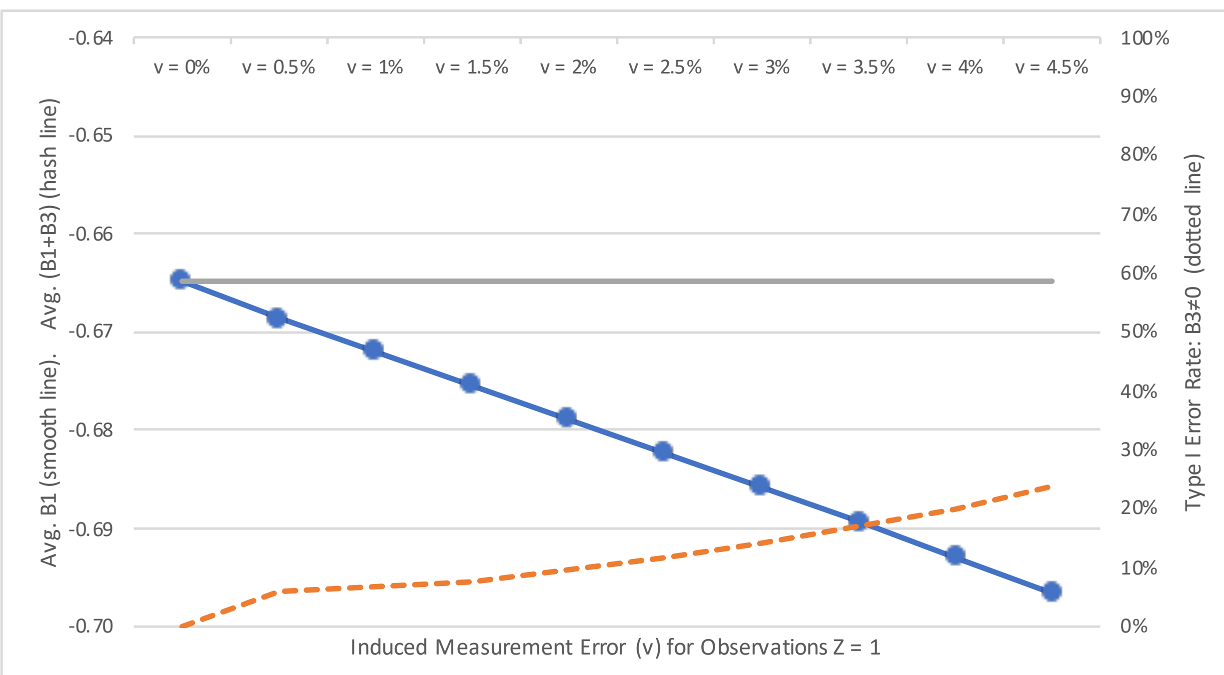


Table 3 – Google Search Sample Details

Panel A details our sample selection process. [A] We obtained the S&P 500 list of firms as of January 2016 consisting of 500 firms. In total 11 firms have two corresponding ticker symbols: Brown-Forman (BFA, BFB), Berkshire Hathaway (BRKA, BRKB), CBS Corp. (CBS, CBSA), Discovery Inc. (DISCA, DISCK), Twenty-First Century Fox (FOX, FOXA), Alphabet Inc (GOOG, GOOGL), Lennar Corp. (LEN, LENB), McCormick & Co. (MKC, MKCV), Constellation Brands (STZ, STZB), and Molson Coors Brewing (TAP, TAPA). We include both tickers for these firms. The dataset covers 2016 and 2017 trading days, totalling 501 days. [B] For two tickers the proprietary data vendor does not provide search volume data (STZB and MKCV). We verified this with Google SVI that also does not provide SVI data due to very limited search. [C] In total 19 tickers have change in ticker symbol during our sample period due to either a change in firm name (COH, DLPH, TSO, and YHOO) or merger (BHI, DD, DOW, EMC, HAR, HOT, LVTL, MJN, RAI, SPLS, STJ, SE, LLTC, TYC, and WFM). [D] For tickers SPGI and FTV the data is available in CRSP/Compustat/IBES from April 2016 and July 2016, respectively. Panel B presents descriptive statistics for information search and control variables per ticker trading day. Variable definitions are provided in Appendix A.

Panel A: Sample selection details

	Firms	Tickers	Trading Days
[A] Initial Sample of S&P 500 firms as of January 2016	500	511	256,011
[B] Less: firms/tickers without Google SVI data available	0	2	1,002
[C] Less: firms/tickers with a change in ticker symbol	19	19	9,519
[D] Less: missing observations in CRSP / Compustat / IBES	0	0	475
Final Sample	481	490	245,015

Panel B: Sample summary statistics

	N	Mean	Std.Dev.	p25	Median	p75
SVI	245,015	33.484	23.813	12.857	30.186	51.330
ASVI	245,015	0.107	1.928	-0.254	-0.028	0.200
ASVI2	245,015	0.027	0.677	-0.194	0.016	0.268
FISVI	245,015	10.819	16.413	0.000	0.000	17.617
AFISVI	245,015	0.419	35.303	-1.000	-0.176	0.227
AFISVI2	245,015	0.045	1.221	-0.646	0.000	0.558
EA	245,015	0.016	0.125	0.000	0.000	0.000
News Articles	245,015	1.647	2.628	0.000	1.000	2.000
Abs Return	245,015	0.100	0.100	0.003	0.007	0.014
Spread	245,015	0.019	0.011	0.011	0.016	0.023
Total EAs	245,015	5.472	2.884	3.000	5.000	8.000
MVE	245,015	5.535	2.877	3.000	6.000	8.000
Analysts	245,015	2.841	0.481	2.639	2.908	3.164
BTM	245,015	5.459	2.864	3.000	5.000	8.000
Volume	245,015	1.898	1.102	1.186	1.595	2.291
Institutional Ownership	245,015	0.838	0.1511	0.758	0.856	0.938
Fourth Qtr	245,015	0.248	0.431	0.000	0.000	0.000
Leverage	245,015	0.657	0.208	0.529	0.651	0.794
Momentum	245,015	0.050	0.019	0.037	0.045	0.059
Return on Assets	245,015	0.013	0.025	0.005	0.013	0.022
Stock Volatility	245,015	0.014	0.005	0.010	0.013	0.016
Beta	245,015	0.936	0.654	0.403	0.786	1.332
Investor_Search	245,015	0.311	0.287	0.039	0.218	0.567
Noise_Search	245,015	0.689	0.287	0.433	0.782	0.961

Table 4 – Google Search Click-Through Website Categorization

Panel A details the types of websites visited after Google ticker searches. Column (i) presents the average total click-throughs per month, pooled across all firms. Column (iiii) is the portion of click-throughs that are to a website included in the audit procedure detailed in Section 3.3. In the pooled sample, after typing any of the ticker symbols on Google, individuals clicked on 63,263 different websites. In total we audited 4,460 websites covering 94% of all clicks. Column (iv) is the fraction of the audited traffic that is determined to be "investor-related." Panel B lists the top 20 websites that are identified as investor-related. The data is based on the final sample of 490 tickers.

Panel A: Categories of websites visited

Website Category	(i) Avg. Total Clicks per Month	(ii) Percentage of All Traffic	(iii) Percentage of Traffic Audited	(iv) Fraction Investor-Related
Adult	168,318	0.1%	72.3%	0.0%
Arts_and_Entertainment	3,892,503	5.7%	92.2%	0.2%
Autos_and_Vehicles	746,802	0.6%	84.5%	1.1%
Beauty_and_Fitness	3,158,182	2.5%	99.6%	0.0%
Books_and_Literature	24,719	0.0%	71.1%	6.1%
Business_and_Industry	6,656,695	5.0%	91.9%	2.3%
Career_and_Education	598,555	0.5%	73.9%	0.7%
Computer_and_Electronics	1,645,274	1.3%	89.6%	0.2%
Finance	7,800,414	9.6%	98.6%	64.8%
Food_and_Drink	351,913	0.3%	72.9%	0.0%
Gambling	40,726	0.0%	83.0%	0.0%
Games	558,790	0.4%	79.4%	0.0%
Health	4,359,264	3.4%	97.0%	0.0%
Home_and_Garden	40,326	0.0%	76.4%	0.0%
Internet_and_Telecom	9,532,860	7.5%	97.6%	0.2%
Law_and_Government	367,407	0.3%	83.1%	1.2%
News_and_Media	6,964,061	7.7%	92.8%	56.0%
People_and_Society	241,284	0.2%	70.3%	0.0%
Pets_and_Animals	163,674	0.1%	78.0%	0.0%
Recreation_and_Hobbies	310,995	0.3%	82.4%	0.0%
Reference	861,556	0.9%	96.2%	8.7%
Science	149,079	0.1%	79.4%	0.0%
Shopping	43,539,718	35.3%	98.5%	0.2%
Sports	186,356	0.2%	73.0%	0.5%
Travel	1,115,030	0.9%	96.4%	4.5%
Unknown	17,395,398	17.0%	84.0%	12.4%
All categories together	**110,869,899**	**100.0%**	**94.0%**	**31.0%**

Panel B: Top 20 investor-related websites

	URL	**Percentage of Investor Traffic**
1	finance.yahoo.com	28.3%
2	seekingalpha.com	9.0%
3	fool.com	6.4%
4	stocktwits.com	4.9%
5	marketwatch.com	4.9%
6	cnbc.com	3.6%
7	investorplace.com	3.0%
8	thestreet.com	2.8%
9	nasdaq.com	2.8%
10	businessinsider.com	2.3%
11	money.cnn.com	1.9%
12	Bloomberg.com	1.8%
13	invest.ameritrade.com	1.5%
14	stockcharts.com	1.2%
15	investors.com	1.2%
16	barrons.com	1.0%
17	streetinsider.com	0.9%
18	stocknewsjournal.com	0.9%
19	us.etrade.com	0.9%
20	forbes.com	0.8%
21+	All others	19.9%
Total		**100.0%**

Table 5 – Variation in *Investor_Search* across Firms

Panel A details the average clicks per firm-month by ticker type, as well as the average percentage of clicks that are investor-related. Ticker type designations as "Ambiguous" used and obtained from Drake et al (2011): AA, ABC, ALL, AN, CAT, COST, EBAY, ED, FAST, HAS, HD, HOG, KEY, KO, LOW, MAT, MET, PEG, SEE, TAP. Panel B details the average clicks per firm month by Fama-French 12 industry classification. Panel C provides a breakdown of the average *Noise_Search* by quintiles of firm characteristics. *,**,*** indicates statistical significance at the p < 0.10, 0.05, 0.01 level, respectively.

Panel A: Click-throughs by ticker type

Ticker Type	Tickers	Average Ticker Searches Per Month	Average *Investor_Search*
Ambiguous	20	2,438,237	15.1%
Other One-Letter Tickers	11	1,332,255	6.6%
Other Two-Letter Tickers	50	290,641	14.4%
Other Three-Letter Tickers	298	88,688	27.7%
Other Four-Letter Tickers	101	88,484	53.0%
Other Five-Letter Tickers	4	144,602	61.3%
All tickers	**490**	**232,761**	**31.0%**

Panel B: Click-throughs by firm industry

Firm's Industry (FF 12)	Tickers	Average Ticker Searches Per Month	Average *Investor_Search*
Consumer NonDurables	34	82,244	22.4%
Consumer Durables	9	643,897	19.4%
Manufacturing	41	84,843	26.6%
Oil, Gas, and Coal Extraction and Products	27	67,634	30.2%
Chemicals and Allied Products	17	52,418	27.5%
Business Equipment	74	807,153	38.8%
Telephone and Television Transmission	16	251,419	49.1%
Utilities	32	86,130	20.1%
Wholesale, Retail, and Some Services	50	232,515	35.3%
Healthcare, Medical Equipment, and Drugs	39	31,824	49.4%
Finance	99	135,932	24.6%
Other	52	121,665	28.7%
All tickers	**490**	**232,761**	**31.0%**

Panel C: Average Investor_Search by firm characteristic

Firm Characteristic	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Diff Q5-Q1
MVE	57.3%	65.5%	72.8%	68.4%	74.2%	16.9%***
BTM	69.0%	75.7%	62.9%	66.4%	64.4%	-4.6%*
Leverage	71.2%	69.1%	63.7%	70.9%	63.5%	-7.6%**
ROA	61.6%	64.8%	69.7%	74.8%	67.5%	5.8%**
Inst. Own.	71.1%	70.8%	66.2%	65.7%	64.7%	-6.4%**
Analyst Following	60.7%	60.8%	71.0%	66.9%	78.8%	18.1%***
Momentum	57.8%	65.6%	70.2%	69.7%	75.0%	17.2%***
Stock Volatility	61.8%	63.3%	70.0%	70.0%	73.0%	11.2%***
Trading Volume	63.9%	66.0%	68.9%	67.2%	72.4%	8.5%***
Beta	74.0%	73.7%	69.2%	57.9%	63.8%	-10.2%***
Spread	67.0%	68.8%	69.2%	69.6%	70.2%	3.2%***

Table 6 – Regressions of Google Search on Earnings Announcement Days

This table presents the results of Equation (8). The dependent variable is *SVI, ASVI,* or *ASVI2*. Panel A (B) tabulates results excluding (including untabulated) control variables: *News Articles, Abs Return, Spread, Total EAs, MVE, Analysts, BTM, Inst Own*, *Fourth Quarter* and *Week* fixed effects. Variable definitions are provided in Appendix A. Panel C excludes "ambiguous" tickers listed in the header of Table 5. T-statistics are in parentheses. Standard errors are clustered by firm. *,**,*** indicates statistical significance at the p < 0.10, 0.05, 0.01 level, respectively.

Panel A: Without controls or fixed effects

	Pooled	**By Decile of Noise_Search**									
		1 [Low]	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 [High]**
Observations	245,015	24,970	24,048	24,549	24,548	25,050	24,283	24,549	24,048	24,922	24,048
Average *Noise_Search*	0.689	0.177	0.311	0.432	0.574	0.729	0.827	0.912	0.960	0.985	0.997
EA [0] - SVI	11.430***	20.050***	24.920***	23.570***	16.000***	7.705***	8.687***	6.431***	4.229***	2.203*	0.424
	(29.82)	(26.29)	(28.50)	(25.13)	(14.49)	(6.64)	(7.34)	(5.61)	(3.55)	(1.88)	(0.37)
Adjusted R-squared	0.004	0.027	0.033	0.025	0.008	0.002	0.002	0.001	0.001	0.000	0.000
EA [0] - ASVI	0.674***	2.764***	2.461***	2.097***	1.115***	0.404***	0.292	0.282***	0.104	0.095***	-0.008
	(60.99)	(28.28)	(31.87)	(27.50)	(16.95)	(5.42)	(1.50)	(10.25)	(0.65)	(3.61)	(-0.35)
Adjusted R-squared	0.015	0.031	0.041	0.030	0.012	0.001	0.000	0.004	0.000	0.001	0.000
EA [0] – ASVI2	0.462***	1.099***	1.052***	0.948***	0.590***	0.299***	0.256***	0.178***	0.114***	0.077***	-0.003
	(42.49)	(21.96)	(23.61)	(23.96)	(14.95)	(8.62)	(8.17)	(6.76)	(4.94)	(3.73)	(-0.21)
Adjusted R-squared	0.007	0.019	0.023	0.023	0.009	0.003	0.003	0.002	0.001	0.001	0.000

Panel B: With controls and year-week fixed effects

	Pooled	**By Decile of Noise_Search**									
		1 [Low]	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 [High]**
Observations	245,015	24,970	24,048	24,549	24,548	25,050	24,283	24,549	24,048	24,922	24,048
Average *Noise_Search*	0.689	0.177	0.311	0.432	0.574	0.729	0.827	0.912	0.960	0.985	0.997
EA [0] - SVI	7.879***	12.340***	19.330***	14.59***	11.96***	4.801	4.834**	-3.350	-0.221	5.029	4.470**
	(5.17)	(6.60)	(7.73)	(3.98)	(5.31)	(1.60)	(2.09)	(-0.99)	(-0.06)	(1.27)	(2.07)
Adjusted R-squared	0.030	0.132	0.111	0.128	0.064	0.135	0.126	0.105	0.240	0.142	0.147
EA [0] - ASVI	0.470***	1.858***	1.901***	1.539***	0.723***	0.186	-0.012	0.168*	0.106	0.046**	0.001
	(13.08)	(6.19)	(5.98)	(5.28)	(4.76)	(1.36)	(-0.06)	(1.86)	(1.58)	(2.03)	(0.04)
Adjusted R-squared	0.037	0.077	0.084	0.065	0.043	0.015	0.018	0.029	0.008	0.027	0.009
EA [0] – ASVI2	0.297***	0.702***	0.759***	0.672***	0.379***	0.199***	0.147***	0.073*	0.086***	0.0367*	0.002
	(13.50)	(9.32)	(9.31)	(8.02)	(6.41)	(3.77)	(3.74)	(1.68)	(2.69)	(1.87)	(0.12)
Adjusted R-squared	0.025	0.063	0.066	0.063	0.044	0.017	0.024	0.021	0.021	0.032	0.016

Panel C: With controls and fixed effects, and dropping "ambiguous" tickers listed in Table 3

	Pooled	By Decile of Noise_Search									
		1 [Low]	2	3	4	5	6	7	8	9	10 [High]
Observations	234,995	24,469	24,048	24,549	23,546	24,048	23,281	23,547	22,545	22,918	22,044
Average *Noise_Search*	0.683	0.176	0.311	0.432	0.574	0.731	0.828	0.912	0.961	0.985	0.997
EA [0] - SVI	7.772***	12.400***	19.330***	14.590***	11.800***	4.257	5.483**	-3.190	0.262	4.074	4.894**
	(5.02)	(6.49)	(7.73)	(3.99)	(5.36)	(1.45)	(2.36)	(-0.91)	(0.07)	(1.02)	(2.15)
Adjusted R-squared	0.034	0.142	0.111	0.128	0.101	0.124	0.183	0.098	0.250	0.133	0.146
EA [0] - ASVI	0.492***	1.855***	1.901***	1.539***	0.757***	0.203	-0.023	0.181*	0.123*	0.052**	-0.001
	(13.23)	(6.05)	(5.98)	(5.28)	(4.84)	(1.43)	(-0.11)	(1.92)	(1.81)	(2.10)	(-0.09)
Adjusted R-squared	0.039	0.080	0.084	0.065	0.044	0.016	0.019	0.029	0.008	0.028	0.009
EA [0] – ASVI2	0.311***	0.706***	0.759***	0.672***	0.396***	0.210***	0.151***	0.079*	0.095***	0.039*	0.001
	(13.69)	(9.18)	(9.31)	(8.02)	(6.53)	(3.84)	(3.66)	(1.76)	(2.84)	(1.83)	(0.03)
Adjusted R-squared	0.026	0.065	0.066	0.063	0.045	0.018	0.024	0.022	0.020	0.034	0.018

Panel D: With controls and fixed effects, and dropping "ambiguous" tickers listed in Table 3 and one-letter and two-letter tickers

	Pooled	By Decile of Noise_Search									
		1 [Low]	2	3	4	5	6	7	8	9	10 [High]
Observations	205,702	24,469	23,547	23,547	22,544	19,539	19,539	19,539	18,036	17,908	17,034
Average *Noise_Search*	0.655	0.176	0.311	0.431	0.572	0.730	0.830	0.908	0.959	0.985	0.997
EA [0] - SVI	9.958***	12.400***	19.060***	17.500***	11.850***	4.863***	5.933**	0.150	7.331***	0.997	3.122
	(8.91)	(6.49)	(7.59)	(5.33)	(5.21)	(3.22)	(2.29)	(0.05)	(2.87)	(0.47)	(1.54)
Adjusted R-squared	0.026	0.142	0.123	0.101	0.101	0.063	0.132	0.071	0.188	0.085	0.193
EA [0] - ASVI	0.547***	1.855***	1.946***	1.547***	0.785***	0.251	-0.038	0.198*	0.108	0.057*	-0.002
	(13.20)	(6.05)	(6.04)	(4.95)	(4.85)	(1.45)	(-0.15)	(1.78)	(1.06)	(1.76)	(-0.11)
Adjusted R-squared	0.044	0.080	0.086	0.067	0.046	0.018	0.022	0.032	0.010	0.034	0.011
EA [0] – ASVI2	0.345***	0.706***	0.775***	0.673***	0.408***	0.257***	0.177***	0.081	0.118***	0.045	0.001
	(13.69)	(9.18)	(9.44)	(7.45)	(6.57)	(3.87)	(3.66)	(1.58)	(2.75)	(1.63)	(0.03)
Adjusted R-squared	0.030	0.065	0.067	0.065	0.047	0.021	0.029	0.025	0.025	0.040	0.020

Table 7 – *Noise_Search* and Induced Increase in Abnormal Google Search [ASVI] on Random Days

This table presents the Monte Carlo results of Equation (8) for random EA days. The dependent variable is *ASVI* given a fixed inducement of search. Controls and fixed effects are untabulated. Standard errors are clustered by firm. For each level of induced search, the upper row presents the average coefficient estimate across 1,000 trials, and the bottom row presents the number of trials that rejected the null of no change in SVI. See Section 5.2 for further details.

		(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)	(xi)
		Pooled	Decile Partitions on Noise_Search									
Induced Increase			1 [Low]	2	3	4	5	6	7	8	9	10 [High]
0%	Avg. coefficient	0.000	0.004	-0.004	0.000	0.002	0.001	0.003	0.002	-0.005	0.001	0.001
	rejected at 5%	0.1%	0.8%	0.3%	0.3%	0.4%	0.6%	0.1%	2.0%	0.1%	0.9%	0.8%
5%	Avg. coefficient	0.018	0.054	0.035	0.031	0.025	0.015	0.012	0.006	-0.003	0.002	0.001
	rejected at 5%	2.8%	3.8%	3.8%	2.3%	2.4%	1.4%	0.1%	3.1%	0.1%	1.0%	0.7%
10%	Avg. coefficient	0.035	0.104	0.074	0.063	0.048	0.03	0.022	0.011	-0.001	0.003	0.001
	rejected at 5%	21.5%	14.0%	13.9%	8.0%	7.8%	3.3%	0.1%	4.8%	0.2%	1.0%	0.7%
15%	Avg. coefficient	0.053	0.154	0.113	0.094	0.072	0.044	0.032	0.015	0.001	0.003	0.001
	rejected at 5.0%	57.9%	29.4%	34.2%	21.2%	17.9%	7.9%	0.1%	6.7%	0.3%	1.4%	0.7%
20%	Avg. coefficient	0.070	0.200	0.154	0.124	0.094	0.059	0.036	0.020	0.002	0.005	0.001
	rejected at 5.0%	86.7%	48.1%	57.6%	38.9%	33.3%	14.2%	0.2%	9.1%	0.5%	1.4%	0.6%
25%	Avg. coefficient	0.088	0.254	0.192	0.157	0.118	0.074	0.051	0.025	0.006	0.005	0.001
	rejected at 5.0%	96.5%	69.6%	78.1%	61.1%	50.6%	22.8%	0.6%	11.9%	0.3%	2.0%	1.1%
50%	Avg. coefficient	0.176	0.505	0.387	0.314	0.235	0.146	0.099	0.042	0.016	0.009	0.002
	rejected at 5.0%	99.9%	99.9%	99.5%	98.4%	96.%	78.3%	6.0%	38.1%	0.9%	2.6%	1.2%
100%	Avg. coefficient	0.353	1.007	0.778	0.629	0.467	0.292	0.194	0.092	0.038	0.017	0.003
	rejected at 5.0%	100.0%	100.0%	100.0%	100.0%	99.9%	98.2%	43.4%	96.3%	3.4%	7.1%	1.7%
200%	Avg. coefficient	0.705	2.011	1.559	1.258	0.933	0.584	0.386	0.182	0.081	0.032	0.003
	rejected at 5.0%	100.0%	100.0%	100.0%	100.0%	100.0%	99.6%	92.3%	100.0%	15.2%	21.8%	3.0%
300%	Avg. coefficient	1.057	3.014	2.340	1.888	1.399	0.875	0.577	0.272	0.124	0.048	0.009
	rejected at 5.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	95.2%	100.0%	32.5%	44.9%	4.5%
500%	Avg. coefficient	1.762	5.021	3.903	3.146	2.330	1.458	0.960	0.452	0.210	0.079	0.014
	rejected at 5.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	95.3%	100.0%	79.8%	87.4%	8.3%

Table 8 – Cross-Sectional Tests

This table presents the results of Equation (9) with high (quartile four) – low (quartiles one to three) partitions on firm size (*MVE*), Analyst following (*Analysts*) and bid-ask spread *(Spread)*. The dependent variable is *SVI, ASVI,* or *ASVI2* and untabulated controls: *News Articles, Abs Return, Spread, Total EAs, BTM, Inst Own*, *Fourth Quarter* and *Week* fixed effects. Panel A tabulates cross-sectional partitions on earnings announcement (EA) days. Panel B tabulates the Monte Carlo results of Equation (9) for random EA days. Standard errors are clustered by firm. *,**,*** indicates statistical significance at the $p < 0.10$, 0.05, 0.01 level, respectively.

Panel A: Cross-sectional partitions of Google Search on earnings announcement days

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
EA	0.408***	1.943***	0.350***	1.845***	0.460***	1.982***
	(10.16)	(16.32)	(9.57)	(16.44)	(11.31)	(16.71)
Large_Firms	0.348***	0.119				
	(3.21)	(1.40)				
*EA * Large_Firms*	0.258**	-0.016				
	(2.06)	(-0.15)				
High_Following			-0.007	-0.011		
			(-0.96)	(-1.54)		
*EA * High_Following*			0.559***	0.401***		
			(5.28)	(5.01)		
Large_Spread					0.001	0.001
					(0.09)	(0.01)
*EA * Large_Spread*					0.118*	0.043
					(1.75)	(0.81)
Noise_Search		-0.087***		-0.083***		-0.081***
		(-9.69)		(-9.07)		(-8.84)
*EA * Noise_Search*		-2.139***		-2.098***		-2.165***
		(-15.70)		(-16.12)		(-15.86)
Controls	Included	Included	Included	Included	Included	Included
Observations	245,015	245,015	245,015	245,015	245,015	245,015
Adjusted R-squared	0.023	0.038	0.025	0.039	0.023	0.039

Panel B: Cross-sectional simulation tests

Induced Investor_Search of:	25%	50%	100%	200%	500%
*Random_Day * Large Firms*					
Average Coefficient	0.025	0.050	0.099	0.194	0.474
Interactions rejected at 5% level	0.0%	28.0%	58.0%	74.0%	86.0%
*Random_Day * High_Following*					
Average Coefficient	0.015	0.033	0.070	0.142	0.353
Interactions rejected at 5% level	0.0%	6.0%	14.0%	32.0%	54.0%
*Random_Day * Large_Spread*					
Average Coefficient	0.021	0.042	0.084	0.168	0.418
Interactions rejected at 5% level	0.0%	6.0%	22.0%	44.0%	68.0%
Controls	Included	Included	Included	Included	Included
Observations	241,105	241,105	241,105	241,105	241,105
Adjusted R-squared	0.018	0.015	0.026	0.039	0.023

Table 9 – Comparison of AIA and SVI as in Ben-Rephael et al. (2017)

Panel A uses Bloomberg AIA as the dependent variable. The model is a probit without control variables, similar to that in BRDI (2017) except that we focus solely on earnings announcements. Panel Buses the same model as in Panel A but it uses a binary version of SVI that mimics AIA, labeled DADSVI as in BRDI (2017). Column (i) presents pooled analyses, and Columns (ii) through (xi) repeat (i) for each decile of *Noise_Search*. The sample is the same as that used in previous tables except due to reductions where AIA is unavailable. We calculate the 95% confidence interval of the pseudo R-squared by bootstrapping the sample with 500 trials. In Panel C, we calculate the Vuong Z-statistic to compare differences in adjusted r-squared by performing a reverse OLS regression with RDQ as the dependent variable and AIA and DADSVI as the respective independent variables.

	(i)	**(ii)**	**(iii)**	**(iv)**	**(v)**	**(vi)**	**(vii)**	**(viii)**	**(ix)**	**(x)**	**(xi)**
	Pooled	**Decile Partitions on Noise_Search**									
		1 [Low]	2	3	4	5	6	7	8	9	10 [High]
Observations	75,079	7,790	7,728	7,720	7,385	7,105	7,432	7,438	7.475	7,547	7,279
Average *Noise_Search*	0.684	0.176	0.311	0.433	0.576	0.728	0.825	0.913	0.960	0.985	0.997
Panel A: AIA regressions											
EA[0] Coefficient	2.754***	2.710***	2.559***	2.362***	2.736***	2.816***	2.865***	3.005***	2.623***	2.902***	3.530***
Z-stat	(35.80)	(11.53)	(14.17)	(10.17)	(9.97)	(10.84)	(13.58)	(9.77)	(12.31)	(12.08)	(9.96)
Pseudo R-Squared	0.0599	0.0551	0.0467	0.0468	0.0510	0.0604	0.0627	0.0651	0.0705	0.0758	0.0821
Panel B: DADSVI regressions											
EA[0] Coefficient	1.025***	1.637***	1.811***	1.839***	1.1051***	0.721***	0.816***	0.302*	0.373**	0.421***	-0.381
Z-stat	(20.55)	(11.12)	(12.43)	(12.37)	(7.72)	(4.93)	(5.14)	(1.70)	(2.21)	(2.84)	(-0.20)
Pseudo R-squared	0.0234	0.0683	0.0851	0.0881	0.0280	0.0108	0.0139	0.0013	0.0020	0.0027	0.0000
Panel C: Vuong Test of R-Squared											
OLS R-sq. Diff: AIA - DADSVI	0.061***	0.005	-0.026*	-0.028*	0.043***	0.071***	0.718***	0.083***	0.091***	0.096***	0.104***
Vuong P-value	0.00	0.72	0.07	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 10 – *Regressions using Google Finance Investor SVI (FISVI)*

This table presents the repeats the analyses from Panel B of Table 6, but using dependent variables based on Google Finance-Investor SVI (FISVI) instead of total SVI. Dependent variables *FISVI*, *AFISVI*, and *AFISVI2* are constructed analogously to *SVI, ASVI,* and *ASVI2* in Table 6. Controls and fixed effects are included but untabulated. Standard errors are clustered by firm. Variable definitions are provided in Appendix A. *,**,*** indicates statistical significance at the p < 0.10, 0.05, 0.01 level, respectively.

Regressions including controls and week fixed effects, using Google Finance-Investor SVI. Comparable to Panel B of Table 6

	Pooled	**By Decile of Noise Search**									
		1 [Low]	2	3	4	5	6	7	8	9	10 [High]
EA [0] - FISVI	7.652***	8.353***	9.182***	12.620***	7.806***	6.403**	8.607***	4.861	7.063***	7.863***	1.709
	(9.96)	(6.68)	(5.29)	(7.48)	(4.95)	(2.49)	(5.69)	(1.50)	(3.95)	(4.55)	(1.36)
Adjusted R-squared	0.066	0.088	0.135	0.125	0.109	0.164	0.078	0.146	0.097	0.098	0.093
EA [0] - AFISVI	1.079***	1.076***	1.787***	1.868***	1.003***	0.878***	1.131***	1.200***	0.894***	0.814***	0.074
	(14.88)	(5.09)	(7.61)	(7.05)	(6.13)	(4.36)	(4.44)	(4.51)	(4.36)	(4.24)	(0.63)
Adjusted R-squared	0.024	0.044	0.046	0.043	0.043	0.027	0.036	0.025	0.021	0.019	0.013
EA [0] – AFISVI2	0.606***	0.640***	0.847***	0.997***	0.567***	0.502***	0.654***	0.621***	0.587***	0.513***	0.073
	(18.76)	(6.72)	(8.50)	(9.61)	(6.99)	(5.15)	(6.76)	(5.44)	(6.02)	(5.33)	(1.00)
Adjusted R-squared	0.024	0.042	0.042	0.041	0.036	0.026	0.031	0.026	0.023	0.021	0.017

Appendix A: Variable Definitions

All continuous variables are winsorized at 1% and 99%.

Variable	Description	Source
Google Ticker Search Variables:		
%InvClicks$_i$	Fraction of investor related click-throughs for firm *i*.	Proprietary
%NonInvClicks$_i$	Fraction of non-investor related click-throughs for firm *i*.	Proprietary
Noise_Search$_i$	Decile rank of *%NonInvClicks*.	Proprietary
Keyword Search$_{i,m}$	Monthly absolute Google ticker searches for firm *i* in month *m*.	Google AdWords
SVI$_{i,t}$	Google ticker search volume index for firm *i* on day *t*. Obtaining daily SVI over a two-year period requires a four-step process. First, we download SVI data for the window of 2004 through 2017. Google provides this data at the monthly level. Second, we download daily SVI for each month in 2016 and 2017. Google provides this data at the daily level. Third, we convert the daily data to a common scale by multiplying the daily data by the monthly SVI scaled by 100. Fourth, we rescale the daily data so that each firm has a maximum value of 100 during our sample period; i.e., we divide each daily value by the maximum value observed for firm *i* over the window of 2016 through 2017.	Google Trends
ASVI$_{i,t}$	SVI for firm *i* on day *t* less the average SVI for firm *i* on the same weekday over prior 10 weeks, scaled by the average SVI for firm *i* on the same weekday over the prior 10 weeks.	Google Trends
ASVI2$_{i,t}$	Natural log of 1 plus SVI for firm *i* on day *t* less the average of natural log of 1 plus SVI for firm *i* on the same weekday over prior 10 weeks.	Google Trends
FISVI$_{i,t}$	Google ticker search volume index sub-category "Finance/Investing" (category 107) for firm *i* on day *t*. Fixed scaling is employed as in SVI.	Google Trends
AFISVI$_{i,t}$	FISVI for firm *i* on day *t* less the average FISVI for firm *i* on the same weekday over the prior 10 weeks, scaled by the average FISVI for firm *i* on the same weekday over the prior 10 weeks.	Google Trends
AFISVI2$_{i,t}$	Natural log of 1 plus FISVI for firm *i* on day *t* less the average of natural log of 1 plus FISVI for firm *i* on the same weekday over prior 10 weeks.	Google Trends
DADSVI$_{i,t}$	Consistent with Ben-Rephael et al. 2017, we follow Bloomberg's methodology and we assign DSVI on day *t* one of the potential 0, 1, 2, 3, or 4 scores using the firm's past 30 trading day DSVI values. For example, if DSVI on day *t* is in the lowest 80% of past DSVI values, it is assigned a value of zero. DADSVI is one on day *t* if the score is 3 or 4, and zero otherwise (i.e. 0, 1, 2). To see how the Bloomberg measure is calculated see the variable AIA.	Google Trends
Events		
EA$_{i,t}$	An indicator variable set equal to one on day *t* if firm *i* announces earnings, and zero otherwise.	Compustat
Determinants of Google Ticker Search:		
News Articles$_{i,t}$	Daily number of news articles for firm *i* on day *t*.	FactSet
Abs Return$_{i,t}$	The absolute raw stock return for firm *i* on day *t*.	CRSP
Spread$_{i,t}$	Bid-ask spread for firm *i* on day *t*. Calculated as [(bid – ask) / price].	CRSP
Large_Spread	Indicator variable equal to one if the average bid-ask spread of the most recent fiscal quarter is in the highest quartile of the sample, and zero otherwise.	CRSP
Total EAs$_t$	The decile rank of the total number of firms announcing earnings on day *t*, calculated across all of Compustat.	Compustat
MVE$_{i,q}$	The decile rank of market capitalization of firm *i* as of most recent fiscal quarter-end *q* (PRCCQ x CSHOQ).	CRSP
Large_Firms$_{i,q}$	Indicator variable set equal to one if the market value of equity of the firm of the most recent fiscal quarter-end is in the highest quartile of the sample, and zero otherwise.	CRSP
Analysts$_{i,t}$	Natural log of 1 plus the number of analysts following firm *i* on day *t*.	I/B/E/S
High_Following$_{i,q}$	Indicator variable set equal to one if the average number of analyst following of the most recent fiscal quarter-end is in the highest quartile of the sample, and zero otherwise.	I/B/E/S
Volume$_{i,t}$	Daily share volume divided by shares outstanding for firm *i* on day *t*, averaged by month.	CRSP
BTM$_{i,q}$	The decile rank of the ratio of book value of equity to market capitalization for firm *i* as of the most recent fiscal quarter-end *q*. (CEQQ/[PRCCQ x CSHOQ]).	Compustat/CRSP

Variable	Description	Source
Inst. Own$_{i,q}$	Percentage institutional ownership in most recent quarter for firm *i*.	FactSet
Fourth Qtr$_{i,t}$	Indicator variable set equal to one if day *t* is in the fourth fiscal quarter for firm *i* and to zero otherwise.	
Other Variables		
Leverage$_{i,q}$	The ratio of long-term and short-term debt to total assets for firm *i* as of the most recent fiscal quarter-end.	Compustat
Momentum$_{i,t}$	The absolute buy-and-hold return for firm *i* on day *t*., averaged by month.	CRSP
ROA$_{i,t}$	The ratio of net income to total assets for firm *i* on day *t* for the trailing 4 quarters.	Compustat
Stock Volatility$_{i,t}$	Monthly average of the standard deviation of daily returns for firm *i* on day *t*.	CRSP
Beta$_{i,t}$	The trailing 12-month monthly beta for firm *i* on day *t*.	CRSP
Ambiguous$_{i,}$	Indicator variable set equal to one if the ticker for firm *i* is deemed ambiguous by Drake et al. (2012). Ticker type designations as "Ambiguous" used and obtained from Drake et al (2011): AA, ABC, ALL, AN, CAT, COST, EBAY, ED, FAST, HAS, HD, HOG, KEY, KO, LOW, MAT, MET, PEG, SEE, TAP.	Drake et al. (2012)
AIA$_{i,t}$	Bloomberg Institutional Attention Measure as per Ben-Rephael et al. 2017 for firm *i* on day *t*, i.e. a dummy variable that receives a value of one if Bloomberg's score is 3 or 4, and zero otherwise. Bloomberg records the number of times news articles on a particular stock are read by its terminal users and the number of times users actively search for news for a specific stock. Bloomberg then assigns a value of one for each article read and ten for each news search. These numbers are then aggregated into an hourly count. Using the hourly count, Bloomberg then creates a numerical attention score each hour by comparing the past eight-hour average count to all hourly counts over the previous month for the same stock. They assign a value of zero if the rolling average is in the lowest 80% of the hourly counts over the previous 30 days. Similarly, Bloomberg assigns a score of 1, 2, 3, or 4 if the average is between 80% and 90%, 90% and 94%, 94% and 96%, or greater than 96% of the previous 30 days' hourly counts, respectively. Finally, Bloomberg aggregates up to the daily frequency by taking a maximum of all hourly scores throughout the day. These are the data provided to us by Bloomberg. Since we are interested in abnormal attention, our AIA measure is a dummy variable that receives a value of one if Bloomberg's score is 3 or 4, and zero otherwise. This captures the right tail of the measure's distribution	Bloomberg

Appendix B: Published Papers Using Google SVI

The following appendix lists all of the published papers that use Google SVI as a measure of investor attention.

Year	Author(s)	Title	Journal
2011	Bank, Larch, and Peter	Google search volume and its influence on liquidity and returns of German stocks	*Financial Markets and Portfolio Management*
2011	Da, Engelberg, and Gao	In search of attention	*Journal of Finance*
2011	Joseph, Wintoki, and Zhang	Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search	*International Journal of Forecasting*
2012	Bordino, Battiston, Caldarelli, Cristelli, Ukkonen and Weber	Web Search Queries can predict stock market volumes	*PLoS ONE*
2012	Choi and Varian	Predicting the present with Google Trends.	*Economic Record 88*
2012	Drake, Roulstone, and Thornock	Investor information demand: Evidence from Google searches around earnings annoucements	*Journal of Accounting Research*
2012	Vlastakis and Markellos	Information demand and stock market volatility	*Journal of Banking & Finance*
2013	Aouadi, Arouri, and Teulon	Investor attention and stock market activity: Evidence from France	*Economic Modelling*
2013	Carrière-Swallow and Labbé	Nowcasting with Google Trends in an emerging marke	*Journal of Forecasting*
2013	Jiang and Li	Investor sentiment and IPO pricing during pre-market and aftermarket periods: Evidence from Hong Kong	*Pacific-Basin Finance Journal*
2013	Korkeamaki and Takalo	Valuation of innovation and intellectual property: The case of the iPhone	*European Management Review*
2013	Kristoufek	Can Google Trends search queries contribute to Risk Diversification	*Nature*
2013	Preis, Moat, and Stanley	Quantifying Trading Behavior in Financial Markets Using Google Trends	*Nature*
2013	Luo, Zhang, and Duan	Social media and firm equity value	*Information Systems Research*
2013	Siganos	Google attention and target price run ups	*International Review of Financial Analysis*
2013	Xiong and Bharadwaj	Asymmetric roles of advertising and marketing capability in financial returns to news: Turning bad into good and good into great	*Journal of Marketing Research*
2013	Xu and Zhang	Impact of Wikpedia on market information: Evidence on management disclosure and investor reaction	*MIS Quarterly*
2013	Zhang, Shen, Zhang, and Xiong	Open source information, investor attention, and asset pricing	*Economic Modelling*
2014	Gwilym, Kita, Wang	Speculate against speculative demand	*International Review of Financial Analysis*
2014	Knittel and Stango	Celebrity endorsements, firm value, and reputation risk: Evidence from the Tiger Woods scandal	*Management Science*
2014	Liu, Ye and Li	Impacts of interactions between news attention and investor attention on stock returns: Empirical investigation on financial shares in China.	*Journal of Management Sciences in China*
2014	Takeda and Wakao	Google search intensity and its relationship with returns and trading volume of Japanese stocks	*Pacific-Basin Finance Journal*
2014	Vaughan	Discovering business infromation from search engine query data	*Online Information Review*
2014	Vozlyublennaia	Investor attention, index performance, and return predictability	*Journal of Banking & Finance*
2014	Zhang, An, Feng	Can online searches be used to forecast stock market performance?	*Journal of Financial Research*
2015	Brown, Stice, and White	Mobile Communication and Local Information Flow: Evidence from Distracted Driving Laws	*Journal of Accounting Research*
2015	Cergol and Omladic	What can Wikipedia and Google tell us about stock prices under different market regimes?	*Ars Mathematica Contemporanea*

2015	deHaan, Shevlin, Thornock	Market (In)Attention and the Strategic Scheduling and Timing of Earnings Announcements	*Journal of Accounting and Economics*
2015	Ding and Hou	Retail investor attention and stock liquidity	*Journal of International Financial Markets, Institutions & Money*
2015	Drake, Roulstone, and Thornock	The Determinants and Consequences of Information Acquisition via EDGAR	*Contemporary Accounting Research*
2015	Goddard, Kita and Wang	Investor attention and FX market volatility	*Journal of International Financial Markets, Institutions & Money*
2015	Hoopes, Reck, Slemrod	Taxpayer Search for Information: Implications for Rational Attention	*American Economic Journal: Economic Policy*
2015	Kristoufek	Power-law correlations in finance-related Google searches, and their cross-correlations with volatility and traded volume	*Physica A: Statistical Mechanics and its Applications*
2015	Li, Ma, Wang, Zhang	How does Google search affect trader positions and crude oil prices?	*Economic Modelling*
2016	Bijl, Kringhaug, Molnar, Sandvik	Google Searches and Stock Returns	*International Review of Financial Analysis*
2016	Curtis, Richardson, and Schmardebeck	Investor attention and the pricing of earnings news	*Handbook of Sentiment Analysis in Finance*
2016	Drake, Jennings, Roulstone and Thornock	The Comovement of Investor Attention	*Management Science*
2016	Fang, Huang, Karpoff	Short Selling and Earnings Management: A Controlled Experiment	*Journal of Finance*
2017	Ben-Rephael, Da, Israelsen	It Depends on Where You Search: Institutional Investor Attention and Underreaction to News	*Review of Financial Studies*
2017	Boulland and Dessaint	Announcing the Announcement	*Journal of Banking & Finance*
2017	Chi and Shantikumar	Local Bias in Google Search and the Market Response around Earnings Announcements	*The Accounting Review*
2017	Colaco, De Cesari, and Hegde	Retail Investor Attention and IPO Valuation	*European Financial Management*
2017	Kong, Lin, Liu	Does Information Acquisition Alleviate Market Anomalies? Categorization Bias in Stock Splits	*Review of Finance*
2017	Madsen	Anticipated Earnings Announcements and the Customer–Supplier Anomaly	*Journal of Accounting Research*
2017	Welagedara, Deb, and Singh	Investor attention, analyst recommendation revisions, and stock prices	*Pacific-Basin Finance Journal*
2018	Chang and Kwon	Ambiguities in valuing information technology firms: Do internet searches T help?	*Journal of Business Research*
2018	Chronopoulos, Papadimitriou, and Vlastakis	Information demand and stock return predictability	*Journal of International Money and Finance*
2018	Frank and Sanati	How does the stock market absorb shocks?	*Journal of Financial Economics*
2018	Hasan, Kumas, Smith	Market ambiguity and individual investor information demand	*Journal of Contemporary Accounting & Economics*
2018	Kupfer and Zorn	Valuable information in early sales proxies: The use of Google search ranks in portfolio optimization	*Journal of Forecasting*
2018	Madsen and Niessner	Is Investor Attention for Sale? The Role of Advertising in Financial Markets	*Journal of Accounting Research*
2018	Mbanga, Darrat, and Park	Investor sentiment and aggregate stock returns: the role of investor attention	*Review of Quantitative Finance and Accounting*
2018	Pantzalis and Ucar	Allergy onset and local investor distraction	*Journal of Banking & Finance*
2018	Reyes	Limited attention and M&A announcements	*Journal of Empirical Finance*
2018	Reyes	Negativity Bias in Attention Allocation: Retail Investors' Reaction to Stock Returns	*International Review of Finance*
2018	Reyes and Waissbluth	Saddled with Attention: Overreaction to Bankruptcy filings	*International Review of Finance*
2018	Wang, Choi, Siraj	Local investor attention and post-earnings announcement drift	*Review of Quantitative Finance and Accounting*
2018	Pantzalis and Ucar	Allergy onset and local investor distraction	*Journal of Banking & Finance*

Forth.	Huang, Huang, and Lin	Attention allocation and return co-movement: Evidence from repeated natural experiments	*Journal of Financial Economics*

Appendix C: Ticker Search Volume and the Fraction Determined to be Investor-Related

This Appendix lists the average ticker searches per firm-month (in units of one) for each of the 490 tickers in our sample (*Keyword_Search*) for our sample period (column (i)). Column (ii) lists the percentage of firms determined to be investor related (*Investor_Search*). "Investor-related" searches are determined based on the contents of the click-through website. Specifically, we designate a website as investor-related if it "likely provides current information for investors about the ticker being searched." See Section 2 for further details. Column (iii) lists the percentage of searches that are not investor related (*Noise_Search*). The columns (iv) and (v) present the results from a firm-level regression of Equation (8) excluding control variables and robust standard errors. Column (iv) presents the increase of ASVI on earnings announcement days and column (iv) the p-value of the coefficient. See Appendix A for variable definitions and Table 3 for the sample composition.

Ticker	Name	*(i)* *Keyword Search*	*(ii)* *Investor Search*	*(iii)* *Noise_Search*	*(iv)* *ASVI on EA days*	*(v)* *Sign. [P-value]*
A	Agilent Technologies Inc	1,519,412	0.66%	99.34%	0.95%	0.619
AA	Alcoa Inc	523,333	1.52%	98.48%	0.17%	0.951
AAL	American Airlines Group	41,855	75.97%	24.03%	60.70%	0.000
AAP	Advance Auto Parts	48,755	17.42%	82.58%	30.10%	0.037
AAPL	Apple Inc.	1,141,600	70.82%	29.18%	278.50%	0.000
ABBV	AbbVie	17,995	62.85%	37.15%	143.00%	0.000
ABC	AmerisourceBergen Corp	1,192,950	0.06%	99.94%	11.70%	0.545
ABT	Abbott Laboratories	85,850	7.64%	92.36%	3.59%	0.568
ACN	Accenture plc	33,610	5.63%	94.37%	58.40%	0.000
ADBE	Adobe Systems Inc	22,750	70.57%	29.43%	324.80%	0.000
ADI	Analog Devices Inc.	41,850	1.26%	98.74%	13.40%	0.071
ADM	Archer-Daniels-Midland Co	39,845	4.97%	95.03%	10.90%	0.187
ADP	Automatic Data Processing	1,396,000	0.29%	99.71%	-2.91%	-0.553
ADS	Alliance Data Systems	192,000	2.65%	97.35%	-20.80%	0.042
ADSK	Autodesk Inc	121,525	84.51%	15.49%	288.00%	0.000
AEE	Ameren Corp	18,020	3.25%	96.75%	-5.01%	-0.437
AEP	American Electric Power	138,750	0.79%	99.21%	-6.29%	-0.272
AES	AES Corp	165,450	1.01%	98.99%	-5.91%	-0.254
AET	Aetna Inc	27,375	6.77%	93.23%	19.30%	0.088
AFL	AFLAC Inc	38,440	0.77%	99.23%	25.70%	0.081
AGN	Allergan plc	23,320	53.45%	46.55%	152.90%	0.000
AIG	American International Group Inc.	71,600	24.86%	75.14%	44.20%	0.012
AIV	Apartment Investment & Mgmt	1,656	10.55%	89.45%	-13.60%	-0.634
AIZ	Assurant Inc	1,859	29.98%	70.02%	42.60%	0.286
AJG	Arthur J. Gallagher & Co.	1,789	3.94%	96.06%	15.00%	0.576
AKAM	Akamai Technologies Inc	13,230	56.16%	43.84%	251.10%	0.000
ALB	Albemarle Corp	247,000	60.10%	39.90%	-6.40%	-0.389
ALK	Alaska Air Group Inc	11,575	48.94%	51.06%	39.40%	0.000
ALL	Allstate Corp	152,625	2.08%	97.92%	2.38%	0.601
ALLE	Allegion	5,824	30.20%	69.80%	6.47%	0.576
ALXN	Alexion Pharmaceuticals	20,240	69.50%	30.50%	166.10%	0.000
AMAT	Applied Materials Inc	20,305	74.12%	25.88%	156.50%	0.000
AME	Ametek	25,630	1.77%	98.23%	-8.67%	0.582
AMG	Affiliated Managers Group Inc	31,370	0.26%	99.74%	-0.91%	-0.840
AMGN	Amgen Inc	25,455	81.56%	18.44%	201.70%	0.000
AMP	Ameriprise Financial	76,300	0.56%	99.44%	-2.54%	-0.124
AMT	American Tower Corp A	39,160	25.90%	74.10%	6.03%	0.635
AMZN	Amazon.com Inc	611,450	66.93%	33.07%	354.10%	0.000
AN	AutoNation Inc	143,571	0.57%	99.43%	2.22%	0.599
ANTM	Anthem Inc.	26,945	1.62%	98.38%	-7.33%	-0.642
AON	Aon plc	29,695	5.26%	94.74%	39.40%	0.066
APA	Apache Corporation	147,450	0.16%	99.84%	2.82%	0.862
APC	Anadarko Petroleum Corp	75,800	0.26%	99.74%	-5.99%	-0.022

Ticker	Name	(i) Keyword Search	(ii) Investor Search	(iii) Noise Search	(iv) ASVI on EA days	(v) Sign. [P-value]
APD	Air Products & Chemicals Inc	53,450	3.23%	96.77%	6.81%	0.530
APH	Amphenol Corp A	5,590	6.83%	93.17%	3.38%	0.773
ATVI	Activision Blizzard	39,565	79.33%	20.67%	349.40%	0.000
AVB	AvalonBay Communities Inc.	6,217	21.98%	78.02%	20.20%	0.410
AVGO	Avago Technologies	23,065	27.95%	72.05%	353.80%	0.000
AVY	Avery Dennison Corp	1,720	66.13%	33.87%	0.54%	0.966
AWK	American Water Works Company Inc	16,685	1.24%	98.76%	1.22%	0.905
AXP	American Express Co	12,960	79.99%	20.01%	306.10%	0.000
AYI	Acuity Brands Inc	12,686	7.35%	92.65%	94.70%	0.000
AZO	AutoZone Inc	66,235	7.27%	92.73%	6.95%	0.387
BA	Boeing Company	181,941	21.56%	78.44%	8.08%	0.003
BAC	Bank of America Corp	316,300	59.79%	40.21%	37.50%	0.000
BAX	Baxter International Inc.	5,953	54.26%	45.74%	29.50%	0.209
BBBY	Bed Bath & Beyond	8,818	74.05%	25.95%	367.60%	0.000
BBT	BB&T Corporation	2,035,000	0.18%	99.82%	-0.05%	-0.991
BBY	Best Buy Co. Inc.	26,015	53.37%	46.63%	161.40%	0.000
BCR	Bard (C.R.) Inc.	11,506	6.69%	93.31%	27.60%	0.271
BDX	Becton Dickinson	8,415	49.23%	50.77%	115.00%	0.002
BEN	Franklin Resources	56,100	4.51%	95.49%	-5.15%	0.496
BFA	Brown-Forman Corporation	14,425	0.07%	99.93%	5.64%	0.793
BFB	Brown-Forman Corporation	2,650	90.33%	9.67%	9.38%	0.838
BIIB	BIOGEN IDEC Inc.	28,805	92.22%	7.78%	235.00%	0.000
BK	The Bank of New York Mellon	50,794	3.46%	96.54%	-5.45%	-0.287
BLK	BlackRock	10,694	19.72%	80.28%	-3.68%	0.404
BLL	Ball Corp	11,900	63.62%	36.38%	37.80%	0.151
BMY	Bristol-Myers Squibb	41,205	48.49%	51.51%	107.60%	0.014
BRKA	Berkshire Hathaway	9,550	80.38%	19.62%	73.80%	0.105
BRKB	Berkshire Hathaway	71,785	73.42%	26.58%	68.60%	0.151
BSX	Boston Scientific	8,030	85.35%	14.65%	183.10%	0.000
BWA	BorgWarner	8,770	1.27%	98.73%	24.50%	0.078
BXP	Boston Properties	938	27.99%	72.01%	16.20%	0.696
C	Citigroup Inc.	1,220,000	8.75%	91.25%	4.55%	0.171
CA	CA Inc.	275,118	0.31%	99.69%	0.57%	0.852
CAG	ConAgra Foods Inc.	18,545	2.62%	97.38%	49.30%	0.000
CAH	Cardinal Health Inc.	17,890	28.59%	71.41%	37.30%	0.013
CAT	Caterpillar Inc.	1,179,000	6.06%	93.94%	8.61%	0.050
CB	Chubb Limited	1,045,588	3.11%	96.89%	-1.52%	-0.692
CBG	CBRE Group	9,000	2.32%	97.68%	31.30%	0.237
CBS	CBS Corp.	1,029,200	0.08%	99.92%	-14.80%	-0.140
CBSA	CBS Corp.	7,935	0.00%	100.00%	32.10%	0.153
CCI	Crown Castle International Corp.	35,080	5.06%	94.94%	14.10%	0.102
CCL	Carnival Corp.	18,015	13.53%	86.47%	60.90%	0.000
CELG	Celgene Corp.	38,310	39.74%	60.26%	130.80%	0.003
CERN	Cerner	77,725	7.11%	92.89%	6.25%	0.697
CF	CF Industries Holdings Inc	59,353	7.32%	92.68%	8.35%	0.005
CFG	Citizens Financial Group	9,595	34.91%	65.09%	45.50%	0.000
CHD	Church & Dwight	22,455	6.36%	93.64%	101.00%	0.015
CHK	Chesapeake Energy	134,250	70.77%	29.23%	56.70%	0.010
CHRW	C. H. Robinson Worldwide	5,284	58.51%	41.49%	156.00%	0.002
CI	CIGNA Corp.	60,500	3.13%	96.87%	-5.37%	-0.339
CINF	Cincinnati Financial	2,023	80.10%	19.90%	94.50%	0.165
CL	Colgate-Palmolive	181,941	7.50%	92.50%	-4.60%	-0.190
CLX	The Clorox Company	4,338	63.10%	36.90%	83.80%	0.011
CMA	Comerica Inc.	58,130	4.68%	95.32%	-31.30%	-0.039
CMCSA	Comcast A Corp	28,220	51.41%	48.59%	173.80%	0.000
CME	CME Group Inc.	52,975	76.22%	23.78%	-2.42%	0.564
CMG	Chipotle Mexican Grill	98,075	80.46%	19.54%	234.80%	0.000
CMI	Cummins Inc.	23,465	14.01%	85.99%	47.20%	0.002
CMS	CMS Energy	202,200	0.06%	99.94%	7.20%	0.025

Ticker	Name	*(i)* ***Keyword Search***	*(ii)* ***Investor Search***	*(iii)* ***Noise Search***	*(iv)* ***ASVI on EA days***	*(v)* ***Sign. [P-value]***
CNC	Centene Corporation	37,910	7.43%	92.57%	7.91%	0.035
CNP	CenterPoint Energy	7,890	24.15%	75.85%	26.10%	0.198
COF	Capital One Financial	27,990	53.25%	46.75%	95.20%	0.000
COG	Cabot Oil & Gas	39,550	9.16%	90.84%	0.47%	0.948
COL	Rockwell Collins	50,450	11.37%	88.63%	0.49%	0.950
COP	ConocoPhillips	99,350	31.90%	68.10%	-4.11%	-0.609
COST	Costco Co.	62,382	40.11%	59.89%	-1.04%	-0.684
CPB	Campbell Soup	23,765	4.63%	95.37%	65.10%	0.002
CRM	Salesforce.com	142,750	8.16%	91.84%	15.70%	0.019
CSCO	Cisco Systems	81,975	47.29%	52.71%	288.80%	0.000
CSRA	CSRA Inc.	22,155	1.81%	98.19%	24.00%	0.152
CSX	CSX Corp.	56,150	26.56%	73.44%	31.50%	0.026
CTAS	Cintas Corporation	92,750	27.00%	73.00%	24.40%	0.488
CTL	CenturyLink Inc	19,347	52.91%	47.09%	61.00%	0.037
CTSH	Cognizant Technology Solutions	67,075	66.67%	33.33%	197.00%	0.000
CTXS	Citrix Systems	6,055	84.48%	15.52%	199.60%	0.000
CVS	CVS Health	3,424,000	0.35%	99.65%	-3.41%	-0.489
CVX	Chevron Corp.	60,800	42.98%	57.02%	66.10%	0.001
CXO	Concho Resources	5,160	6.81%	93.19%	44.60%	0.034
D	Dominion Resources	823,000	6.23%	93.77%	-3.95%	-0.347
DAL	Delta Air Lines	39,180	30.34%	69.66%	29.60%	0.000
DE	Deere & Co.	165,000	33.69%	66.31%	-5.20%	-0.105
DFS	Discover Financial Services	31,670	3.77%	96.23%	-0.22%	-0.979
DG	Dollar General	35,276	22.10%	77.90%	15.40%	0.053
DGX	Quest Diagnostics	5,089	33.40%	66.60%	42.30%	0.043
DHI	D. R. Horton	4,929	14.63%	85.37%	56.60%	0.027
DHR	Danaher Corp.	17,935	21.76%	78.24%	4.64%	0.621
DIS	The Walt Disney Company	136,805	64.54%	35.46%	55.60%	0.000
DISCA	Discovery Communications-A	1,906	77.17%	22.83%	159.20%	0.019
DISCK	Discovery Communications-C	1,212	97.38%	2.62%	13.50%	0.668
DLR	Digital Realty Trust	5,580	25.71%	74.29%	17.70%	0.230
DLTR	Dollar Tree	5,550	41.56%	58.44%	270.20%	0.000
DNB	Dun & Bradstreet	14,529	2.34%	97.66%	-6.74%	-0.300
DO	Diamond Offshore Drilling	208,941	0.65%	99.35%	6.37%	0.460
DOV	Dover Corp.	5,300	82.82%	17.18%	-19.20%	-0.063
DPS	Dr. Pepper Snapple Group	131,500	0.18%	99.82%	-10.10%	-0.093
DRI	Darden Restaurants	42,700	9.27%	90.73%	11.80%	0.122
DTE	DTE Energy Co.	126,500	18.22%	81.78%	-13.40%	-0.469
DUK	Duke Energy	15,455	59.66%	40.34%	38.50%	0.009
DVA	DaVita Inc.	44,830	5.17%	94.83%	54.40%	0.381
DVN	Devon Energy Corp.	10,647	46.74%	53.26%	59.00%	0.148
EA	Electronic Arts	187,235	16.94%	83.06%	-3.33%	-0.636
EBAY	eBay Inc.	44,300,000	0.00%	100.00%	-2.05%	0.409
ECL	Ecolab Inc.	7,100	39.95%	60.05%	19.90%	0.436
ED	Consolidated Edison	245,118	3.63%	96.37%	-2.45%	-0.778
EFX	Equifax Inc.	11,535	53.69%	46.31%	26.00%	0.077
EIX	Edison Int'l	4,508	67.57%	32.43%	31.60%	0.320
EL	Estee Lauder Cos.	90,676	5.77%	94.23%	-7.02%	-0.001
EMN	Eastman Chemical	5,706	82.25%	17.75%	59.00%	0.106
EMR	Emerson Electric Company	35,390	13.61%	86.39%	1.66%	0.790
ENDP	Endo International	64,588	59.21%	40.79%	192.70%	0.000
EOG	EOG Resources	8,665	24.57%	75.43%	37.20%	0.449
EQIX	Equinix	12,937	17.66%	82.34%	188.10%	0.009
EQR	Equity Residential	2,035	17.01%	82.99%	146.20%	0.024
EQT	EQT Corporation	17,915	13.95%	86.05%	19.10%	0.343
ES	Eversource Energy	348,294	0.30%	99.70%	-0.73%	-0.894
ESRX	Express Scripts	11,330	83.44%	16.56%	135.60%	0.000
ESS	Essex Property Trust Inc	70,625	2.19%	97.81%	-6.14%	-0.214
ETFC	E*Trade	6,118	96.65%	3.35%	192.20%	0.002

Ticker	Name	(i) Keyword Search	(ii) Investor Search	(iii) Noise Search	(iv) ASVI on EA days	(v) Sign. [P-value]
ETN	Eaton Corporation	18,340	16.02%	83.98%	29.10%	0.258
ETR	Entergy Corp.	10,230	14.52%	85.48%	19.80%	0.289
EW	Edwards Lifesciences Corp.	170,631	0.86%	99.16%	1.00%	0.226
EXC	Exelon Corp.	18,235	67.73%	32.27%	21.80%	0.040
EXPD	Expeditors Int'l	4,329	28.48%	71.52%	178.60%	0.000
EXPE	Expedia Inc.	15,035	51.52%	48.48%	240.10%	0.000
EXR	Extra Space Storage	3,494	29.89%	70.11%	10.10%	0.521
F	Ford Motor	3,999,412	1.13%	98.87%	-0.02%	-0.994
FAST	Fastenal Co	91,775	1.40%	98.60%	9.41%	0.335
FB	Facebook	5,671,176	1.60%	98.40%	30.90%	0.002
FBHS	Fortune Brands Home & Security	1,456	5.47%	94.53%	122.10%	0.067
FCX	Freeport-McMoran Cp & Gld	57,220	60.15%	39.85%	81.80%	0.006
FDX	FedEx Corporation	13,650	81.27%	18.73%	332.50%	0.000
FE	FirstEnergy Corp	78,853	1.57%	98.43%	-3.80%	0.193
FFIV	F5 Networks	9,795	64.96%	35.04%	198.20%	0.000
FIS	Fidelity National Information Services	29,500	2.69%	97.31%	-7.04%	-0.513
FISV	Fiserv Inc	34,200	69.93%	30.07%	225.30%	0.000
FITB	Fifth Third Bancorp	8,339	19.81%	80.19%	117.10%	0.006
FL	Foot Locker Inc	89,265	25.55%	74.45%	-8.49%	0.031
FLIR	FLIR Systems	37,990	0.60%	99.40%	-5.57%	-0.473
FLR	Fluor Corp.	8,385	12.63%	87.37%	-1.99%	-0.817
FLS	Flowserve Corporation	5,178	3.15%	96.85%	39.50%	0.018
FMC	FMC Corporation	22,115	9.39%	90.61%	9.97%	0.060
FOX	Twenty-First Century Fox Class B	1,471,250	0.42%	99.58%	23.70%	0.186
FOXA	Twenty-First Century Fox Class A	3,288	88.31%	11.69%	205.90%	0.001
FRT	Federal Realty Investment Trust	3,950	47.05%	52.95%	29.20%	0.039
FSLR	First Solar Inc	37,610	69.03%	30.97%	189.60%	0.000
FTI	FMC Technologies Inc.	4,900	1.82%	98.18%	3.48%	0.608
FTR	Frontier Communications	33,465	38.84%	61.16%	48.80%	0.004
FTV	Fortive Corp	52,250	1.58%	98.42%	11.00%	0.052
GD	General Dynamics	37,453	38.07%	61.93%	4.46%	0.409
GE	General Electric	170,143	32.60%	67.40%	17.30%	0.113
GGP	General Growth Properties Inc.	6,080	10.14%	89.86%	16.00%	0.601
GILD	Gilead Sciences	91,700	61.25%	38.75%	199.20%	0.000
GIS	General Mills	68,600	2.73%	97.27%	-1.86%	-0.678
GLW	Corning Inc.	14,585	64.95%	35.05%	169.20%	0.000
GM	General Motors	1,486,471	2.53%	97.47%	7.53%	0.002
GOOG	Alphabet Inc Class C	1,026,300	27.67%	72.33%	98.20%	0.000
GOOGL	Alphabet Inc Class A	547,300	19.51%	80.49%	44.70%	0.000
GPC	Genuine Parts	14,195	10.60%	89.40%	-3.57%	-0.743
GPN	Global Payments Inc	3,220	23.08%	76.92%	90.30%	0.004
GPS	Gap (The)	437,800	1.06%	98.94%	-2.36%	0.679
GRMN	Garmin Ltd.	9,340	52.31%	47.69%	252.50%	0.000
GS	Goldman Sachs Group	118,824	52.47%	47.53%	10.40%	0.227
GT	Goodyear Tire & Rubber	75,588	5.85%	94.15%	-1.98%	0.583
GWW	Grainger (W.W.) Inc.	4,350	10.66%	89.34%	227.80%	0.000
HAL	Halliburton Co.	29,680	31.02%	68.98%	-4.45%	-0.421
HAS	Hasbro Inc.	53,200	20.09%	79.91%	13.70%	0.163
HBAN	Huntington Bancshares	10,806	47.49%	52.51%	101.60%	0.001
HBI	Hanesbrands Inc	10,165	46.17%	53.83%	75.60%	0.081
HCA	HCA Holdings	40,660	3.73%	96.27%	3.78%	0.624
HCN	Welltower Inc.	22,476	23.40%	76.60%	33.70%	0.224
HCP	HCP Inc.	13,045	47.33%	52.67%	69.10%	0.001
HD	Home Depot	101,971	30.34%	69.66%	5.10%	0.348
HES	Hess Corporation	14,275	8.68%	91.32%	-0.65%	0.729
HIG	Hartford Financial Svc.Gp.	10,340	13.00%	87.00%	28.10%	0.035
HOG	Harley-Davidson	39,470	7.48%	92.52%	-6.72%	0.505
HOLX	Hologic	3,072	75.10%	24.90%	253.00%	0.000
HON	Honeywell Int'l Inc.	33,060	12.91%	87.09%	20.50%	0.029

Ticker	Name	*(i)* *Keyword Search*	*(ii)* *Investor Search*	*(iii)* *Noise Search*	*(iv)* *ASVI on EA days*	*(v)* *Sign. [P-value]*
HP	Helmerich & Payne	363,647	0.51%	99.49%	5.51%	0.049
HPE	Hewlett Packard Enterprise	32,465	30.48%	69.52%	137.90%	0.003
HPQ	HP Inc.	16,455	70.53%	29.47%	234.10%	0.000
HRB	Block H&R	4,915	52.71%	47.29%	100.50%	0.002
HRL	Hormel Foods Corp.	4,700	17.37%	82.63%	131.90%	0.000
HRS	Harris Corporation	10,270	4.36%	95.64%	-0.34%	0.944
HSIC	Henry Schein	2,446	81.74%	18.26%	198.40%	0.001
HST	Host Hotels & Resorts	9,650	4.90%	95.10%	4.94%	0.606
HSY	The Hershey Company	4,430	58.27%	41.73%	64.60%	0.003
HUM	Humana Inc.	84,150	1.53%	98.47%	-4.51%	0.784
IBM	International Bus. Machines	169,200	21.39%	78.61%	38.00%	0.000
ICE	Intercontinental Exchange	213,450	3.08%	96.92%	-7.92%	0.253
IFF	Intl Flavors & Fragrances	10,030	0.77%	99.23%	-8.22%	0.201
ILMN	Illumina Inc	60,925	77.33%	22.67%	151.30%	0.003
INTC	Intel Corp.	109,900	63.57%	36.43%	311.50%	0.000
INTU	Intuit Inc.	6,826	53.65%	46.35%	297.80%	0.000
IP	International Paper	332,529	0.08%	99.92%	-3.35%	0.055
IPG	Interpublic Group	12,015	1.78%	98.22%	47.20%	0.041
IR	Ingersoll-Rand PLC	49,682	3.19%	96.81%	7.59%	0.097
IRM	Iron Mountain Incorporated	9,905	13.53%	86.47%	28.60%	0.206
ISRG	Intuitive Surgical Inc.	18,410	76.96%	23.04%	338.80%	0.000
ITW	Illinois Tool Works	13,920	4.66%	95.34%	48.10%	0.011
IVZ	Invesco Ltd.	1,395	41.90%	58.10%	184.10%	0.006
JBHT	J. B. Hunt Transport Services	1,671	100.00%	0.00%	229.90%	0.000
JCI	Johnson Controls	12,775	16.47%	83.53%	33.10%	0.045
JEC	Jacobs Engineering Group	4,119	26.73%	73.27%	64.70%	0.182
JNJ	Johnson & Johnson	32,235	70.84%	29.16%	74.20%	0.000
JNPR	Juniper Networks	8,763	65.51%	34.49%	275.60%	0.000
JPM	JPMorgan Chase & Co.	124,700	56.30%	43.70%	246.30%	0.000
JWN	Nordstrom	11,295	76.62%	23.38%	279.10%	0.000
K	Kellogg Co.	805,353	0.07%	99.93%	-2.06%	0.522
KEY	KeyCorp	131,500	1.38%	98.62%	5.23%	0.231
KHC	Kraft Heinz Co	9,425	42.28%	57.72%	142.50%	0.000
KIM	Kimco Realty	86,800	13.39%	86.61%	-3.74%	0.748
KLAC	KLA-Tencor Corp.	5,416	21.29%	78.71%	140.90%	0.004
KMB	Kimberly-Clark	5,621	50.20%	49.80%	161.30%	0.000
KMI	Kinder Morgan	27,750	81.79%	18.21%	74.70%	0.001
KMX	Carmax Inc	6,869	60.54%	39.46%	313.30%	0.000
KO	The Coca Cola Company	129,118	48.21%	51.79%	6.12%	0.214
KORS	Michael Kors Holdings	6,540	34.51%	65.49%	8.66%	0.634
KR	Kroger Co.	44,871	72.64%	27.36%	53.20%	0.001
KSS	Kohl's Corp.	11,570	12.69%	87.31%	246.30%	0.000
KSU	Kansas City Southern	40,760	1.61%	98.39%	1.31%	0.894
L	Loews Corp.	1,220,000	2.29%	97.71%	1.55%	0.493
LB	L Brands Inc.	43,865	24.37%	75.63%	14.50%	0.281
LEG	Leggett & Platt	58,850	0.28%	99.72%	-4.66%	0.584
LEN	Lennar Corp.	16,353	2.30%	97.70%	14.30%	0.055
LENB	Lennar Corp.	90	0.00%	100.00%	-16.60%	0.336
LH	Laboratory Corp. of America Holding	32,641	20.74%	79.26%	-3.87%	0.404
LKQ	LKQ Corporation	160,500	0.32%	99.68%	2.94%	0.526
LLL	L-3 Communications Holdings	52,800	17.55%	82.45%	10.00%	0.028
LLY	Lilly (Eli) & Co.	14,820	67.92%	32.08%	26.10%	0.119
LM	Legg Mason	27,100	2.16%	97.84%	1.98%	0.799
LMT	Lockheed Martin Corp.	49,675	47.37%	52.63%	29.40%	0.010
LNC	Lincoln National	3,993	37.93%	62.07%	26.90%	0.149
LNT	Alliant Energy Corp	2,756	12.16%	87.84%	-39.30%	0.074
LOW	Lowe's Cos.	62,775	29.06%	70.94%	-3.05%	0.267
LRCX	Lam Research	14,805	79.46%	20.54%	196.90%	0.000
LUK	Leucadia National Corp.	5,012	50.44%	49.56%	25.02%	0.190

Ticker	Name	(i) Keyword Search	(ii) Investor Search	(iii) Noise Search	(iv) ASVI on EA days	(v) Sign. [P-value]
LUV	Southwest Airlines	41,145	61.25%	38.75%	15.90%	0.271
LYB	LyondellBasell	2,289	41.39%	58.61%	149.50%	0.024
M	Macy's Inc.	1,502,941	2.20%	97.80%	3.52%	0.358
MA	Mastercard Inc.	258,941	8.69%	91.31%	9.24%	0.356
MAC	Macerich	634,000	0.52%	99.48%	1.83%	0.535
MAR	Marriott Int'l.	197,300	21.15%	78.85%	-2.37%	0.795
MAS	Masco Corp.	61,175	0.28%	99.72%	0.93%	0.492
MAT	Mattel Inc.	53,450	18.94%	81.06%	3.38%	0.245
MCD	McDonald's Corp.	42,806	54.09%	45.91%	137.20%	0.002
MCHP	Microchip Technology	5,700	68.12%	31.88%	170.60%	0.010
MCK	McKesson Corp.	14,720	74.56%	25.44%	81.60%	0.083
MCO	Moody's Corp	46,225	1.51%	98.49%	-10.70%	0.073
MDLZ	Mondelez International	4,685	33.76%	66.24%	231.90%	0.000
MDT	Medtronic plc	37,540	17.91%	82.09%	24.60%	0.018
MET	MetLife Inc.	79,775	82.21%	17.79%	46.20%	0.291
MHK	Mohawk Industries	2,426	9.25%	90.75%	39.10%	0.153
MKC	McCormick & Co.	5,912	5.73%	94.27%	15.60%	0.383
MLM	Martin Marietta Materials	32,015	13.03%	86.97%	-0.66%	0.925
MMC	Marsh & McLennan	17,110	3.80%	96.20%	6.27%	0.351
MMM	3M Company	70,100	21.22%	78.78%	33.30%	0.019
MNK	Mallinckrodt Plc	11,294	50.62%	49.38%	195.70%	0.000
MNST	Monster Beverage	5,453	58.72%	41.28%	265.60%	0.000
MO	Altria Group Inc	224,824	28.07%	71.93%	1.31%	0.653
MON	Monsanto Co.	44,500	30.87%	69.13%	-9.55%	0.199
MOS	The Mosaic Company	41,130	4.99%	95.01%	-10.80%	0.017
MPC	Marathon Petroleum	45,730	8.16%	91.84%	12.90%	0.032
MRK	Merck & Co.	32,241	47.44%	52.56%	87.60%	0.012
MRO	Marathon Oil Corp.	40,170	27.57%	72.43%	23.00%	0.088
MS	Morgan Stanley	334,500	2.83%	97.17%	4.03%	0.438
MSFT	Microsoft Corp.	191,050	70.85%	29.15%	334.70%	0.000
MSI	Motorola Solutions Inc.	127,575	0.19%	99.81%	24.40%	0.243
MTB	M&T Bank Corp.	93,875	0.31%	99.69%	0.23%	0.972
MU	Micron Technology	188,467	36.36%	63.64%	22.60%	0.000
MUR	Murphy Oil	7,806	59.48%	40.52%	-6.53%	0.616
MYL	Mylan N.V.	14,605	74.04%	25.96%	96.70%	0.076
NAVI	Navient	27,870	2.78%	97.22%	9.00%	0.487
NBL	Noble Energy Inc	13,285	1.39%	98.61%	26.80%	0.457
NDAQ	NASDAQ OMX Group	2,173	68.32%	31.68%	169.40%	0.020
NEE	NextEra Energy	50,600	8.25%	91.75%	-7.47%	0.289
NEM	Newmont Mining Corp. (Hldg. Co.)	17,465	65.61%	34.39%	-1.13%	0.922
NFLX	Netflix Inc.	205,450	51.17%	48.83%	370.80%	0.000
NFX	Newfield Exploration Co	2,130	18.43%	81.57%	17.60%	0.652
NI	NiSource Inc.	82,735	0.30%	99.70%	-3.26%	0.360
NKE	Nike	43,265	53.03%	46.97%	354.40%	0.000
NLSN	Nielsen Holdings	1,681	77.76%	22.24%	225.50%	0.002
NOC	Northrop Grumman Corp.	29,625	14.78%	85.22%	10.30%	0.454
NOV	National Oilwell Varco Inc.	25,765	13.85%	86.15%	22.00%	0.635
NRG	NRG Energy	47,035	5.44%	94.56%	2.95%	0.920
NSC	Norfolk Southern Corp.	19,715	9.03%	90.97%	25.90%	0.323
NTAP	NetApp	14,415	64.69%	35.31%	359.00%	0.000
NTRS	Northern Trust Corp.	3,476	59.30%	40.70%	200.80%	0.001
NUE	Nucor Corp.	78,275	21.76%	78.24%	-8.33%	0.459
NVDA	Nvidia Corporation	477,150	73.93%	26.07%	257.70%	0.000
NWL	Newell Rubbermaid Co.	5,184	66.31%	33.69%	159.10%	0.001
NWS	News Corp. Class B	410,500	0.01%	99.99%	-24.60%	0.001
NWSA	News Corp. Class A	3,115	1.91%	98.09%	96.90%	0.038
O	Realty Income Corporation	658,529	7.39%	92.61%	3.86%	0.160
OI	Owens-Illinois Inc.	7,113	4.94%	95.09%	7.90%	0.165
OKE	ONEOK	7,089	20.97%	79.03%	4.55%	0.611

Ticker	Name	(i) Keyword Search	(ii) Investor Search	(iii) Noise Search	(iv) ASVI on EA days	(v) Sign. [P-value]
OMC	Omnicom Group	10,890	1.24%	98.76%	12.70%	0.499
ORCL	Oracle Corp.	137,730	73.33%	26.67%	317.10%	0.000
ORLY	O'Reilly Automotive	10,815	13.32%	86.68%	23.40%	0.055
OXY	Occidental Petroleum	21,255	18.08%	81.92%	14.80%	0.020
PAYX	Paychex Inc.	2,919	47.16%	52.84%	178.30%	0.002
PBCT	People's United Financial	2,331	72.16%	27.84%	48.20%	0.388
PBI	Pitney-Bowes	14,540	8.35%	91.65%	5.61%	0.517
PCAR	PACCAR Inc.	4,144	22.54%	77.46%	75.90%	0.000
PCG	PG&E Corp.	10,695	20.58%	79.42%	-5.99%	0.624
PCLN	Priceline.com Inc	40,475	73.32%	26.68%	322.10%	0.000
PDCO	Patterson Companies	3,475	55.05%	44.95%	276.40%	0.000
PEG	Public Serv. Enterprise Inc.	55,635	5.24%	94.76%	1.46%	0.728
PEP	PepsiCo Inc.	90,714	5.92%	94.08%	7.50%	0.069
PFE	Pfizer Inc.	46,190	78.29%	21.71%	94.00%	0.000
PFG	Principal Financial Group	15,490	38.31%	61.69%	-2.23%	0.797
PG	Procter & Gamble	52,735	18.41%	81.59%	13.60%	0.073
PGR	Progressive Corp.	8,130	30.18%	69.82%	33.20%	0.036
PH	Parker-Hannifin	96,588	0.62%	99.38%	5.02%	0.603
PHM	Pulte Homes Inc.	23,400	24.17%	75.83%	-30.50%	0.003
PKI	PerkinElmer	10,360	0.07%	99.93%	-8.35%	0.638
PLD	Prologis	7,242	45.34%	54.66%	-5.69%	0.297
PM	Philip Morris International	142,059	13.96%	86.04%	1.89%	0.855
PNC	PNC Financial Services	1,472,000	0.10%	99.90%	1.95%	0.538
PNR	Pentair Ltd.	6,485	1.14%	98.86%	13.10%	0.541
PNW	Pinnacle West Capital	21,050	2.17%	97.83%	18.70%	0.270
PPG	PPG Industries	40,000	61.53%	38.47%	-5.74%	0.284
PPL	PPL Corp.	121,250	0.84%	99.16%	-3.27%	0.407
PRGO	Perrigo	5,067	81.00%	19.00%	163.80%	0.003
PRU	Prudential Financial	92,305	55.09%	44.91%	-3.37%	0.577
PSA	Public Storage	122,500	0.54%	99.46%	2.80%	0.633
PSX	Phillips 66	33,090	28.49%	71.51%	-4.78%	0.725
PVH	PVH Corp.	15,190	6.49%	93.51%	82.60%	0.000
PWR	Quanta Services Inc.	7,665	2.74%	97.26%	-4.02%	0.766
PX	Praxair Inc.	31,324	0.87%	99.13%	-3.88%	0.417
PXD	Pioneer Natural Resources	5,539	44.14%	55.86%	154.90%	0.003
PYPL	PayPal	36,355	77.07%	22.93%	310.10%	0.000
QCOM	QUALCOMM Inc.	165,800	74.84%	25.16%	233.90%	0.000
QRVO	Qorvo	9,489	70.46%	29.54%	228.00%	0.000
R	Ryder System	805,353	0.37%	99.63%	-0.75%	0.904
RCL	Royal Caribbean Cruises Ltd	15,085	10.93%	89.07%	67.70%	0.000
REGN	Regeneron	27,515	68.16%	31.84%	183.70%	0.000
RF	Regions Financial Corp.	40,688	19.32%	80.68%	1.60%	0.709
RHI	Robert Half International	3,973	23.15%	76.85%	-4.76%	0.717
RHT	Red Hat Inc.	11,375	58.99%	41.01%	289.40%	0.000
RIG	Transocean	50,450	50.50%	49.50%	-6.82%	0.165
RL	Polo Ralph Lauren Corp.	23,159	25.68%	74.32%	6.93%	0.439
ROK	Rockwell Automation Inc.	15,500	11.96%	88.04%	20.30%	0.088
ROP	Roper Industries	14,155	0.79%	99.21%	12.80%	0.286
ROST	Ross Stores	6,955	61.63%	38.37%	20.30%	0.386
RRC	Range Resources Corp.	8,870	28.89%	71.11%	16.20%	0.178
RSG	Republic Services Inc	5,432	16.66%	83.34%	-9.33%	0.150
RTN	Raytheon Co.	17,245	83.16%	16.84%	41.70%	0.008
SBUX	Starbucks Corp.	67,750	49.46%	50.54%	254.40%	0.000
SCG	SCANA Corp	27,920	1.50%	98.50%	14.50%	0.108
SCHW	Charles Schwab Corporation	8,465	39.34%	60.66%	87.20%	0.003
SEE	Sealed Air Corp.	84,050	4.69%	95.31%	-5.32%	0.422
SHW	Sherwin-Williams Company	17,408	47.59%	42.41%	101.71%	0.000
SIG	Signet Jewelers	46,450	1.12%	98.88%	2.33%	0.796
SJM	Smucker (J.M.)	4,711	72.08%	27.92%	235.70%	0.000

Ticker	Name	(i) *Keyword Search*	(ii) *Investor Search*	(iii) *Noise Search*	(iv) *ASVI on EA days*	(v) *Sign. [P-value]*
SLB	Schlumberger Ltd.	14,380	38.07%	61.93%	116.70%	0.000
SLG	SL Green Realty	13,436	0.74%	99.26%	-10.30%	0.616
SNA	Snap-On Inc.	34,280	3.97%	96.03%	1.16%	0.850
SNI	Scripps Networks Interactive Inc.	6,435	16.17%	83.83%	33.10%	0.147
SO	Southern Co.	173,471	14.93%	85.07%	5.53%	0.402
SPG	Simon Property Group Inc	260,250	0.18%	99.82%	-8.59%	0.083
SPGI	S&P Global Inc.	2,054	82.65%	17.35%	244.50%	0.000
SRCL	Stericycle Inc	2,950	72.10%	27.90%	247.10%	0.000
SRE	Sempra Energy	18,755	13.18%	86.82%	4.40%	0.507
STI	SunTrust Banks	85,700	1.08%	98.92%	-1.74%	0.672
STT	State Street Corp.	12,384	42.81%	57.19%	22.20%	0.016
STX	Seagate Technology	17,750	30.89%	69.11%	52.40%	0.001
STZ	Constellation Brands	9,236	69.49%	30.51%	321.00%	0.000
SWK	Stanley Black & Decker	5,550	80.79%	19.21%	142.50%	0.000
SWKS	Skyworks Solutions	20,240	38.83%	61.17%	227.40%	0.000
SWN	Southwestern Energy	12,725	72.30%	27.70%	35.30%	0.048
SYF	Synchrony Financial	5,947	93.73%	6.27%	73.80%	0.091
SYK	Stryker Corp.	5,357	67.95%	32.05%	136.70%	0.000
SYMC	Symantec Corp.	60,875	73.99%	26.01%	184.30%	0.000
SYY	Sysco Corp.	7,368	53.79%	46.21%	252.90%	0.000
T	AT&T Inc	1,077,647	25.10%	74.90%	0.20%	0.880
TAP	Molson Coors Brewing Company	140,500	0.16%	99.84%	-3.44%	0.333
TAPA	Molson Coors Brewing Company	123,300	0.00%	100.00%	-16.20%	0.266
TDC	Teradata Corp.	16,160	21.83%	78.17%	32.10%	0.058
TDG	TransDigm Group	4,559	55.63%	44.37%	41.40%	0.086
TEL	TE Connectivity Ltd.	22,115	0.61%	99.39%	-15.20%	0.084
TGNA	Tegna	989	60.33%	39.67%	268.50%	0.000
TGT	Target Corp.	20,840	48.13%	51.87%	246.30%	0.000
TIF	Tiffany & Co.	20,890	12.71%	87.29%	28.80%	0.000
TJX	TJX Companies Inc.	38,730	4.10%	95.90%	60.20%	0.000
TMK	Torchmark Corp.	2,870	17.11%	82.89%	-10.80%	0.507
TMO	Thermo Fisher Scientific	24,950	5.51%	94.49%	25.60%	0.150
TRIP	TripAdvisor	71,500	1.64%	98.36%	-6.83%	0.492
TROW	T. Rowe Price Group	7,455	17.43%	82.57%	21.50%	0.194
TRV	The Travelers Companies Inc.	4,635	35.25%	64.75%	112.10%	0.000
TSCO	Tractor Supply Company	5,540	68.00%	32.00%	99.50%	0.010
TSN	Tyson Foods	73,125	1.82%	98.18%	3.39%	0.877
TSS	Total System Services	53,780	5.23%	94.77%	6.25%	0.389
TWX	Time Warner Inc.	14,120	75.61%	24.39%	148.00%	0.002
TXN	Texas Instruments	23,585	84.14%	15.86%	254.80%	0.000
TXT	Textron Inc.	15,790	34.39%	65.61%	10.70%	0.051
UA	Under Armour	102,618	19.77%	80.23%	24.60%	0.000
UAL	United Continental Holdings	79,675	23.46%	76.54%	73.90%	0.093
UDR	UDR Inc	7,545	1.28%	98.72%	27.60%	0.356
UHS	Universal Health Services Inc.	35,260	0.37%	99.63%	-5.77%	0.419
ULTA	Ulta Salon Cosmetics & Fragrance Inc	2,818,500	0.45%	99.55%	2.35%	0.740
UNH	United Health Group Inc.	41,720	6.82%	93.18%	1.76%	0.794
UNM	Unum Group	35,020	0.00%	100.00%	2.58%	0.618
UNP	Union Pacific	13,235	64.42%	35.58%	92.00%	0.000
UPS	United Parcel Service	3,645,000	0.28%	99.72%	-15.00%	0.003
URBN	Urban Outfitters	10,825	16.65%	83.35%	182.70%	0.000
URI	United Rentals Inc.	59,775	3.96%	96.04%	4.54%	0.725
USB	U.S. Bancorp	74,000	2.22%	97.78%	-1.73%	0.667
UTX	United Technologies	26,585	60.68%	39.32%	102.70%	0.000
V	Visa Inc.	854,235	3.31%	96.69%	-12.30%	0.248
VAR	Varian Medical Systems	16,450	1.05%	98.95%	3.18%	0.576
VFC	V.F. Corp.	8,100	21.26%	78.74%	44.10%	0.062
VIAB	Viacom Inc.	4,200	83.34%	16.66%	231.20%	0.000
VLO	Valero Energy	14,729	50.85%	49.15%	41.20%	0.059

Ticker	Name	(i) Keyword Search	(ii) Investor Search	(iii) Noise Search	(iv) ASVI on EA days	(v) Sign. [P-value]
VMC	Vulcan Materials	8,085	8.86%	91.14%	4.18%	0.680
VNO	Vornado Realty Trust	1,525	42.04%	57.96%	-25.90%	0.428
VRSK	Verisk Analytics	3,050	82.69%	17.31%	79.80%	0.054
VRSN	Verisign Inc.	3,188	78.40%	21.60%	214.70%	0.001
VRTX	Vertex Pharmaceuticals Inc	9,070	59.82%	40.18%	98.50%	0.001
VTR	Ventas Inc	6,330	28.83%	71.17%	33.00%	0.085
VZ	Verizon Communications	75,676	56.74%	43.26%	66.60%	0.000
WAT	Waters Corporation	72,275	0.16%	99.84%	-2.99%	0.494
WBA	Walgreens Boots Alliance	18,347	33.98%	66.02%	150.30%	0.000
WDC	Western Digital	20,635	45.61%	54.39%	90.40%	0.025
WEC	Wisconsin Energy Corporation	9,680	3.56%	96.44%	38.30%	0.130
WFC	Wells Fargo	82,825	63.66%	36.34%	122.80%	0.000
WHR	Whirlpool Corp.	9,422	66.72%	33.28%	74.20%	0.010
WLTW	Willis Towers Watson	1,056	37.72%	62.28%	106.30%	0.152
WM	Waste Management Inc.	33,971	2.04%	97.96%	-2.97%	0.501
WMB	Williams Cos.	11,685	63.80%	36.20%	33.20%	0.192
WMT	Wal-Mart Stores	66,975	69.08%	30.92%	204.80%	0.000
WRK	Westrock Co	2,720	13.05%	86.95%	7.37%	0.551
WU	Western Union Co	33,100	1.80%	98.20%	-7.44%	0.423
WY	Weyerhaeuser Corp.	18,147	21.93%	78.07%	6.68%	0.269
WYN	Wyndham Worldwide	4,129	52.35%	47.65%	79.30%	0.004
WYNN	Wynn Resorts Ltd	40,253	41.91%	58.09%	9.66%	0.157
XEC	Cimarex Energy	1,750	83.98%	16.02%	46.70%	0.267
XEL	Xcel Energy Inc	4,333	35.27%	64.73%	-4.39%	0.872
XL	XL Capital	18,965	4.69%	95.31%	2.32%	0.819
XLNX	Xilinx Inc	6,495	67.08%	32.92%	204.30%	0.000
XOM	Exxon Mobil Corp.	122,300	83.96%	16.04%	36.20%	0.000
XRAY	Dentsply Sirona	98,300	2.38%	97.62%	9.60%	0.061
XRX	Xerox Corp.	5,774	78.79%	21.21%	131.10%	0.017
XYL	Xylem Inc.	1,185	25.37%	74.63%	118.40%	0.061
YUM	Yum! Brands Inc	31,845	42.45%	57.55%	12.80%	0.014
ZBH	Zimmer Biomet Holdings	2,900	59.49%	40.51%	260.40%	0.000
ZION	Zions Bancorp	74,450	0.26%	99.74%	7.19%	0.387
ZTS	Zoetis	2,917	95.56%	4.44%	52.30%	0.003

Appendix D: Retrieving Google SVI (SVI) and Google Finance Investor SVI (FISVI)

This Appendix documents how to retrieve Google SVI (SVI) and Google "Finance-Investor" SVI (FISVI). Google search volume is obtained via www.trends.google.com. As shown below in Panel A, the search term can be entered (i.e. the relevant ticker), the region can be selected (typically set to the United States), the period can be selected, and a category can be selected with the default "All categories". When "All categories" is selected, Google SVI (SVI) is extracted. Google allows to refine SVI using "category", as they state this "show the search interest for that term in in a specific context". [29] We use the sub-category Finance-Investing [Category 107].[30] Panel B shows examples of retrieved SVI and FISVI for Apple. Inc [AAPL] and Starbuck Corp. [SBUX].

Panel A: Retrieving SVI and FISVI

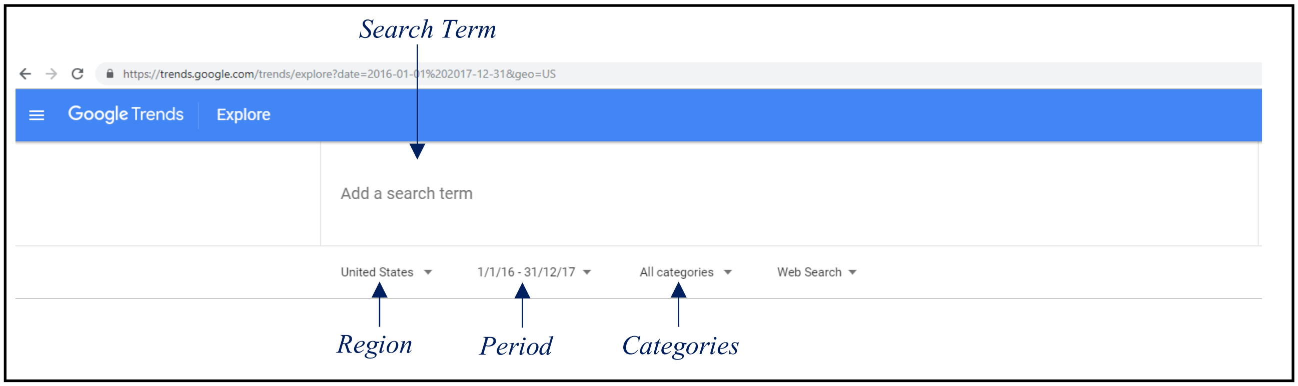
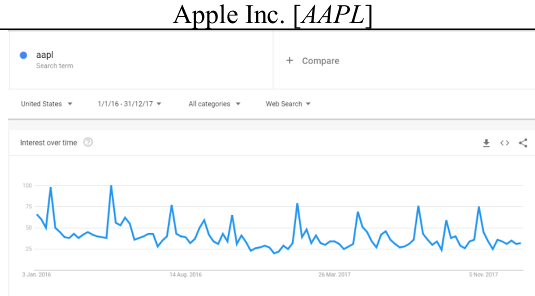
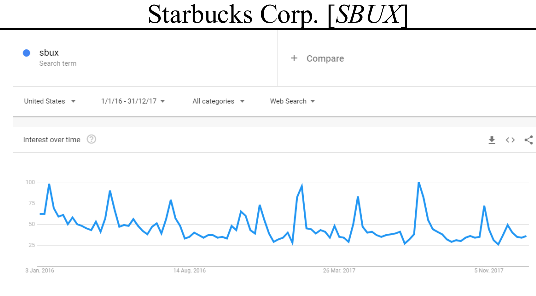
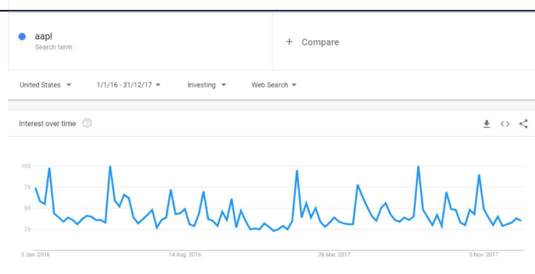
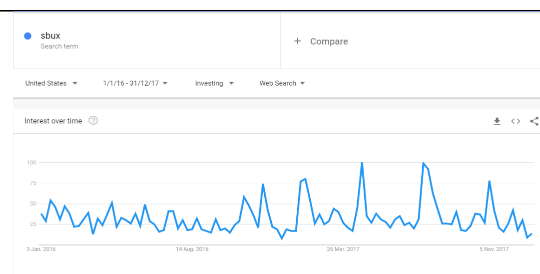


Figure 2: Examples of retrieved SVI and FISVI

	Apple Inc. [*AAPL*]	Starbucks Corp. [*SBUX*]
SVI		
FISVI		

[29] "**Refine Trends results by category.** If you're using Trends to search for a word that has multiple meanings, you can filter your results to a certain category to get data for the right version of the word. For example, if you search for "jaguar," you can add a category to indicate if you mean the animal or the car manufacturer." https://support.google.com/trends/answer/4359597?hl=en [accessed 10/23/2018].

[30] For a complete overview of categories see: https://github.com/pat310/google-trends-api/wiki/Google-Trends-Categories [accessed 10/23/2018].