Investor Inattention and the Market Impact of Summary Statistics*

Thomas Gilbert, Shimon Kogan, Lars Lochstoer, and Ataman Ozyildirim†

December 8, 2008

Abstract

Investors with limited attention have an incentive to focus on summary statistics rather than individual pieces of information. We use this observation to form a test of the impact of limited attention on the aggregate stock market. We examine the market response to a macroeconomic release that is purely a summary statistic, the U.S. Leading Economic Index (LEI). Consistent with the limited attention hypothesis, we show that the LEI announcement has an impact on aggregate stock returns, return volatility, and trading volume.

*We would like to thank The Conference Board for providing us with the data. We wish to thank Frank Tortorici and Ken Goldstein for their help, as well as Francisco Gomes, Rick Green, Terry Hendershot, Rich Lyons, Chris Malloy, Miguel Palacios, and Christophe Périgouillon for helpful comments. We thank seminar participants at EFMA 2006, Carnegie Mellon, U.C. Berkeley, and Penn State. The views expressed in this paper are those of the authors and do not necessarily represent those of The Conference Board. All errors remain ours.

†Affiliations: Foster School of Business at the University of Washington, McCombs School of Business at the University of Texas at Austin, Graduate School of Business at Columbia University, and The Conference Board, respectively. Corresponding author: Shimon Kogan, 1 University Station, Austin, TX-78712, USA. E-mail: shimon.kogan@mccombs.utexas.edu
1 Introduction

For most investors, the costs of processing all relevant information are prohibitively high. As such, these investors can be viewed as having limited attention. Previous studies provide evidence suggesting that investors’ limited attention is important for the pricing of individual securities. For example, DellaVigna and Pollet (2007) show that publicly available demographic information relating to future firm earnings is not completely impounded in stock prices. Hirshleifer, Lim, and Teoh (2007) suggest that investors do not process all the relevant information when many firms release their earnings on the same day, leading to weaker stock price reaction. Huberman and Regev (2001) document an instance where a re-release of news had a large effect on the stock price of a biotech firm. In this paper, we ask whether limited attention also has a systematic impact on the aggregate stock market.

To answer this question, we examine how the release of summary statistics affect the market. Barberis and Shleifer (2003) and Peng and Xiong (2006) highlight that investors with limited attention economize on information processing by grouping stocks into categories. In the same spirit, investors with limited attention may choose to focus on summary statistics instead of attending to every individual piece of information. This implication of limited attention has not, to our knowledge, been tested in the previous literature.

The summary statistic we identify is The Conference Board’s U.S. Leading Economic Index (LEI). This index is designed to track business cycle fluctuations and signal turning points in the business cycle. It is economically relevant as it contains information about future cash flows and discount rates, and it has a leading relationship relative to macroeconomic aggregates such as output and employment. Like most macroeconomic announcements, the LEI is released on a pre-determined schedule.

What makes the LEI an important variable in our context is that the components of the composite index are publicly available or can be easily calculated using publicly available data in advance of the announcement. In fact, several of the constituent macroeconomic variables are usually released weeks before the LEI scheduled releases. Furthermore, the methodology used to compile the LEI is also publicly available. These are well-known facts, publicized among other places on The Conference Board’s website and Bloomberg.\(^1\) It is therefore possible to calculate the change in the index before its release. Based on semi-strong market efficiency, the announcement of the LEI should not impact market returns

since the information in the component data relevant to fundamentals should already have been incorporated in stock prices.

Our null hypothesis is that there should be no market reaction to the announcements of the index since 1) the LEI is based on previously released data, 2) the components and methodology of the LEI are readily available to the public, and 3) the index is easy to reproduce. However, if limited attention is important for the aggregate stock market, the information in the LEI may be news to investors and as such have a market impact.

Looking at high-frequency intraday data over 72 announcement days over the period 1997-2005, we find that the release of the LEI is associated with measures of information arrival such as market returns, return volatility, and trading volume. The market return is positively related to the changes in the LEI: A one standard deviation increase in the LEI on average leads to a 4 basis-point increase in the subsequent five-minute market returns. Aggregate return volatility and trading volume increase by 25% and 7%, respectively, following the announcement. The volatility increase is significant for the five-minute interval following the announcement, while the volume increase persists for the subsequent 30 minutes.

The level of the return response is small in magnitude. However, this result is expected given that limited attention should affect markets less if the information is market-wide, as is the case here. Investors have a higher incentive to be attentive to information that is more important for their utility and, as a group, investors therefore focus more on market-wide information. Further, since the release is recurring and perfectly forecastable, one would expect attentive arbitrageurs to eliminate profit opportunities arising from the release up to transaction costs. The tests in this paper are therefore stacked against the alternative hypothesis of limited attention. Nevertheless, when we use quote and volume data from the Spider (SPDR) exchange-traded fund, we find that a simple trading strategy, which goes long (short) the market prior to positive (negative) LEI announcements, delivers on average positive daily volume-weighted profits.

The impact of investors with limited attention on the valuation of financial assets is important because its existence and causes have implications for the efficiency of financial markets. This paper is in that sense related to the large body of literature on market efficiency and more specifically to studies that evaluate the impact of news about fundamentals on asset prices. Examples of such studies include Schwert (1981), Huberman and Schwert (1985), Cutler, Poterba, and Summers (1989), and Mitchell and Mulherin (1994). Fama (1970, 1991) provides summaries of the market efficiency literature.

---

2 The Spider is an exchange-traded fund (ETF) designed to track the S&P500 Index.

market efficiency along the following dimensions. First, our test is on the aggregate market whereas much of the previous evidence of limited attention concerns company specific news events: for example, Huberman and Regev (2001) and Meschke (2004) who finds that stock prices and trading volume react to CNBC broadcasts of interviews with the relevant company’s CEO. Second, since the LEI announcement is recurring monthly at a known time, agents have the opportunity to learn about the nature of the announcement over time. However, we do not find evidence that the announcement effect is weaker in the latter part of the sample. Third, the results are unlikely to be explained by risk. As an example, Liu, Whited and Zhang (2007) propose a rational risk-based explanation for the earnings-announcement drift of Bernard and Thomas (1990) and Sloan (1996).

We do acknowledge that our experiment has two main drawbacks: namely that the magnitude of the announcement response and the sample size are relatively small. In a related study, Tetlock (2008) analyzes the market reaction of S&P500 stocks to news stories that may contain stale information. Using various proxies for staleness, such as the presence of another news story in the previous week, he finds that markets do overreact to stale news but that this overreaction is partially reversed in the subsequent week. Even though his set of information releases is much larger than ours (more than 350,000 stories compared to 72), his measures of staleness are approximate. In contrast, we know that the LEI contains only already-released information and therefore is 100% a stale summary statistic.

2 The Leading Economic Index

The Composite Index of Leading Economic Indicators (LEI), calculated and published monthly by The Conference Board (TCB), is designed to predict turning points (peaks and troughs) in the business cycle. TCB took over the responsibility to publish and maintain the LEI and the Business Cycle Indicators database from the Bureau of Economic Analysis starting with the December 6, 1995 release.

Leading indicators are series that have an established tendency to decline before recessions and rise before recoveries. For instance, new orders for machinery and equipment are placed well before investment plans are completed. By design, the LEI should help predict changes in real economic activity. Figure 1 shows that the LEI systematically declines ahead of the recessions dated by the NBER. Filardo (2004) provides evidence that the LEI performs well

---

4For more details on the indicator approach to measuring and analyzing business cycles, see Burns and Mitchell (1946) and Zarnowitz (1992).
as a variable to forecast cyclical movements in the economy. McGuckin, Ozyildirim, and Zarnowitz (2007) also report evidence on the significant out-of-sample forecasting ability of the LEI.

After TCB assumed responsibility for the Business Cycle Indicators program, it reviewed and revised the LEI in 1996. Notably, the composition of the LEI was changed: Two components were deleted due to their excessive volatility which led to “false signals” of recessions and a new component was added (Interest Rate Spread). After this major revision (first released December 30, 1996), TCB also started to publish the LEI press release at 10:00am EST to be consistent with its other economic data releases. Previously, the LEI releases were made at 8:30am, following the BEA schedule.
In the current indexing methodology, which changed very little since the 1960s when the U.S. Department of Commerce began publishing composite indexes, the volatility of each component is standardized before the component contributions are averaged together, using equal weights. This adjustment is made so that relatively more volatile series do not exert undue influence on the index (the standardization factors are updated every year in January and are available in the monthly press releases). The average contribution becomes the monthly change in the LEI. Using this monthly change, the index level is calculated recursively starting from a value of 100 in January 1959, and it is normalized to have an average value of 100 in 1996.

Seven of the ten indicators used every month in the LEI calculation are available at least 24 hours before each release. The monthly values of the three remaining components which are not available on the publication date are based on estimates by TCB. These components (Manufacturers’ New Orders for Consumer Goods and Materials, Manufacturers’ New Orders for Nondefense Capital Goods, and the Personal Consumption Expenditure Deflator used to get real Money Supply (M2)) are estimated using a simple AR(2) time series regression.\(^5\) The Appendix provides more background information and details on why this procedure was selected and how it was implemented by TCB. It is important to note that the procedures used in the LEI calculation are explained with extensive details on TCB’s website and in their manual (The Conference Board (2001)), which is publicly available.

3 Empirical Analysis

This section presents the impact of the LEI announcements on aggregate measures of information arrival: stock returns, return volatility, and trading volume. Our null hypothesis is that the LEI announcements have no effect on any of the three variables. We focus on intraday market activity for two reasons. First, previous research has shown that the effect of news on aggregate stock market prices are mainly manifested in intraday returns data (Andersen et al. (2003, 2007)). Second, focusing on intraday returns makes our study less sensitive to the presence of other news effects over the same day (including the time since the close of the prior day until market open at 9:30am) that we may not have captured in our econometric specification.

\(^5\)When the unavailable data becomes available in the next month, the index is revised.
3.1 Data

We combine three different data sources: macroeconomic news, intraday index prices and individual stock transactions. The LEI release dates and index series were provided by TCB. It is important to note that we use the original release series (subsequent revisions to macroeconomic data resulted in ex-post updates of the index). In our sample (1/1997 - 8/2005), the index is always reported at 10:00am.\footnote{Before 1997, the LEI was reported at 8:30am, which coincides with the reporting time for a number of other macroeconomic releases (Nonfarm Payroll, Housing Starts, PPI, CPI, among others). The move to the 10:00am announcement time reflected in part a desire to make the announcement during market open hours.}

The market returns data is constructed using S&P500 futures prices, while the trading strategy analysis and our measure of volume relies on data from the NYSE Trade and Quote (TAQ) database. The futures data was purchased from Price-Data.com and includes five-minute interval data on open, high, low, and close prices for each of the futures contacts traded between 1997 and 2005. For each date, we determine which of the multiple contracts available are “on-the-run” and construct the intraday return series for each day using prices from that day’s on-the-run contract. Since aggregate intraday volume data is not readily available, we construct it by gathering tick-by-tick data from TAQ for all firms that were in the S&P500 index on a given day. We add transactions across all firms for each five-minute interval to arrive at the market volume for that time period.

We use data from the Census Bureau, Bureau of Economic Analysis, Federal Reserve Board, National Association of Purchasing Managers, and The Conference Board to screen out all dates on which other macroeconomic announcements were released simultaneously at 10:00am. The specific announcement are New Home Sales, Factory Orders, Construction Spending, Business Inventories, Consumer Confidence Index, NAPM Index, and the Target Federal Funds Rate.\footnote{These announcements were identified by Andersen et al. (2007) as having a significant impact on S&P500 futures returns.} Therefore, out of a total of 104 announcements in our sample (1/1997 - 8/2005), we exclude 30 due to the simultaneous macroeconomic releases and 2 due to unavailable futures prices at the time of the announcements.\footnote{We present results including all announcements days in Section 3.6.}

3.2 General Methodology

We evaluate the aggregate return, volatility and volume of announcement days over different five-minute intervals around the time of announcement to investigate any pre- and post-
announcement effects. The LEI release is at 10:00 throughout the sample, and we focus on the 9:30 to 10:30 interval.\textsuperscript{9} Andersen et al. (2007) note that looking at five-minute futures returns strikes a good balance between capturing fundamental dynamics operating at high-frequencies and minimizing the noise in returns caused by bid-ask bounce and other microstructure issues. The futures contracts on the S&P500 Index are extremely liquid, so empirically neither stale prices nor the bid-ask bounce should be important issues for our purposes.\textsuperscript{10} Further, this approach allows us to compare our results with those obtained in similar studies. For all five-minute time intervals between 9:30 and 10:30, we test whether there is a return response at the time of the announcement.

A concern when evaluating high-frequency data is the well-known presence of intraday patterns in volatility and volume (Admati and Pfleiderer (1988)). Rather than attempt a parametric description of such patterns, for which at present there is no agreed upon model, we test whether realized volatility and volume are different on LEI announcement days compared to a matched sample of non-announcement days. Specifically, we designate the day one week after each announcement date as the non-announcement date, unless there is another important macro news release on that date, in which case we pick the date following the LEI release.\textsuperscript{11}

### 3.3 Returns Analysis

In this section, we investigate the effect of the LEI announcement on S&P500 futures’ returns. We first define the shocks to the change in the LEI index, $\Delta \text{LEI}_t$, as

$$\Delta \text{LEI}_t \equiv \Delta \text{LEI}_\text{index}_t - E[\Delta \text{LEI}_\text{index}_t | I_t],$$

where $I_t$ denotes the information set of the inattentive investors. Since the information contained in the LEI index release is based on publicly available information, the shocks are always zero for an attentive investor: it is possible to perfectly replicate the change in the index ahead of time. Instead, to capture the information set of the inattentive investors, we use three naive time-$t$ proxies for the expected change in the LEI index:

\textsuperscript{9}In the remainder of the paper, we use the 24:00 time convention when quoting time intervals. Thus, 10:00 is 10:00am.
\textsuperscript{10}We verify that our results indeed are robust to the bid-ask bounce using trades and quotes data on the Spider S&P500 ETF (SPDR) in Section 3.3.3.
\textsuperscript{11}This allows us to control for both intraday and day-of-the-week effects.
1. Expected change is zero:

\[ E[\Delta \text{LEI}_{it}|I_t] = 0 \]  

(2)

2. Expected change is the historical average change up until time \( t \):

\[ E[\Delta \text{LEI}_{it}|I_t] = \frac{1}{t - 1} \sum_{i=0}^{t-1} \Delta \text{LEI}_{i} \]  

(3)

3. Expected change is the average change over last year:

\[ E[\Delta \text{LEI}_{it}|I_t] = \frac{1}{12} \sum_{i=t-12}^{t-1} \Delta \text{LEI}_{i} \]  

(4)

These measures are simple proxies for the actual expectations of inattentive investors. Using only variations of the sample mean to form expectations effectively holds the inattentive investors only to weak-form market efficiency. The functional forms chosen reflect simple rules for forming expectations. For instance, the moving average of previous price changes implies an extrapolative rule as in, for example, Murdock (1962). Regardless of the proxy, the null hypothesis is well-defined since all the measures are in a fully attentive investors’ information set at time \( t \).

We run the regression

\[ R_{it} = \alpha_i + \beta_i \Delta \text{LEI}_t + \epsilon_{i,t} \]  

(5)

where \( t \in [1, 2, ..., T = 72] \) and \( R_{it} \) is the intraday interval \( i \)'s log return on announcement day \( t \). Thus, if the interval \( i \) is before 10:00, the regressor is the same-day future percentage change in the LEI index, whereas if the interval \( i \) is after 10:00, the regressor is the same day’s already reported LEI change.

Table 1 presents the LEI regression coefficients, standard errors, and \( R^2 \)'s resulting from these regressions. At the announcement (interval 10:00-10:05), the announcement day regression coefficient is positive and significant at the 5% level in all the specifications. Thus, the LEI announcement appears to be moving aggregate stock prices in the direction of the change in the LEI index. The \( R^2 \) of the regressions at the announcement interval is 5.3\% and 8.5\%, where the moving average formulation of the expected change has the highest explanatory power. This result implies that a one standard deviation change in the LEI
Table I
Return Regressions

This table reports estimates from OLS regressions of five-minute S&P500 futures return on the same-day LEI announcement using three different methods to calculate the expectations of naive investors: zero, historical average, and moving average. There are 72 observations in each regression. Returns are multiplied by 100, standard errors are corrected for heteroskedasticity (White standard errors). * denotes significant at the 10 percent level, while ** denotes significant at the 5 percent level in a two-tailed test.

\[ R_{i,t_0-t_1} = \alpha_i + \beta_i \Delta LEI_t + \varepsilon_{i,t} \]

<table>
<thead>
<tr>
<th>Time</th>
<th>( E_t[\Delta LEI] = 0 )</th>
<th>( E_t[\Delta LEI] = Mean_{0,t-1}(\Delta LEI) )</th>
<th>( E_t[\Delta LEI] = Mean_{t-12,t-1}(\Delta LEI) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta ) (s.e.)</td>
<td>( R^2 )</td>
<td>( \beta ) (s.e.)</td>
</tr>
<tr>
<td>9:30 - 9:35</td>
<td>0.0497 (0.0442)</td>
<td>0.021</td>
<td>0.0509 (0.0447)</td>
</tr>
<tr>
<td>9:35 - 9:40</td>
<td>0.0708* (0.0367)</td>
<td>0.042</td>
<td>0.0729** (0.0359)</td>
</tr>
<tr>
<td>9:40 - 9:45</td>
<td>-0.0397 (0.0363)</td>
<td>0.016</td>
<td>-0.0386 (0.0362)</td>
</tr>
<tr>
<td>9:45 - 9:50</td>
<td>0.0231 (0.0462)</td>
<td>0.004</td>
<td>0.0179 (0.0459)</td>
</tr>
<tr>
<td>9:50 - 9:55</td>
<td>-0.0406 (0.0418)</td>
<td>0.014</td>
<td>-0.0417 (0.0418)</td>
</tr>
<tr>
<td>9:55 - 10:00</td>
<td>0.0104 (0.0456)</td>
<td>0.001</td>
<td>0.0101 (0.0459)</td>
</tr>
<tr>
<td>10:00 - 10:05</td>
<td>0.0968** (0.0446)</td>
<td>0.053</td>
<td>0.0957** (0.0442)</td>
</tr>
<tr>
<td>10:05 - 10:10</td>
<td>-0.0234 (0.0423)</td>
<td>0.004</td>
<td>-0.0239 (0.0418)</td>
</tr>
<tr>
<td>10:10 - 10:15</td>
<td>-0.0248 (0.0506)</td>
<td>0.005</td>
<td>-0.0257 (0.0501)</td>
</tr>
<tr>
<td>10:15 - 10:20</td>
<td>-0.0344 (0.0323)</td>
<td>0.011</td>
<td>-0.0304 (0.0317)</td>
</tr>
<tr>
<td>10:20 - 10:25</td>
<td>-0.0056 (0.0396)</td>
<td>0.000</td>
<td>-0.0084 (0.0386)</td>
</tr>
<tr>
<td>10:25 - 10:30</td>
<td>-0.0020 (0.0328)</td>
<td>0.000</td>
<td>-0.0042 (0.0324)</td>
</tr>
</tbody>
</table>
gives, on average, a 3.5bp to 4.5bp return response on the S&P500 futures. The regression coefficients for the five-minute intervals before the announcement (from 9:30-10:00) are on average positive, but not significant at the 5% level, with one exception: The coefficients for the 9:35-9:40 interval regression on announcement days are positive and marginally significant (p-value of 0.06 and 0.05). It is interesting to note that it is in this interval that TCB releases its announcement to the news agencies that report the release at 10:00. After the announcement interval, from 10:05 to 10:30, the regression coefficients tend to be negative, indicating a reversal. However, none of these are individually significant.

### 3.3.1 Return Reversal

To assess whether the price impact around the LEI announcement is permanent or transitory, we regress post-announcement returns from a telescoping return window, starting with the 10:05 to 10:10 return, then the 10:05 to 10:15 return, etc., until the 10:05 to 10:30 return, on the LEI announcement. Table 2 displays the results.

Under the transitory return response interpretation, the demand of inattentive investors exerts temporary price pressure which is partly reversed over the latter half of the event window. Under the permanent return response interpretation, the increased demand of inattentive investors has a lasting effect on prices as these investors update their expectations of economic conditions and thus their demand for risky assets. The regression coefficients are negative and increasing in absolute value and, in terms of magnitudes, there appears to be a reversal in the returns data. However, none of the regression coefficients are statistically significant. Thus, the data unfortunately cannot reliably tell us whether the return response is transitory or permanent, but only provide suggestive evidence that the return response is transitory. Tetlock (2008) provides evidence of return reversal after partially stale individual stock announcements, consistent with the sign of the regressions presented here.

### 3.3.2 Benchmarking Against Other Macroeconomic News Announcements

The return response documented in the previous section is small in magnitude. However, when benchmarked against other major macroeconomic news announcements such as Non-Farm Payroll and New Home Sales, the response is comparable. Figure 2 shows the return response of five-minute returns to the LEI announcement using the moving average expecta-

---

12 The standard deviation of predicted returns is equal to the square root of the regression’s $R^2$ multiplied by the historical sample standard deviation of returns in the five-minute interval.
Table II
Return Reversal Regressions

This table reports estimates from OLS regressions of the S&P500 futures returns of a telescoping return window on the same-day LEI announcement using three different methods to calculate the expectations of naive investors: zero, historical average, and moving average. There are 72 observations in each regression. Returns are multiplied by 100 and standard errors are corrected for heteroskedasticity (White standard errors). * denotes significant at the 10 percent level, while ** denotes significant at the 5 percent level in a two-tailed test.

\[ R_{t_0-t_1} = \alpha_i + \beta_i \Delta LEI_{t_1} + \varepsilon_{i,t} \]

<table>
<thead>
<tr>
<th>Time</th>
<th>( E_t[\Delta LEI] = 0 )</th>
<th>( E_t[\Delta LEI] = \text{Mean}_{0,t-1}(\Delta LEI) )</th>
<th>( E_t[\Delta LEI] = \text{Mean}_{t-12,t-1}(\Delta LEI) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_0 - t_1 )</td>
<td>( \beta )</td>
<td>( R^2 )</td>
<td>( \beta )</td>
</tr>
<tr>
<td>10:05 - 10:10</td>
<td>-0.0234 (0.0423)</td>
<td>0.004</td>
<td>-0.0239 (0.0418)</td>
</tr>
<tr>
<td>10:05 - 10:15</td>
<td>-0.0482 (0.0614)</td>
<td>0.009</td>
<td>-0.0496 (0.0607)</td>
</tr>
<tr>
<td>10:05 - 10:20</td>
<td>-0.0826 (0.0641)</td>
<td>0.018</td>
<td>-0.0800 (0.0633)</td>
</tr>
<tr>
<td>10:05 - 10:25</td>
<td>-0.0882 (0.0874)</td>
<td>0.015</td>
<td>-0.0885 (0.0862)</td>
</tr>
<tr>
<td>10:05 - 10:30</td>
<td>-0.0902 (0.0944)</td>
<td>0.011</td>
<td>-0.0927 (0.0925)</td>
</tr>
</tbody>
</table>

...next to the statistically significant five-minute return responses to macroeconomic news announcements taken from Andersen et al. (2007).\(^{13}\) These authors run similar regressions as in our equation (5), also using S&P500 futures returns, but with certain differences in the data period and construction. Their sample period is different from ours, 1994 to 2002 versus 1997 to 2006, but with a six-year overlap. Further, Andersen et al. (2007) define macroeco-

\(^{13}\)Andersen et al. (2007) investigate all macroeconomic releases including the LEI, but do not find significant evidence that it has a price impact. We offer two explanations for the discrepancy between our results. First, in their Table 4, they state that the LEI is released at 8:30. This is true only for the beginning of their sample. In our sample, from 1997 and onwards, the release time is always at 10:00. At present we do not know if the authors corrected the change in release time over the sample, but they give no indication in their paper that they do. Second, and more fundamentally, they investigate normalized “surprises” based on market estimates obtain from a survey database (MMS). As we discuss above, it is unclear what these “shocks” represent since the index is perfectly forecastable.
nomic news surprises by subtracting the conditional expectations of professional forecasters, obtained from survey data, from the actual announcement. In our case, subtracting a measure of expert expectations is not appropriate since under the null hypothesis, the expected value is equal to the announced value, so all “news” values would be zero. Instead, we apply three naive proxies for inattentive investor expectations as discussed earlier. It is important to note that the figure reports the average return response given a one standard deviation change in the respective independent variables. This makes the numbers comparable across macroeconomic variables with the caveats that the sample periods and the construction of the variables somewhat differ.
3.3.3 Trading Strategies

The results using the S&P500 futures returns are based on transaction prices. This raises two questions. First, is the return response due to a bid-ask bounce? In this case, a high LEI leads to buys at the ask, while a low LEI leads to sells at the bid, without the midpoint of the bid and the ask prices actually changing. Second, even if midpoint prices are moving in the direction of the LEI, which indicates failure of market efficiency in the Grossman and Stiglitz (1980) sense, is the market response to the LEI within the bid-ask spread and thus could not have been traded on?

Unfortunately, quotes data for the S&P500 futures for our sample period is not publicly available. Therefore, we turn the analysis to the Spider (SPDR): an ETF that tracks the S&P500 Index (ticker symbol SPY). The price of the Spider is 1/10 of the index level and ranges from $80 to $153 in the sample period. Intraday trades and quotes on the Spider can be obtained from the NYSE Trade and Quote (TAQ) database, which we clean using standard filters.\(^{14}\) This data allows us to run return regressions based on index prices determined as the midpoint of the bid-ask spread, as well as evaluate trading strategies taking into account bid-ask spreads. The trading strategies are only suggestive as we cannot know if the prescribed trades would have moved the subsequent bid and ask prices had they been executed. We also ignore any fees, such as brokerage fees, incurred when trading.

**Mid-Point vs. Bid - Ask.** Table 3 shows the regression of midpoint returns on the Spider on the normalized LEI announcement. As for the S&P500 futures return, the 10:00 - 10:05 return regression gives a positive and statistically significant (at the 5% level) regression coefficient. The coefficients are insignificant in the other five-minute intervals. Thus, the results using prices based on the midpoint confirm the finding in Table 1 for the S&P500 futures return. The \(R^2\)'s are also of a similar magnitude. We conclude from this that the regression results are not due to a bid-ask bounce but are instead the consequence of fundamental price changes.

**Trading Strategies.** Next, we ask whether trading strategies based on the LEI release
Table III
Spider Midpoint Return Regressions

This table reports estimates from OLS regressions of the Spider five-minute returns on the same-day LEI announcement using three different methods to calculate the expectations of naive investors: zero, historical average, and moving average. The number of observations in each regression varies between 63 and 69, as there are a few instances early in the sample where the Spider does not trade in the relevant five-minute interval. Returns are multiplied by 100 and standard errors are corrected for heteroskedasticity (White standard errors). * denotes significant at the 10 percent level, while ** denotes significant at the 5% level in a two-tailed test.

\[ R_{t_0-t_1}^{SPY} = \alpha_i + \beta_i \Delta LEI_t + \varepsilon_{t_i} \]

<table>
<thead>
<tr>
<th>Time</th>
<th>( E_t[\Delta LEI] = 0 )</th>
<th>( E_t[\Delta LEI] = Mean_0,t-1(\Delta LEI) )</th>
<th>( E_t[\Delta LEI] = Mean_{t-12,t-1}(\Delta LEI) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_0 - t_1 )</td>
<td>( \beta ) (s.e.)</td>
<td>( R^2 )</td>
<td>( \beta ) (s.e.)</td>
</tr>
<tr>
<td>9:50 - 9:55</td>
<td>-0.0368 (0.0379)</td>
<td>0.013</td>
<td>-0.0384 (0.0380)</td>
</tr>
<tr>
<td>9:55 - 10:00</td>
<td>0.0319 (0.0552)</td>
<td>0.006</td>
<td>0.0289 (0.0557)</td>
</tr>
<tr>
<td>10:00 - 10:05</td>
<td>0.0909** (0.0465)</td>
<td>0.047</td>
<td>0.0923** (0.0462)</td>
</tr>
<tr>
<td>10:05 - 10:10</td>
<td>0.0497 (0.0894)</td>
<td>0.005</td>
<td>0.0545 (0.0878)</td>
</tr>
<tr>
<td>10:10 - 10:15</td>
<td>0.0462 (0.0868)</td>
<td>0.004</td>
<td>0.0357 (0.0865)</td>
</tr>
</tbody>
</table>

can be profitable after accounting for transaction costs as implied by the bid-ask spread. Panel A of Table 4 shows the average nominal bid-ask spread and the average five-minute volume for the Spider in the first hour of market open as well as for the interval 9:55-10:05 for each year in the sample. The average nominal bid-ask spread decreases strongly between 2001 and 2003, while volume increases. In the first part of the sample, the average bid-ask spread is roughly 20 basis points of the price, whereas in the second half of the sample the average bid-ask spread is about 3 basis points of the price. The bid-ask spread will naturally affect the profitability of a trading strategy, and we therefore focus on both the full sample, 1997-2005, and the second half, 2002-2005.

We first consider a return-based measure of the profitability of trading on the LEI release.
Table IV
Spider Trading Strategies

Panel A reports the average nominal bid-ask spread and average five-minute volume for each year in the sample for the given intraday time period. The price of the Spider ranged between $80 and $153 in this period so, to give an example, the average bid-ask spread in 2005 was about 2.5 basis points. Panel B shows the average return to the trading strategy based on bid and ask prices, while Panel C shows the average profits of the volume-weighted trading strategy. The trading strategies are conducted both over the full sample and over the sub-period 2002-2005 (after 2001 there was a large drop in the bid-ask spreads which increases the profitability of the trading strategies).

### Panel A: Bid-Ask Spreads and 5-Minute Average Volume

<table>
<thead>
<tr>
<th>Year</th>
<th>Bid-Ask</th>
<th>Volume</th>
<th>Year</th>
<th>Bid-Ask</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>0.182</td>
<td>63,431</td>
<td>1998</td>
<td>0.170</td>
<td>106,110</td>
</tr>
<tr>
<td>1999</td>
<td>0.324</td>
<td>88,704</td>
<td>2000</td>
<td>0.213</td>
<td>120,999</td>
</tr>
<tr>
<td>2001</td>
<td>0.216</td>
<td>278,935</td>
<td>2002</td>
<td>0.048</td>
<td>468,839</td>
</tr>
<tr>
<td>2003</td>
<td>0.042</td>
<td>549,987</td>
<td>2004</td>
<td>0.042</td>
<td>327,091</td>
</tr>
<tr>
<td>2005</td>
<td>0.028</td>
<td>541,598</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Panel B: Trading Strategy: Returns Net of Bid-Ask Spread

<table>
<thead>
<tr>
<th>Start time</th>
<th>Average Return 1997 - 2005</th>
<th>Average Return 2002 - 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:30</td>
<td>-0.0605%</td>
<td>0.0504%</td>
</tr>
<tr>
<td>9:35</td>
<td>-0.0490%</td>
<td>0.0449%</td>
</tr>
<tr>
<td>9:40</td>
<td>-0.0507%</td>
<td>0.0478%</td>
</tr>
<tr>
<td>9:45</td>
<td>-0.0514%</td>
<td>0.0402%</td>
</tr>
<tr>
<td>9:50</td>
<td>-0.0511%</td>
<td>0.0288%</td>
</tr>
<tr>
<td>9:55</td>
<td>-0.0544%</td>
<td>0.0219%</td>
</tr>
</tbody>
</table>

### Panel C: Trading Strategy: $ Profit Net of Bid-Ask Spread

<table>
<thead>
<tr>
<th>Start time</th>
<th>Average Profit 1997 - 2005</th>
<th>Average Profit 2002 - 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:30</td>
<td>$44,062</td>
<td>$157,071</td>
</tr>
<tr>
<td>9:35</td>
<td>$35,419</td>
<td>$115,988</td>
</tr>
<tr>
<td>9:40</td>
<td>$31,334</td>
<td>$89,076</td>
</tr>
<tr>
<td>9:45</td>
<td>$15,399</td>
<td>$54,569</td>
</tr>
<tr>
<td>9:50</td>
<td>$3,595</td>
<td>$21,167</td>
</tr>
<tr>
<td>9:55</td>
<td>$-2,037</td>
<td>$3,853</td>
</tr>
</tbody>
</table>
To ensure robustness of the results, we consider a range of trading strategies constructed as follows. We start trading at a given time before 10:00 and trade an equal dollar amount every minute up until, but not including, 10:00. We then hold this portfolio until 10:05, when we unload the entire position. The trade is in the direction of the LEI change: If the release is above a certain threshold described below, we start buying the Spider before the announcement and sell at 10:05, and vice versa. Buys are made at the then-available ask, while sells are made at the bid. To avoid using future information, we let the cut-off value of the LEI change, which determines the threshold between buys and sells, be equal to its long-run historical median using data up until 1997 (which is 0.1). Note that this cut-off point leads to an equal number of long versus short days. Thus, the returns to the trading strategies are not due to an unconditional long or short bias.

Panel B of Table 4 shows the average returns from this strategy. As is apparent, the average announcement day return for the full sample is negative and about -5 basis points. However, for the latter part of the sample, the returns become positive: about 5 basis points if trading commences between 9:30 and 9:45, and decreasing afterwards. The change in trading strategy returns between the two samples is likely to be primarily driven by the change in bid-ask spreads, which are much higher in the first half of the sample. Given that the return response is fairly low, high bid-ask spreads lead to negative returns. Further, the bid-ask spread closer to 10:00 is slightly higher than the average for the hour, which partly explains the lower returns for the trading strategies that start at 9:50 and 9:55.

Panel C of Table 4 shows the average announcement-day dollar profits from a volume-weighted trading strategy using the available bid and ask prices. Instead of trading an equal amount in the interval from the given start time until, but not including, 10:00, we now trade a number of Spider shares equal to half the volume in a given minute, which we assume is the number of buys or sells in each minute. At 10:05, we begin unloading the shares, again according to half the volume in each minute thereafter, until we hold no more shares. If we still have shares left at 10:29, the remainder of the position is assumed to be liquidated at 10:29 prices. While this trading strategy gives the maximal amount of trading that would have been possible, it is easy to scale: If trading only 1/5 of half of the volume seems more reasonable, one can simply divide the dollar profits we report by five. For the full sample, the average announcement day trading profit is $44,062 if trading starts at market open. If instead trading starts at 9:55, the strategy leads to an average daily loss of $2,037. While the latter loss seems consistent with the reported negative return from the full sample reported in Panel B, the gain if one starts trading at the open does not. However, higher volume
is accompanied by lower bid-ask spread in the data. Thus, the volume-weighted trading strategy in effect takes advantage of periods with lower bid-ask spreads both within the hour and, more importantly for these results, across time. The results for the 2002-2005 period confirms this: Volume is higher, bid-ask spreads are lower and the dollar profits are consistently positive, ranging from $157,071 per announcement day if trading commences at market open, to $3,853 if trading starts at 9:55.

Thus, using historical data on available bid and ask prices for the Spider ETF, we show that the market response to the LEI release gives rise to profitable hypothetical trading strategies, especially in the latter half of the sample when the bid-ask spreads were markedly lower. However, we do not claim that these trades were in fact feasible as that would require estimating the market impact of additional trades. At the same time, this analysis excludes parallel trades that an arbitrageur may have also been able to execute profitably in other instruments, such as the S&P500 futures. Further, the trading strategy is not “optimal” in that we assume that the hypothetical trader does not consider the size of the bid-ask spread when implementing the trades. A complete evaluation would therefore have to take into account all assets that respond to the announcement as well as estimates of their trading costs. Nevertheless, the results in Table 4 suggest that the magnitude of the market response to the LEI release is economically significant.

3.4 Volatility Analysis

In this section, we test whether five-minute stock return volatility is higher on announcement days compared to non-announcement days in each of the five-minute intervals in the hour around the announcement. If the LEI announcement constitutes news to a set of investors, we should expect volatility to be higher at the announcement interval on announcement days compared to the non-announcement days. As was discussed earlier, we designate the day one week after each announcement date as the non-announcement date, unless there is another important macro news release on that date, in which case we pick the date following the LEI release.\textsuperscript{15}

It is well-documented that aggregate stock return volatility is time-varying. To control for this, we employ a matching study. First, we calculate the volatility of five-minute returns for each non-announcement day for the relevant trading hour. Next, we divide the five-minute returns on both the corresponding announcement day and the non-announcement

\textsuperscript{15}We experimented with other matching rules, without any qualitative changes in the results.
day by this volatility measure.\textsuperscript{16} We use only the non-announcement days’ volatility in order to capture any overall higher levels of volatility on announcement days in the subsequent volatility tests. This normalization is valid under the null hypothesis that the volatility over matched announcement and non-announcement days are equal.

Next, we calculate the volatility of five-minute (normalized) announcement and non-announcement day returns for each interval as follows. Let \( \tilde{R}_{i,t} \) be the normalized five-minute log return for the interval \( i \), where \( i \in \{9:30-9:35, 9:35-9:40, ..., 10:25-10:30\} \). Interval \( i \)'s variance estimate is then

\[
\hat{\sigma}^2_i = \frac{1}{T} \sum_{t=1}^{T} \tilde{R}_{i,t}^2
\]

where the subscript \( t \) corresponds to the announcement or non-announcement days in our sample, which are indexed 1 to 72. To test whether the variance on announcement days is different than on non-announcement days, we apply a Levene F-test for each interval \( i \).\textsuperscript{17}

Column 2 in Table 5 shows the results. The ratios of announcement vs. non-announcement days’ volatility exhibit a significant spike for the interval 10:00-10:05, which corresponds to the time the LEI index is announced. The increase is not only statistically significant (at the 5% level), but also economically sizable – volatility increases by an average of 25%. Before 10:00, there appears to be no overall pattern in the volatility ratio: volatility is about the same on announcement vs. non-announcement days. There is one statistically significant observation at 9:35-9:40 for which announcement days seem to have lower volatility than non-announcement days. After 10:00, the volatility ratios are all above 1, indicating that volatility is overall higher on announcement days in the half hour following the LEI release.

### 3.5 Volume Analysis

In this section, we test whether volume is higher on announcement days compared to non-announcement days in each of the five-minute intervals in the hour around the announcement. To control for the strong increase in aggregate volume over the sample period and the

\textsuperscript{16}We calculate standard deviations assuming the expected five-minute returns are equal to zero. This is a standard assumption given the short time-interval and yields more robust volatility estimates. Using the residuals of a regression of intraday returns on their lagged value (to capture any bid-ask bounce, which we do not find significant) does not produce qualitatively different results.

\textsuperscript{17}It is common in empirical work to use modified Levene F-tests (for example the Brown-Forsythe modified Levene-test), as these are generally more robust to departures from normality of returns. We assume that the expected five-minute return is equal to zero, which is neither the sample mean, median nor the 10% trimmed mean, but which empirically turns out to be very close to the median.
Table V
Return Volatility and Volume Ratios
This table reports estimates of standard deviation of normalized five-minute returns and five-
minute volume on announcement and non-announcement days. There are 72 observations
in each group. The variance ratio test is a Levene F-test, where zero is assumed to be the
median/mean return. The volume ratio is regressed on a constant. The null hypothesis
is $\alpha = 1$. Standard errors are corrected for heteroskedasticity (White standard errors). *
denotes significance at the 10 percent level, while ** denotes significance at the 5 percent
level in a two-tailed test.

<table>
<thead>
<tr>
<th>$t_0 - t_1$</th>
<th>Volatility Ratio</th>
<th>Volume Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>($p - value$)</td>
<td>($p - value$)</td>
</tr>
<tr>
<td>9:30-9:35</td>
<td>1.186 (0.23)</td>
<td>0.978 (0.46)</td>
</tr>
<tr>
<td>9:35-9:40</td>
<td>0.785** (0.02)</td>
<td>1.016 (0.64)</td>
</tr>
<tr>
<td>9:40-9:45</td>
<td>0.934 (0.72)</td>
<td>1.049 (0.14)</td>
</tr>
<tr>
<td>9:45-9:50</td>
<td>1.146 (0.41)</td>
<td>1.031 (0.31)</td>
</tr>
<tr>
<td>9:50-9:55</td>
<td>0.963 (0.90)</td>
<td>1.048 (0.10)*</td>
</tr>
<tr>
<td>9:55-10:00</td>
<td>1.179 (0.27)</td>
<td>1.046 (0.19)</td>
</tr>
<tr>
<td>10:00-10:05</td>
<td>1.252** (0.05)</td>
<td>1.068** (0.04)</td>
</tr>
<tr>
<td>10:05-10:10</td>
<td>1.133 (0.46)</td>
<td>1.067** (0.03)</td>
</tr>
<tr>
<td>10:10-10:15</td>
<td>1.037 (0.93)</td>
<td>1.070** (0.03)</td>
</tr>
<tr>
<td>10:15-10:20</td>
<td>1.326** (0.01)</td>
<td>1.071** (0.04)</td>
</tr>
<tr>
<td>10:20-10:25</td>
<td>1.017 (0.98)</td>
<td>1.071** (0.02)</td>
</tr>
<tr>
<td>10:25-10:30</td>
<td>1.240 (0.10)</td>
<td>1.099** (0.00)</td>
</tr>
</tbody>
</table>
well known intraday patterns in volume, we create normalized five-minute volume for each announcement day, $v_{i,t}$, by dividing the volume of the same five-minute interval on the matched non-announcement day:

$$v_{i,t} = 1 + \ln\left(\frac{\text{volume in 5-minute interval } i \text{ on announcement day } t}{\text{volume in 5-minute interval } i \text{ on non-announcement day } t'}\right)$$  \hspace{1cm} (7)

We then regress this normalized volume on a constant for each five-minute interval from market open at 9:30 until 10:30:

$$v_{i,t} = \alpha_i + \varepsilon_{i,t} \text{ where } i \in \{9:30-9:35, 9:35-9:40, ..., 10:25-10:30\}$$  \hspace{1cm} (8)

The null hypothesis we are testing is $\alpha = 1$ and column 3 of Table 5 reports the results. The volume at market open on non-announcement days is slightly higher than on non-announcement days, but the difference is insignificant. However, as we get closer to the 10:00 announcement, the volume ratio becomes larger than unity. It is larger than unity and statistically significant following the announcement, as it was for both returns and volatility. The volume effect, however, persists significantly throughout the half hour following the announcement.

### 3.6 Robustness Checks

To examine the robustness of the results, we repeat the same analysis on the full set of announcement days. Recall that we initially excluded days on which other macro announcements were released. The full sample consists of 102 announcement days in the sample period January 1997 to August 2005. If the full sample is affected by the presence of other macro announcements, we would expect to find a weaker return response but a stronger volatility and volume responses to the LEI announcements.

Overall, we find that the LEI announcement has very similar impact on aggregate stock market returns, volatility and volume in the full sample compared to the filtered sample. Specifically, we find that the return response at the time of the announcement (10:00-10:05) is positively related to the change in the normalized LEI, although the coefficient is smaller (0.028 compared with 0.035 in the sub-sample) and its statistical significance is weaker ($t$-statistic of 1.74). This is to be expected as the other announcements are imperfectly correlated with changes in the LEI index and therefore add noise.

In contrast to the return tests which depend on the sign of the announcement, volatility
and volume are affected by the presence of other announcements. Indeed, we find that in the full sample, volatility increases by 28% following the announcement, which is stronger than for the filtered sample (25%) where the LEI is the only announcement that occurs. The same applies to trading volume. It increases by 8.8% following the announcement for the full sample, relative to 6.8% in the filtered sample.

4 Conclusion

In this paper we present evidence that investors act on summary information, impacting aggregate stock market returns, volatility and volume. The paper uses a weak restriction on aggregate prices to test for the presence of limited investor attention: Markets should not respond to the release of summary statistics that are based on information already available. We identify a unique stream of events, the U.S. Leading Economic Index (LEI), that is released on an monthly basis at pre-determined times, consists of previously published macro data, is calculated using a publicly available methodology, and is widely followed by the mass media.

We show that the LEI has statistically significant effects on proxies for information arrival such as instantaneous market-level returns (which move in the direction of the announcement), return volatility, and trading volume. The effect is short-lived for returns and volatility, consistent with previous studies of the impact of news announcements on aggregate prices. Volume, however, exhibits a more prolonged reaction. These findings are consistent with the presence of investors with limited attention, who find it optimal to focus on summary statistics as opposed to the individual components of the index.

Since the test pertains to aggregate information, the effects of limited attention on returns should be constrained by information gathering incentives and arbitrageurs. Peng and Xiong (2006) show that investors with limited attention choose to expend a larger fraction of their resources on aggregate information, such as the index constituents of the LEI. Further, since the LEI is released every month at a pre-determined time and date, it is relatively easy for arbitrageurs to profit from a market return reaction to its release. Therefore, if investors suffer from limited attention, the market return impact should be small, which is what we find. However, the other proxies for the arrival of information, volatility and volume, increase substantially following the announcement. The evidence provided in this paper therefore indicates that summary statistics like the LEI are valuable to a significant fraction of investors, even though the information provided is technically stale. More broadly, this
suggests a role for other summaries of information such as the ones provided by the financial press.
References


5 Appendix: LEI Calculation

Let $\Delta \text{LEI}_{t,t-1}$ denote the monthly change in the LEI for month $t - 1$ published in month $t$. This monthly change is calculated as the sum of component contributions which are derived from a symmetric percentage change formula:

$$
\Delta \text{LEI}_{t,t-1} = \left( \sum_{i=1}^{10} \sigma_i \times 200 \times \frac{X_{i,t} - X_{i,t-1}}{X_{i,t} + X_{i,t-1}} \right)
$$

where $\sigma_i$ is the standardization factor calculated by dividing the inverse standard deviation of component $i$ by the sum of the inverse standard deviations over all components. As the notation makes clear, the index published in month $t$ refers to past data for $t - 1$ which has already been published.

Since January 2001, leading indicator components for month $t - 1$ that are not available at the time of publication, month $t$, are estimated by TCB using a univariate autoregressive model to forecast each unavailable component. This procedure seeks to address the problem of varying availability in its components (publication lags). Without it, the index would contain incomplete components or it would not be available promptly under the current schedule.

In the publication schedule prior to January 2001, the index published in month $t$ referred to the month $t - 2$. In the new schedule after January 2001, the index published in month $t$ refers to the preceding month $t - 1$ (this information is available from The Conference Board). For example, in the old publication schedule the index would be calculated in the first week of March ($t$) for January ($t - 2$), and the January value of the LEI would use a complete set of components. According the new schedule, the index is calculated in the third week of March for February ($t - 1$), and the February value of the index uses 70% of the components which are already available and the remaining 30% are forecast. As seen in this example, users of the LEI would have had to wait for two more weeks until April for the February index.

The missing components (Manufacturers’ New Orders for Consumer Goods and Materials, Manufacturers’ New Orders for Nondefense Capital Goods, and the personal consumption expenditure used to deflate the Money Supply (M2)) are estimated using a time series regression that uses two lags (see McGuckin, Ozyildirim, and Zarnowitz (2001) for more on this model and a comparison with other alternative lag structures). The procedure used to estimate the current month’s Personal Consumption Expenditure Deflator (used in the
calculation of Real Money Supply and Commercial and Industrial Loans Outstanding) incor-
porates the current month’s Consumer Price Index when it is available before the release
of the LEI. When the unavailable data become available in the next month, the index is
revised.

The missing components could be forecast through alternative means. However, The
Conference Board has focused on simplicity, stability, and low costs of production and argues
for concentrating on easily implementable autoregressive model. Note that under the pre-
2001 release schedule of the LEI, it would have been possible to perfectly forecast the new
value each month just by collecting the individual data components and following the index
calculation methodology. In the post-2001 schedule, this is still possible, but the estimated
components require one additional step.