

Vakrman, Tomas; Kristoufek, Ladislav

Working Paper

Underpricing, underperformance and overreaction in initial public offerings: Evidence from investor attention using online searches

FinMaP-Working Paper, No. 35

Provided in Cooperation with:

Collaborative EU Project FinMaP - Financial Distortions and Macroeconomic Performance, Kiel University et al.

Suggested Citation: Vakrman, Tomas; Kristoufek, Ladislav (2015) : Underpricing, underperformance and overreaction in initial public offerings: Evidence from investor attention using online searches, FinMaP-Working Paper, No. 35, Kiel University, FinMaP - Financial Distortions and Macroeconomic Performance, Kiel

This Version is available at:

<https://hdl.handle.net/10419/108897>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

FinMaP-Working Paper No.35



This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 612955



FINMAP –

**FINANCIAL DISTORTIONS AND MACROECONOMIC
PERFORMANCE: EXPECTATIONS, CONSTRAINTS AND
INTERACTION OF AGENTS**

DATE: 03/23/2015

TITLE

Underpricing, underperformance and overreaction in initial public offerings: Evidence from investor attention using online searches

by: Tomas Vakra and Ladislav Kristoufek

ABSTRACT

Online activity of Internet users has proven very useful in modeling various phenomena across wide range of scientific disciplines. In our study, we focus on two stylized facts or puzzles surrounding the initial public offerings (IPOs) – underpricing and long-term underperformance. Using the Internet searches on Google, we proxy the investor attention before and during the day of the offering to show that the high attention IPOs have different characteristics than the low attention ones. After controlling for various effects, we show that investor attention still remains a strong component of the high initial returns (underpricing), primarily for the high sentiment periods. Moreover, we demonstrate that the investor attention partially explains overoptimistic market reaction and thus also a part of the long-term underperformance.

Keywords: initial public offerings, Google Trends, puzzles
JEL codes: E44, G02

AUTHORS

1. Tomas Vakrman

Institute of Information Theory and Automation,
Academy of Sciences of the Czech Republic
Pod Vodarenskou Vezi 4
CZ-18208 Prague

Email: tomas.vakrman@gmail.com

2. Ladislav Kristoufek

Institute of Information Theory and Automation,
Academy of Sciences of the Czech Republic
Pod Vodarenskou Vezi 4
CZ-18208 Prague

Institute of Economic Studies
Charles University
Opletalova 21
CZ-110 00 Prague

Warwick Business School,
University of Warwick, Coventry,
West Midlands, CV4 7AL, United
Kingdom, EU

Email: kristouf@utia.cas.cz

Underpricing, underperformance and overreaction in initial public offerings: Evidence from investor attention using online searches

Tomas Vagrman[†], Ladislav Kristoufek^{*,†,‡}

**Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodarenskou Vezi 4, 182 08, Prague, Czech Republic, EU*

†Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Opletalova 26, 110 00, Prague, Czech Republic, EU

‡Warwick Business School, University of Warwick, Coventry, West Midlands, CV4 7AL, United Kingdom, EU

Abstract

Online activity of Internet users has proven very useful in modeling various phenomena across wide range of scientific disciplines. In our study, we focus on two stylized facts or puzzles surrounding the initial public offerings (IPOs) – underpricing and long-term underperformance. Using the Internet searches on Google, we proxy the investor attention before and during the day of the offering to show that the high attention IPOs have different characteristics than the low attention ones. After controlling for various effects, we show that investor attention still remains a strong component of the high initial returns (underpricing), primarily for the high sentiment periods. Moreover, we demonstrate that the investor attention partially explains overoptimistic market reaction and thus also a part of the long-term underperformance.

Keywords: initial public offerings, Google Trends, puzzles

JEL codes: E44, G02

[☆]This article has been published as “Vagrman, T. & Kristoufek, L. (2015): Underpricing, underperformance and overreaction in initial public offerings: Evidence from investor attention using online searches, *SpringerPlus* 8:84” with the DOI reference 10.1186/s40064-015-0839-4.

Email address: kristouf@utia.cas.cz (Ladislav Kristoufek)

1. Introduction

The Internet, a revolutionary invention from 1965 with more than two billion users by 2014, has undoubtedly changed the world we live in. It allows its users to access an unprecedented amount of information in very short time. Due to abundance of available information, attention has become a scarce resource that needs to be efficiently allocated in order to acquire the information of interest. For a vast majority of Internet users, search engines serve as a gateway to all that information, and Google, with its dominant market share and more than one billion unique visitors every month, is the uncrowned king among them. Such online behavior leaves a digital trace. All individual search queries typed into search bar are stored by Google and processed statistics on searches are made publicly available by the company via its online facility Google Trends. Thus, the Google search volume data produces a direct measure of people's attention that is freely available, timely and representative to a whole population of Internet users.

Such an extreme potential of Internet search data has been put into practice and it is now used to track or even anticipate various social phenomena. The utilization ranges from influenza tracking (Ginsberg et al., 2008; Dugas et al., 2012), consumer interest and its impact on product sales (Choi and Varian, 2012; Goel et al., 2010; Kulkarni, 2012) to macroeconomic indicators (Cooper et al., 2005; Askitas and Zimmermann, 2009; Preis et al., 2010). The work of Merton Merton (1987) suggests that attention may be also relevant for complex reality of financial markets and Preis et al. (2008) are among the the first ones to prove this hypothesis right using attention proxied by web search data. Since then, many researchers used online attention to either track or forecast various financial indicators. Here, we focus on utilizing Google searches to help us explain two stylized facts of the initial public offerings (IPOs) – the long-term underperformance and the high initial returns, also known as the IPO underpricing.

The long-term underperformance (i.e. inferior performance to non-issuing firms) is arguably the most attractive area of the IPO academic research. Stern and Bornstein (1985) show that issuing firms under-perform the S&P 500 index by 22% in the long-term. The underperformance was confirmed by several studies (Ritter, 1991; Spiess and Affleck-Graves, 1995,?), most notably by Loughran and Ritter (1995) who labelled the long-term performance of newly issued stocks as a puzzle. Existence of the puzzle has been questioned by several studies. Brav et al. (2000) report that the underperformance disappears if the benchmarks are matched on firm size and book-to-market ratios. Conversely, Eckbo and Norli (2000) attribute the potential underperformance to a lower risk of IPO stocks, providing evidence that issuing companies have lower leverage ratios and higher liquidity than matched firms in years following the IPO. After controlling for additional risk of peer companies, the authors do not reject the hypothesis of zero abnormal returns of IPO stocks. Ritter and Welch (2002), in their comprehensive review of the IPO related literature, argue that benchmarking of long-term performance of IPOs is highly sensitive to employed methodology as well as to choice of a sample period. In addition, they note that despite the similar (unappealing) performance of issuers and their peers with comparable characteristics, the equally weighted post-IPO returns still underperform market indices.

The existence of the second IPO stylized fact – underpricing – is rather indisputable. Ritter and Welch (2002) report that between 1980 and 2001, the average difference between the offer price and the first day closing price was 18.8% for the US issuers. Furthermore, there was a positive price change for 70% of the issuing firms, while negative initial return was exhibited only by 14% of the IPOs. What remains unclear about the underpricing phenomenon is why the issuing firms would voluntarily leave money on the table. Ritter and Welch (2002) offer a wide variety of explanations based on both symmetric and asymmetric information arguments. The most promising stream of literature struggling to explain the underpricing seems to be focused on behavioral side of investors. Ritter (1991) sheds some light on the topic by pointing out that investors tend to be periodically overoptimistic about the potential of issuing firms, and that the firms take advantage of it by timing the offerings so they correspond with these periods. Loughran and Ritter (1995) provide some support to the hypothesis by showing the first day returns are significantly higher following the periods when the market has grown. In line with the investor sentiment theory, it has been shown that the underpricing is positively associated with news and non-lead analyst research coverage of IPOs (Demers and Lewellen, 2003; Aggarwal et al., 2002).

Ljungqvist et al. (2006) and Derrien Derrien (2005) offer theoretical models for the IPO pricing and initial returns in presence of investor sentiment. The former study (Ljungqvist et al., 2006) builds the model on assumption that sentiment investors are budget-constrained and cannot buy the entire IPO. Thus, in order to induce rational investors to participate, firms must set the offer price below the price that noise traders are willing to pay. The latter study (Derrien, 2005), on the other hand, stresses out the assumption that “*aftermarket price support is costly for the underwriter*”. (Derrien, 2005, p. 490). While the models are different in construction, their predictions are rather similar. They predict the high underpricing in presence of high investor sentiment and consequently the poor long-term performance. Derrien aptly notes that it is not the firms who leave the money on table but rather “*the overoptimistic noise traders who pay excessive prices for IPO shares on the aftermarket*” (Derrien, 2005, p. 490).

The empirical evidence favors these models. Cook et al. (2006) reveal that underwriters promote IPOs in order to induce the sentiment investors into the market for it. It is also reported that sentiment influences initial pricing and that underwriters do not solely base their valuation on fundamentals and comparable valuation. The most notable empirical validation of the sentiment theories are the higher initial returns of IPOs that exhibited an above average abnormal attention (measured by Google search volume), and subsequent return reversal of such stocks in the long-term (Da et al., 2011). Here, we focus on the IPOs and specifically the two stylized facts in the USA between 2004 and 2010. As a measure of attention, we utilize search queries provided by Google and we examine whether such attention can be used to explain and describe the IPO underpricing and long-term underperformance.

2. Methods

2.1. Data

2.1.1. Variables construction

Studying the two stylized facts about IPOs stems in defining two types of returns – an initial return and a long-term return. We define the initial return (which we also refer to as a first day return or we abbreviate it as IR) as

$$IR_i = \log(P_i^{Close}) - \log(P_i^{Offer}), \quad (1)$$

where P_i^{Close} and P_i^{Offer} refer to the closing price on the first day of trading and the offering price, respectively, for the IPO i . The long-term cumulative logarithmic return is defined as

$$CLR_i = \log(P_{i,t+k}^{Close}) - \log(P_{i,t}^{Close}), \quad (2)$$

where t either refers to closing price on the first day of trading or the closing price one month after IPO, and k is either 91, 183 or 366 days, depending on the used definition of the long-term. The two starting dates are considered to control for potential immediate drop in price after the first day of trading.

For the Google search volume (usually referred to as *GSV* in the literature), we utilize the daily statistics provided by the Google Trends database (trends.google.com). Google provides *GSV* as a normalized measure of online searches and as such, the value shows the changes in proportion of given searched term in the whole sample of searches rather than dynamics of the searches themselves. Again in correspondence to the standards in the literature, we utilize abnormal *GSV* usually labeled as *ASVI* (Abnormal Search Volume Index) which is defined as a logarithmic deviation of the actual logarithmic *GSV* from the logarithm of median *GSV* over a specific period. In our application, we use the median period of the last 26 trading days¹. Therefore, if we refer to *GSV* in the text, it represents the original Google search queries, and *ASVI* stands for the logarithmic deviation from the 26-day median value.

2.1.2. Dataset

We use the firm database of emerging growth IPOs Kenney and Patton (2013) for the identification of firms going public between years 2004 and 2010. The database contains a complete list of emerging growth firms going public at the US exchanges between 1990 and 2010. We limit ourselves to the period between 2004 and 2010 due to the Google searches data span, which starts in 2004. The complete list of variables can be found in respective guide written by the authors².

¹The median period of 26 trading days is chosen as it is close to a trading month and such choice delivers the best results. However, it needs to be noted that the results do not change qualitatively for the median periods between 20 and 30 trading days.

²Available at http://hcd.ucdavis.edu/faculty/webpages/kenney/misc/Firm_IPO_Database_Guide.pdf

The database excludes following types of firms and filings from the Thomson Financial Venture Expert, SDC data and other comprehensive lists of IPOs: mutual funds, real estate investment trusts (REITs), asset acquisition or blank check companies, foreign F-1 filers, and all spin-offs and other firms that are not true emerging growth firms (Da et al., 2011).

We use all the companies included in the Kenney-Patton database that went public between years 2004 and 2010, with the exception of the unit offerings and one firm that went public on the OTC (over the counter) market. Totally, it encompasses 547 companies. For the identification of relevant search queries, we follow the steps of Bank et al. (2011) and Vlastakis and Markellos (2012). The complete list of search terms is available from the authors upon request. Out of the 547 companies, the daily data were available only for 75 of them³. Using the daily rather than weekly data thus comes at a cost. However, the missing values are not much more frequent than in other studies (Da et al., 2011) considering the additional information value provided by higher frequency of the series.

The IPOs database (Kenney and Patton, 2013) does not contain data on the post-IPO performance. Therefore, the financial data on the first day closing prices come from *SCOOP Track Record from 2000 to Present* IPO database⁴, which has been controlled against data from Yahoo! Finance, Google Finance, NASDAQ web site database and IPO news coverage. For the long-term performance, the data availability is also poor as some of the companies have been already acquired, merged or delisted, and therefore do not anymore appear in the freely available databases. Thus, we utilize the Quantshare Trading Software⁵, or more specifically the *Historical EOD data Downloader for Delisted/Bankrupt Stocks* plug-in⁶ for such stocks. When possible, these have been again checked against the SCOOP Track Record database, Yahoo! Finance, Google Finance, NASDAQ web site and news coverage for comparison. The final IPO data set contains search volumes and stock prices for 75 firms, even though long-term cumulative returns are available only for 62 firms. Table 1 lists and describes all variables used in the computational sections for IPO data set.

2.2. Regression analysis

IPO regressions are all estimated by the cross-sectional ordinary least squares (OLS) procedure. We perform a widely applied methodology to test for the OLS assumptions. First, the presence of heteroskedasticity is tested by Breusch-Pagan (Breusch and Pagan, 1979) and White (White, 1980) tests. No severe heteroskedasticity is detected in the sample. However, if any of the tests suggest presence of mild heteroskedasticity, White's heteroskedasticity consistent standard errors are used (White, 1980). Second, the existence of multicollinearity is tested by variance inflation factors. Lastly, the normality of residuals

³Google Trends system allows to download daily series for a period of up to three months. For our given dataset, we have selected a three-month period covering the IPO date for each company.

⁴Available at https://www.iposcoop.com/index.php?option=com_trackrecord&Itemid=200.

⁵Available at <http://www.quantshare.com/>.

⁶Available at <http://www.quantshare.com/item-1270-historical-eod-data-downloader-for-delisted-bankrupt-stocks>.

is tested by Shapiro-Wilk test (Shapiro and Wilk, 1964). When the Shapiro-Wilk test suggests the residuals are non-normally distributed, we use bootstrapping (1000 replications) method to estimate the confidence intervals and p -values.

3. Results

We study 75 initial public offerings, which took place in the USA between 2004 and 2010, based on the Kenney-Patton database (Kenney and Patton, 2013). As a measure of investor attention, we utilize Google searches provided by the Google Trends database⁷. For more details about the dataset selection process and variable construction, please refer to the Methods/Data section. Basic descriptive statistics are provided in Table ???. The initial returns are on average positive, positively skewed and fat-tailed, strongly rejecting normality. The long-term returns show opposite statistics with a negative mean and longer left tail, again strongly rejecting normality. These findings are independent of the long-term return definition. We thus observe a reversal between initial and long-term returns, at least on average. More detailed examination is provided in the following text. The true discount is on average positive and the market reaction is very close to zero. And the offering size varies strongly across the examined IPOs.

To illustrate the importance and potential usefulness of the Google searches in the IPO setting, we start with the average dynamics of the Google Search Volume (GSV) before IPO takes place. Fig. 1 shows the average *GSV* for the studied 75 IPOs together with the 95% confidence intervals. Dynamics up to 30 days before IPO takes place is presented. We can see that the investor attention starts rising around 5 five days prior to IPO. This strongly justifies using daily data in the IPO analysis contrary to the standardly used weekly frequency. We now focus on two IPO stylized facts – high initial returns and long-term underperformance.

3.1. Initial returns

We analyze whether search volume may bring some information or predictive power regarding the IPO first day return, which is labelled as *IR* in the following text. The investor sentiment theory (Loughran and Ritter, 2002; Demers and Lewellen, 2003; Aggarwal et al., 2002) states that the these initial returns tend to be higher in periods of positive sentiment. Da et al. (2011) argue that investor sentiment and investor attention is closely related for retail investors as these are prone to sentiment while attention is a necessary condition for sentiment. Nonetheless, we measure the effect both of attention (firm specific) and of sentiment (market level) on the first day returns.

Before proceeding to the regression analysis, we examine how initial returns and investor attention are related on a basic level. Thus, we divide the firms from the sample into three groups based on their *ASVI* values (Abnormal Search Volume Index, see the Methods/Data section for more details) prior IPO – high, medium and low attention groups – based

⁷Freely available at trends.google.com. Google data are registered trademarks of Google Inc., used with permission.

on quantiles. The results show that the high attention group’s average initial return is 22.85%, while low attention group’s initial return only equals to 12.23%. The difference is statistically significant at 5%. Thus, the first look at the data suggests that investor attention, very likely, drives the first day returns up.

Relationship between the initial return IR and the investor attention $ASVI$ is estimated via the following model

$$IR_i = \beta_0 + \beta_1 ASVI_i + \sum_j \gamma_j CON_{j,i} + \varepsilon_i \quad (3)$$

in order to estimate how an increase in attention prior IPO influences the size of the initial return in more detail. CON represents a set of control variables, specifically offering size and investor sentiment (both in levels and a change to previous month). Table 2 provides the results. Column (1) shows that the steeper the increase in attention prior to an IPO is, the higher the corresponding initial returns are. The effect is highly significant and has a notable size – a standard deviation increment in $ASVI$ leads to an increase in initial return by a magnitude of 41.4% of its standard deviation.

Columns (2) to (9), which display the results of the robust-check regressions, suggest that neither the offering size nor the investor sentiment (both in levels and changes from the previous month level) are able to predict initial returns. The insignificance of the offering size variable is in contradiction with results of Da et al. (2011), who used IPO data set with 185 firms that went public between 2004 and 2007. Thus it seems that the offering size effect over the initial return largely depends on selected sample of firms as well as quality and availability of the Google data, which are increasing in time. The authors have also found the change in investor sentiment modestly significant (at 10% level), which is not significant in our results either.

To test the sentiment hypothesis, we construct dummy variables for positive, normal and negative values of sentiment and use them in the interaction with $ASVI$ in regressions (10) to (13) in Table 2. The results show that attention significantly increases initial returns only in positive sentiment periods. For the negative and normal sentiment times, attention boosts initial returns as well, albeit the effect is not significant. Nevertheless, the difference between the three coefficients in (13) is insignificant when tested by F -test. In addition, regressions (11) and (12) show that the results are robust if one controls for the original sentiment measures.

3.2. Long-term returns

We now approach the second stylized fact about IPOs – the long-term underpricing of the IPO firms compared to their already traded peers. The sentiment-based hypothesis regarding high first day returns works well with the subsequent long-term underperformance. The overoptimism of investors about the offering may lead to overly escalated initial returns, which should be followed by a price reversion towards the fundamental value afterward, i.e. the long-term underperformance (Ljungqvist et al., 2006; Ritter and Welch, 2002).

We consider five different horizons for long-term performance for which the cumulative log-returns are calculated: first day closing price to the (1) closing price one year, (2) half a year (3) and quarter of the year after IPO; and the closing price one month after IPO to (4) the closing price one year (5) and half a year after IPO. Such approach is used to avoid coincidental results based on randomly selected period marked as the long-term. Fig. 2 provides an overview of the cumulative returns over the five specified horizons for the low and high attention IPOs. It seems that, with an exception of the shortest horizon, the high attention IPOs clearly under-perform the low attention ones in long-term. Thus, the first results are in line with Da et al. (2011) findings and the attention/sentiment based theory on IPOs.

We proceed by regressing the long-term returns on the abnormal search volume on the IPO date. Table 3 compares the predictive power of *ASVI* over the long-term cumulative returns (LR) for the five defined periods. The results provide only weak evidence for *ASVI* ability to forecast the negative LR returns. For the half-year horizon (measured both from the opening day (2) and one month after IPO (5)), *ASVI* negatively correlates with the LR returns. Nevertheless, we see no significant effect on the one year (1, 4) or quarter of the year (3) cumulative returns regardless all coefficients being negative in sign.

Da *et al.* Da et al. (2011) construct an interaction variable between *ASVI* and initial return ($ASVI \times IR$) as the high initial return of IPOs that also experience increases in retail investor attention should be partly driven by the price pressure and hence revert in the long-term. We follow their procedure and regress the cumulative long-term returns on initial returns and the interaction variables. Table 4 shows that there is, as expected, a higher price reversion for IPOs that experienced high initial returns (1-5), albeit the effect is significant only for cumulative returns measured from one month after IPO. The performance of the interaction variable (5-10) matches the findings of Da et al. (2011) – it is obvious that high attention IPOs with high first day return experience severe price reversion in the long-term. The effect is significant for all considered horizons with the exception of the quarter of the year horizon measured from the offering day. It seems, and the results from the other regressions support this claim, that a quarter of the year horizon is too short for the prices to revert to long-term level.

We further employ the sentiment (dummy) interaction with *ASVI* to account for the effect of attention on the long-term returns in positive, medium and negative sentiment periods. Thus, we regress the long-term returns on *ASVI* in different sentiment periods. Results are provided in Table 5. Interestingly, only IPOs that went public in high sentiment periods and get abnormal attention show the price reversion in long-term. Nevertheless, also sentiment itself is able to predict the long-term reversal, albeit for fewer horizons and with lower significance.

3.3. Initial returns versus underpricing

Terms “initial return” and “underpricing” are usually used interchangeably. However, Ma and Tsai (2002) argue that under the sentiment hypothesis, the interchangeability is not correct. According to their definition, initial return has two components, true discount

(TD) and market reaction (MR) and it is split in the following way

$$IR = TD + MR = \frac{P_m - P_o}{P_o} = \frac{P_e - P_o}{P_o} + \frac{P_m - P_e}{P_o} \quad (4)$$

where P_m is the first day closing price, P_o is the offer price and P_e is the equilibrium (fundamental) market price. In the previous section, we have shown that the price revision and reversion for high attention IPOs happens approximately half a year after the offering. Moreover, if return variance is calculated for 30-day periods up to one year after IPO, the lowest variance corresponds to a horizon between 150 and 180 days after emission. Therefore, we use the average price between $t + 150$ and $t + 180$, where t is the IPO date, as an estimate for P_e . Note that any estimate of the fundamental price is rather arbitrary so that other definitions are indeed feasible.

According to Ma and Tsai (2002), positive values of MR mean that investors overreact, while negative values suggest under-reaction of investors. True discount, on the other hand, corresponds to actual underpricing. Thus, we use this setting to confirm the results that $ASVI$, especially if combined with positive sentiment on the market, drives the investor overreaction. In contrast, we expect that $ASVI$ should not possess any significant information about the underpricing term TD . To see whether such expectations may be valid, we calculate mean TD and MR for high and low attention IPOs. Fig. 3 displays the comparison. As expected, the true discount does not seem to be influenced by attention. Conversely, market reaction and attention devoted to IPO show strong interdependence.

The relationship is mainly confirmed by the regression results. We regress TD and MR on attention measured by $ASVI$, on the $ASVI$ interaction with the initial return, and on the attention-sentiment interaction variables. Results are presented in Table 6. On the one hand, it can be observed that no attention-based variable predicts the underpricing term. On the other hand, market seems to overreact on high attention IPOs, albeit the effect is significant only at 10%. The effect is more pronounced if we take into account the interaction with initial return, which is logical as MR is one of the two terms which the initial return consists of (thus the stronger is the evidence against $ASVI$ and TD interdependence, as the interaction term is insignificant in TD). Surprisingly, we see only insignificant effect of the sentiment interaction variables and market reaction. While the coefficient is positive for attention in positive sentiment periods, it is insignificant (albeit on the edge of 10% significance). Even more surprising is the positive coefficient for the attention in negative sentiment periods, as one would expect this term to be negative. It suggests that investors overreact to IPOs also in low sentiment period and that it is the attention that drives the overreaction and not sentiment. This is confirmed by regression (8), which shows that sentiment is not able to predict the market reaction on its own. The insignificance is indisputable in this case.

4. Discussion

We confirm that initial returns are higher for IPOs that receive above average attention. However, we argue that the effect is significantly present only for firms going public in the

positive sentiment periods. In addition, since the daily data are used, we are able to demonstrate that Google search volume is capable of forecasting the initial returns within a few days horizon.

Contrary to Da et al. (2011), we observe a weak evidence of Google data ability to forecast (with negative sign) the long-term cumulative returns. Nevertheless, in line with the authors, we show that high attention IPOs leaving a lot of money on the table experience a price reversal in long-term. In correspondence with the results for initial returns, the long-term cumulative returns seem to be inversely proportional to investor attention to IPO only for firms that emitted shares during the positive sentiment periods. The findings correspond to predictions Derrien (2005) claiming that it is the overoptimistic investors who leave the money on the table rather than the issuing firms.

Finally, we test Google search volume in the setting of the model proposed by Ma and Tsai (2002), which questions the interchangeability of terms initial return and underpricing. The results suggest that the Google search volume is able to predict one part of initial returns – the market overreaction to the offering –, while the other – the true IPO discount (i.e. the underpricing) – is unpredictable by Google data, which is in fact expected.

Acknowledgements

We want to thank Martin Kenney and Donald Patton for providing the IPO database (Kenney and Patton, 2013). Google data are registered trademarks of Google Inc., used with permission.

The research leading to these results has received funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreement No. FP7-SSH-612955 (FinMaP), the Czech Science Foundation project No. P402/12/G097 “DYME – Dynamic Models in Economics” and the Research Councils UK via Grant EP/K039830/1.

References

- Aggarwal, R. K., L. Krigman, and K. L. Womack (2002). Strategic ipo underpricing, information momentum, and lockup expiration selling. *Journal of Financial Economics* 66(1), 105–137.
- Askitas, N. and K. F. Zimmermann (2009, April). Google econometrics and unemployment forecasting. *Applied Economics Quarterly* 55(2), 107–120.
- Baker, M. and J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance* 61(4), 1645–1680.
- Bank, M., M. Larch, and G. Peter (2011, July). Google search volume and its influence on liquidity and returns of german stocks. *Financial Markets and Portfolio Management* 25(3), 239–264.
- Brav, A., C. Geczy, and P. A. Gompers (2000). Is the abnormal return following equity issuances anomalous? *Journal of Financial Economics* 56(2), 209–249.

- Breusch, T. S. and A. R. Pagan (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica* 47, 1287–1294.
- Choi, H. and H. Varian (2009/2012). Predicting the present with google trends. *Economic Record* 88, 2–9.
- Cook, D. O., R. Kieschnick, and R. A. Van Ness (2006, October). On the marketing of IPOs. *Journal of Financial Economics* 82(1), 35–61.
- Cooper, C. P., K. P. Mallon, S. Leadbetter, L. A. Pollack, and L. A. Peipins (2005, July). Cancer internet search activity on a major search engine, united states 2001-2003. *Journal of Medical Internet Research* 7(3), e36.
- Da, Z., J. Engelberg, and P. Gao (2011). In search of attention. *The Journal of Finance* 66(5), 1461–1499.
- Demers, E. and K. Lewellen (2003, June). The marketing role of IPOs: evidence from internet stocks. *Journal of Financial Economics* 68(3), 413–437.
- Derrien, F. (2005). Ipo pricing in hot market conditions: who leaves money on the table? *The Journal of Finance* 60(1), 487–521.
- Dugas, A. F., Y.-H. Hsieh, S. R. Levin, J. M. Pines, D. P. Mareiniss, A. Mohareb, C. A. Gaydos, T. M. Perl, and R. E. Rothman (2012, January). Google flu trends: Correlation with emergency department influenza rates and crowding metrics. *Clinical Infectious Diseases* 54(4), 463–469.
- Eckbo, B. E. and Ø. Norli (2000). Leverage, liquidity and long-run IPO returns. *BORANHH Discussion Paper* 29, 1–38.
- Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant (2008, November). Detecting influenza epidemics using search engine query data. *Nature* 457(7232), 1012–1014.
- Goel, S., J. M. Hofman, S. Lahaie, D. M. Pennock, and D. J. Watts (2010, September). Predicting consumer behavior with web search. *Proceedings of the National Academy of Sciences* 107(41), 17486–17490.
- Kenney, M. and D. Patton (2013). Firm database of emerging growth initial public offerings (IPOs) from 1990 through 2010.
- Kulkarni, G. (2012). Patterns of consumer interest across developed and emerging markets. *International Business & Economics Research Journal* 11(6), 603–610.
- Ljungqvist, A., V. Nanda, and R. Singh (2006). Hot markets, investor sentiment, and IPO pricing. *The Journal of Business* 79(4), 1667–1702.

- Loughran, T. and J. R. Ritter (1995). The new issues puzzle. *The Journal of Finance* 50(1), 23–51.
- Loughran, T. and J. R. Ritter (2002). Why don't issuers get upset about leaving money on the table in IPOs? *Review of Financial Studies* 15(2), 413–444.
- Ma, T. and P. Tsai (2002). Are initial return and IPO discount the same thing? a comparison of direct public offerings and underwritten IPOs. *EFA 2002 Berling Meetings Presented Paper*, 1–23.
- Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance* 42(3), 483–510.
- Preis, T., W. Paul, and J. J. Schneider (2008, June). Fluctuation patterns in high-frequency financial asset returns. *EPL (Europhysics Letters)* 82(6), 68005.
- Preis, T., D. Reith, and H. E. Stanley (2010, November). Complex dynamics of our economic life on different scales: insights from search engine query data. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368(1933), 5707–5719.
- Ritter, J. R. (1991). The long-run performance of initial public offerings. *The Journal of Finance* 46(1), 3–27.
- Ritter, J. R. and I. Welch (2002). A review of IPO activity, pricing, and allocations. *The Journal of Finance* 57(4), 1795–1828.
- Shapiro, S. S. and M. B. Wilk (1964). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591–611.
- Spiess, D. K. and J. Affleck-Graves (1995). Underperformance in long-run stock returns following seasoned equity offerings. *Journal of Financial Economics* 38(3), 243–267.
- Stern, R. L. and P. Bornstein (1985). Why new issues are lousy investments. *Forbes* 136, 152–190.
- Vlastakis, N. and R. N. Markellos (2012, June). Information demand and stock market volatility. *Journal of Banking & Finance* 36(6), 1808–1821.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48, 817–838.

Table 1: Used variables and their definition

Variable	Definition
<i>GSV</i>	Original Google search volume for given keyword
<i>ASVI</i>	The log of <i>GSV</i> for given day minus the log of median <i>GSV</i> during previous 26 days
<i>IR</i>	Log initial return of IPO calculated from the offering price to the first day closing price
<i>LR</i> ⁽¹⁾	Log cumulative return calculated from the first day closing price to the closing price one year after IPO
<i>LR</i> ⁽²⁾	Log cumulative return calculated from the first day closing price to the closing price half a year after IPO
<i>LR</i> ⁽³⁾	Log cumulative return calculated from the first day closing price to the closing price a quarter after IPO
<i>LR</i> ⁽⁴⁾	Log cumulative return calculated from the closing price one month after IPO to the closing price one year after IPO
<i>LR</i> ⁽⁵⁾	Log cumulative return calculated from the closing price one month after IPO to the closing price half a year after IPO
<i>TD_i</i>	True discount of IPO defined as in Ma & Tsai Ma and Tsai (2002). $TD = \frac{P_e - P_o}{P_o}$ where P_o is the offering price and P_e is the so-called equilibrium price – in our case the average price between $t + 150$ and $t + 180$, where t is the IPO date
<i>MR_i</i>	Market reaction to IPO defined as in Ma & Tsai Ma and Tsai (2002). $MR = \frac{P_m - P_e}{P_o}$ where P_o is the offering price, P_m is the first day closing price and P_e is the so-called equilibrium price - in this case the average price between $t + 150$ and $t + 180$, where t is the IPO date
<i>POSSENT</i>	Dummy variable that takes value of one if the level of <i>SENTIMENT</i> exceeds the third quartile, and zero otherwise
<i>NOSENT</i>	Dummy variable that takes value of one if the level of <i>SENTIMENT</i> is between the first and the third quartile, and zero otherwise
<i>NEGSENT</i>	Dummy variable that takes value of one if the level of <i>SENTIMENT</i> is below the first quartile, and zero otherwise
<i>ASVI</i> × <i>SENT</i>	<i>ASVI</i> and <i>SENTIMENT</i> interaction variable
<i>ASVI</i> ^{POSSENT}	Interaction variable that takes value of <i>ASVI</i> if the level of <i>SENTIMENT</i> exceeds the third quartile, and zero otherwise
<i>ASVI</i> ^{NOSENT}	Interaction variable that takes value of <i>ASVI</i> if the level of <i>SENTIMENT</i> is between the first and the third quartile, and zero otherwise
<i>ASVI</i> ^{NEGSENT}	Interaction variable that takes value of <i>ASVI</i> if the level of <i>SENTIMENT</i> is below the first quartile, and zero otherwise
<i>ASVI</i> × <i>IR</i>	<i>ASVI</i> and <i>IR</i> interaction variable
<i>Offering size</i>	Log size of the offering measured in the US dollars
<i>NYSE</i>	Dummy variable that take one if the offering emits its shares at NYSE and zero if it emits its shares at NASDAQ
<i>Crisis</i>	Dummy variable that takes value of one for days in interval $\langle 3, December\ 2007; 30, June\ 2009 \rangle$, and zero otherwise
<i>Sentiment</i>	Monthly time-varying aggregate market sentiment orthogonalized with respect to a set of macroeconomic conditions developed by Baker & Wurgler Baker and Wurgler (2006)
Δ <i>Sentiment</i>	Month on month difference in time-varying aggregate market sentiment orthogonalized with respect to a set of macroeconomic conditions developed by Baker & Wurgler Baker and Wurgler (2006)

Table 2: IPO first-day return and ASVI

The dependent variable in each regression is the IPO first day return IR_i , IR_i and independent variables are defined in Table 1. *, **, and *** represent significance at the 10%, 5%, and 1% level, respectively, standard errors are shown in the parentheses. N is number of observations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
$ASVI_i$	0.414*** (3.311)				0.400*** (3.138)	0.437*** (3.768)	0.397*** (3.169)	0.404*** (3.217)	0.357*** (3.084)				
$Offering\ size_i$		-0.119 (-1.090)			-0.094 (-0.929)		-0.015 (-0.153)	-0.130 (-1.269)	-0.068 (-0.685)				
$Sentiment_i$			0.112 (0.876)			0.003 (0.025)		-0.020 (-0.177)			-0.039 (-0.242)		
$\Delta Sentiment_i$				0.044 (0.345)				0.082 (0.711)				-0.037 (-0.291)	
$ASVI_i^{POSSENT,st}$										0.297** (2.600)	0.275** (2.034)	0.268** (2.276)	0.344*** (2.815)
$ASVI_i^{NEGSENT,st}$										0.163 (1.231)	0.152 (1.253)	0.136 (1.260)	0.280 (1.539)
$ASVI_i^{NOSENT,st}$													0.268 (1.365)
$Constant$	0.003 (0.026)	-0.045 (-0.443)	-0.044 (-0.398)	0.023 (0.203)	-0.065 (-0.681)	-0.034 (-0.369)	-0.019 (-0.190)	-0.073 (-0.763)	-0.089 (-0.936)	-0.011 (-0.109)	-0.048 (-0.484)	-0.028 (-0.275)	0.064 (0.574)
N	70	72	67	70	67	65	67	63	66	69	68	67	66

Table 3: IPO long-term performance and ASVI

Long-term performance LR_i and independent variables are defined in Table 1. The columns show over which period the cumulative return is calculated: first day closing price to the (1) closing price one year, (2) half a year (3) and 91 days after IPO; and the closing price one month after IPO to (4) the closing price one year (5) and half a year after IPO. *, **, and *** represent significance at the 10%, 5%, and 1% level, respectively, standard errors are shown in the parentheses. N is number of observations.

	(1)	(2)	(3)	(4)	(5)
$ASVI_i$	-0.171 (-1.19)	-0.204* (-1.97)	-0.0662 (-0.63)	-0.190 (-1.29)	-0.187** (-2.15)
$Constant$	0.0292 (0.22)	0.0711 (0.84)	0.102 (1.30)	0.0265 (0.19)	0.0775 (1.00)
N	59	60	59	59	60

Table 4: IPO long-term performance, ASVI and initial returns

The dependent variable in each regression is cumulative long-term return LR_i . LR_i and independent variables are defined in Table 1. The columns show over which period the cumulative return is calculated: first day closing price to the (1) closing price one year, (2) half a year (3) and 91 days after IPO; and the closing price one month after IPO to (4) the closing price one year (5) and half a year after IPO. *, **, and *** represent significance at the 10%, 5%, and 1% level, respectively, standard errors are shown in the parentheses. N is number of observations.

	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
IR_i	-0.143 (-1.263)	-0.053 (-0.566)	-0.020 (-0.237)	-0.221** (-2.083)	-0.162** (-2.052)	-0.387* (-1.94)	-0.317** (-2.19)	0.112 (1.07)	-0.411** (-2.47)	-0.293** (-2.16)
$ASVI \times IR_i$						0.0185 (0.14)	0.0768 (0.93)	0.176*** (2.76)	-0.0229 (-0.18)	0.104 (1.35)
$Constant$	0.221** (2.477)	0.218*** (3.094)	0.197*** (3.048)	0.195** (2.245)	0.200*** (3.185)					
N	56	56	57	57	57	58	59	58	60	60

Table 5: IPO long-term performance, ASVI and sentiment

The dependent variable in each regression is cumulative long-term return LR_i . LR_i and independent variables are defined in Table 1. The columns show over which period the cumulative return is calculated: first day closing price to the (1) closing price one year, (2) half a year (3) and 91 days after IPO; and the closing price one month after IPO to (4) the closing price one year (5) and half a year after IPO. *, **, and *** represent significance at the 10%, 5%, and 1% level, respectively, standard errors are shown in the parentheses. N is number of observations.

	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
$ASVI_i^{POSSENT,st}$	-0.388** (-2.340)	-0.206** (-2.300)	0.007 (0.067)	-0.326*** (-2.932)	-0.094 (-0.813)	-0.414** (-2.066)	-0.058 (-0.273)	-0.003 (-0.014)	-0.423** (-2.116)	-0.053 (-0.252)
$ASVI_i^{NOSENT,st}$	-0.033 (-0.189)	-0.019 (-0.099)	0.247 (1.127)	-0.008 (-0.048)	-0.019 (-0.101)	0.189 (0.900)	-0.073 (-0.329)	-0.044 (-0.200)	0.163 (0.780)	-0.118 (-0.538)
$ASVI_i^{NEGSENT,st}$	0.056 (0.449)	-0.052 (-0.573)	0.034 (0.410)	0.085 (0.729)	0.013 (0.134)	0.309 (1.363)	0.158 (0.665)	0.055 (0.232)	0.350 (1.550)	0.206 (0.866)
$POSSENT_i$										
$NOSENT_i$										
$NEGSENT_i$										
<i>Constant</i>	0.031 (0.281)	0.111 (1.232)	0.208** (2.302)	0.056 (0.552)	0.160* (1.823)					
<i>N</i>	54	57	56	55	55	62	62	62	62	62

Table 6: Ma-Tsai model and ASVI

The dependent variables are true discount TD_i and market reaction MR_i as defined by Ma & Tsai Ma and Tsai (2002). TD_i , MR_i and independent variables are defined in Table 1. *, **, and *** represent significance at the 10%, 5%, and 1% level, respectively, standard errors are shown in the parentheses. N is number of observations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	TD_i	MR_i	TD_i	MR_i	TD_i	MR_i	TD_i	MR_i
$ASVI_i$	0.00754 (0.06)	0.221* (1.88)						
$ASVI \times IR_i$			0.109 (0.62)	0.428* (1.86)				
$ASVI_i^{POSSENT,st}$					-0.130 (-0.717)	0.252 (1.641)		
$ASVI_i^{NOSENT,st}$					-0.081 (-0.494)	0.083 (0.482)		
$ASVI_i^{NEGSENT,st}$					0.025 (0.214)	0.153 (1.136)		
$POSSENT_i$							0.098 (0.464)	-0.013 (-0.063)
$NOSENT_i$							-0.130 (-0.590)	-0.042 (-0.190)
$NEGSENT_i$							0.027 (0.114)	0.066 (0.276)
$Constant$	-0.0451 (-0.39)	0.0946 (0.83)	-0.0406 (-0.39)	0.0389 (0.31)	-0.067 (-0.590)	0.051 (0.457)		
N	58	56	56	57	56	55	62	62

Figure 1: Increase in investor attention prior IPO

The vertical axis shows the average *GSV* for the analysed sample, dashed lines represent the 95% confidence intervals. The horizontal axis shows days left to IPO.

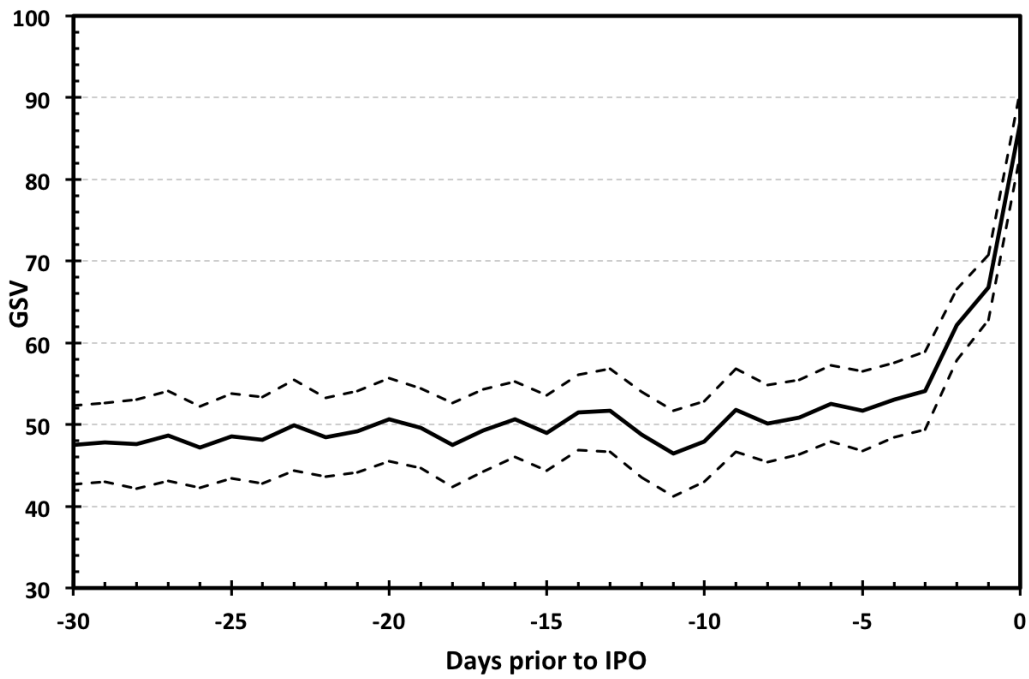


Figure 2: Long-term cumulative returns for low and high attention IPOs

Average cumulative log-returns: first day closing price to the (1) closing price one year, (2) half a year (3) and 91 days after IPO; and the closing price one month after IPO to (4) the closing price one year (5) and half a year after IPO.

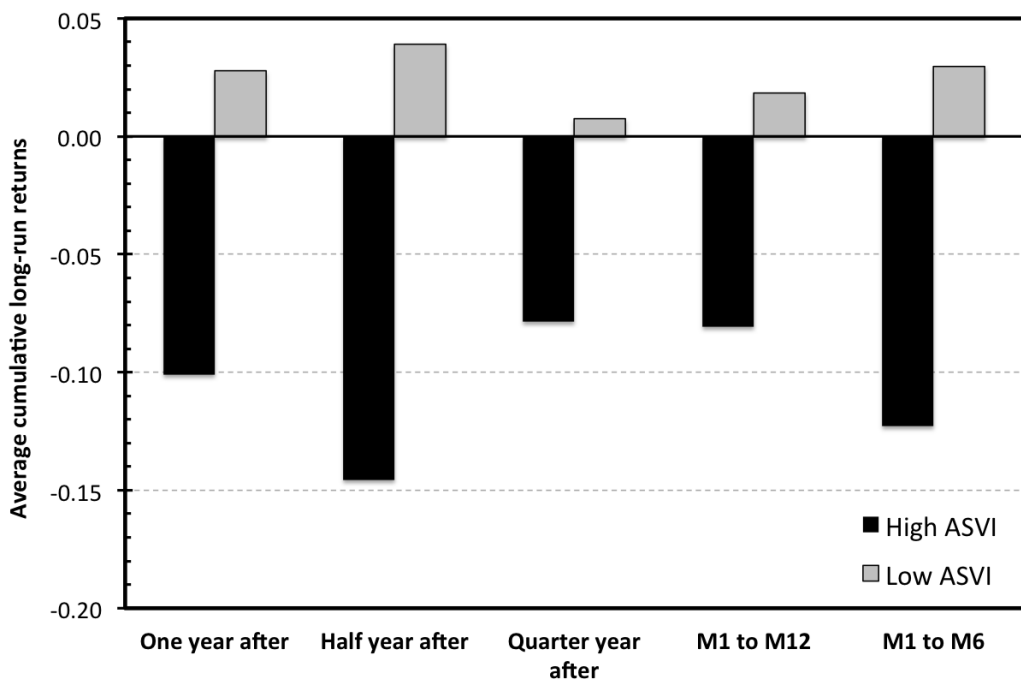


Figure 3: True discount and market overreaction for low and high attention IPOs.

