The Impacts of Social Media on Bitcoin Performance

Completed Research Paper

Feng Mai
Howe School of Business
Stevens Institute of Technology
Hoboken, NJ, USA
feng.mai@stevens.edu

Qing Bai
College of Business
University of Wisconsin-Eau Claire
Eau Claire, WI, USA
baiq@uwec.edu

Zhe Shan
Carl H. Lindner College of Business
University of Cincinnati
Cincinnati, OH, USA
zhe.shan@uc.edu

Xin (Shane) Wang
Ivey Business School
Western University
London, Ontario, Canada
xwang@ivey.uwo.ca

Roger H.L. Chiang
Carl H. Lindner College of Business
University of Cincinnati
Cincinnati, OH, USA
roger.chiang@uc.edu

Abstract

As the world's first completely decentralized digital payment system, the emergence of Bitcoin represents a revolutionary phenomenon in financial markets. This study examines the dynamic relationships between social media and bitcoin performance. We consider the distinct effects of different social media platforms and different user groups subdivided by posting volume. The results suggest that more bullish forum posts have a positive effect on bitcoin returns, and the effect is stronger when we only include the posts by users who are less likely to contribute. In addition, messages on Internet forum have stronger impacts on future bitcoin market measures at a daily frequency, but microblogs' effects are more significant at an hourly frequency.

Keywords: Social media, online communities, user-generated content, Bitcoin

Introduction

Since the 2008 invention of Bitcoin by an unidentified programmer known as Satoshi Nakamoto, the virtual currency has achieved great success, such that by July 2015, the value of all bitcoins in the world surpassed 3 billion USD.[1] Remarkable for a cryptocurrency that exists solely in digital form and is backed by no central bank or other authority, a bitcoin in September 2014 was worth US$400–500 and traded by approximately 1.5 million Bitcoin users, who exchange more than 100,000 bitcoins daily. Well-known companies such as Dell and Newegg.com accept bitcoin; several Bitcoin ATMs now operate in five cities on four continents. According to CoinDesk's (an online publication that tracks digital currencies) recent estimate, there are eight million bitcoin trading accounts and 100,000 retailers that accept bitcoin in the first quarter of 2015. Thus the rise of Bitcoin seems unstoppable.

Because of Bitcoin's decentralized structure and exclusively online presence, the value of bitcoins derives not from gold or government fiat but from the value that people assign. The dynamics of the bitcoin price thus should relate to pertinent discussions and opinions on online social media, where investors and business adopters interact and provide feedback about the market. Tirunillai and Tellis (2012) have posited that social media and user-generated content (UGC) constitute important determinants of investments. Social media capture the "wisdom of the crowd" and provide low-cost platforms for connecting with target markets. Bitcoin provides a unique opportunity to observe and understand the interplay of social media with the value of a financial instrument. Accordingly, we examine the influence of social media metrics on bitcoin returns and strive to answer the following research questions:

- What factors characterize trading behavior surrounding bitcoins? Does it behave like a currency or resemble a speculative investment, similar to Internet stocks?
- Can user-generated content, available through social media, affect value in the bitcoin market? Do distinctive social media characteristics (e.g., user and platform differences) affect this relationship?

We conduct an empirical analysis of the bitcoin market, using vector autoregressive (VAR) and vector error correction (VECM) models with exogenous control variables. In cases when some variables share common trends, the VECM provides a more appropriate framework than standard VAR models, because it can explore dynamic movements among cointegrated focal variables. We assemble diverse data sources, from bitcoin and stock markets, traditional Internet measures, and social media; we also construct social media metrics using data from an Internet forum (bitcointalk.org) and a microblogging service site (Twitter) to assess the volume of posts, user sentiments, and measures of bullishness and disagreement across forum contributors. To incorporate the influence analysis in our models, we stratify the sample by the characteristics of social media users, dividing users of the Internet forum and Twitter into the "silent majority" and "vocal minority," according to their contribution levels.

A summary of our major findings is as follows. We find that the number of bullish (bearish) forum posts has a positive (negative), statistically significant effect on future bitcoin returns when we consider daily frequencies. At an hourly frequency, the number of bullish tweets reveals a positive, statistically significant relationship with future bitcoin returns. Disagreement among forum contributors also affect future bitcoin trading volume at a daily frequency. The UGC obtained from different users and platforms exerts different impacts on bitcoin returns, such that UGC contributed by the vocal minority and silent majority display distinct relationships with the future bitcoin market. Finally, regarding the relative effect of both types of social media, the daily data shows that the Internet forum variables influence future bitcoin prices significantly, but Twitter variables do not. The hourly data instead shows that Twitter variables have stronger effects.

In the next section, we develop our theoretical background and hypotheses. We then introduce the measures and data sample for our empirical study, after which we develop our empirical models (VAR and VECM). Subsequent to the discussion of our findings, we conclude this article with some implications and insights.

[1] We use "Bitcoin" to refer to the online payment system; "bitcoin" is the unit of currency.

Theoretical Background and Hypotheses

Can Social Media Influence Bitcoin Returns?

Because bitcoin is the most popular virtual currency, a reasonable expectation is that it behaves similarly to traditional currencies, such that its price would be driven by its use in transactions, its supply, and the price level (of tradable goods and services) (Kristoufek 2014). However, virtual currency is a fundamentally different financial phenomenon, defined as “a type of unregulated, digital money, which is issued and usually controlled by its developers, and used and accepted among the members of a specific virtual community” (European Central Bank 2012). Transactions on the Bitcoin network are not denominated in dollars or any other currency. Bitcoin is a virtual currency, supported by a decentralized payment network. Because the value of bitcoin derives not from gold or government fiat but from the value that people assign to it, its monetary value gets determined on an open market, similar to the exchange rate among world currencies (Brito and Castillo 2013). Yet the bitcoin’s daily exchange rates do not correlate with traditional currencies (Kristoufek 2014), and its exchange rate volatility is orders of magnitude greater than the volatilities of those more widely used currencies (Yermack 2013). In these traits, bitcoin seemingly mimics an Internet stock, rather than a currency.

Researchers have tested whether the Internet in general, and UGC specifically, influences the underlying behavior of stock markets. Those studies lead to contradictory results. Based on a sample of the 50 firms with the greatest Yahoo message board posting volume, Wysocki (1999) finds that changes in daily posting volume are associated with both earnings announcement events and changes in stock trading volume and returns. However, Das and Chen (2007) study a sample of 24 tech-sector firms, and find no significant predictive ability for individual stock returns using message board sentiment. In terms of economical scales, Chen et al. (2014) reveal that the views expressed in both articles and comments on a social media platform predict future stock returns and earnings surprises, with an effect that is both statistically and economically significant. In contrast, Antweiler and Frank (2004) indicate that a positive shock to message board posting predicts negative stock returns on the next day, though the effect is economically small. Moreover, they find that both the level of message posting and disagreement among messages seemingly predict subsequent trading volume. This finding is not consistent as in Tumarkin and Whitelaw (2001), in which the authors examine Internet stocks and find that message board activity cannot predict stock returns; rather, the causality appears to run from the market to the forums.

In the context of bitcoin, online messages can disclose new or private information that fundamentally alters bitcoin evaluations, such as when new stores accept bitcoins or forthcoming regulations limit its use. Online discussions thus offer good indications of the general market sentiment toward bitcoins. In addition, speculative investors tend to follow the trends, which may exaggerate the effects of such information. Hence, we anticipate that the UGC on online platforms should also has an effect on the investment returns and trading volume of bitcoins, especially for the following reasons. First, the decentralized nature of Bitcoin meant that most early users were individuals, rather than large institutional investors. These early adopters arguably contribute to social media more frequently and are more likely to be influenced by social media. Second, an important motivation for early institutional Bitcoin adopters was to capture positive public relations through social media, in that “being noted as a Bitcoin innovator can potentially generate favorable press and social media mentions” (PricewaterhouseCoopers LLP 2014). For instance, when the social gaming company Zynga added Bitcoin to its most popular games in 2014, it garnered thousands of media mentions. Furthermore, the design of Bitcoin’s algorithm ensures that the supply of new coins gets created at a known, geometrically decaying rate, so demand from both businesses and individuals represents the main driver of the bitcoin’s value. Finally, according to a recent survey (Duggan and Brenner 2013), Bitcoin users largely share the demographic characteristic of being heavy social media users. Therefore, we postulate:

H1. Social media metrics have significant effects on future bitcoin returns, such that (a) increased positive (negative) sentiments indicate higher (lower) future bitcoin prices and (b) disagreements on social media indicate greater future trading volume.

The Significant Role of Social Influence in the Bitcoin Market

The power-law nature of social media implies that most social media users contribute little content; this “silent majority” contributes to conversations sporadically, mostly after important events, and are not particularly interested in generating buzz (Metaxas and Mustafaraj 2012; Mustafaraj et al. 2011). Moreover, social media users from different groups (e.g., silent majority vs. vocal minority) have significant differences in the generated content and tweeting behavior (Chen et al. 2012). As we previously argued, the decentralized nature of Bitcoin meant that most grassroots users can usually be categorized as "silent majority". In this sense, the UGC from the silent majority may be a more compelling metric for actual investors.

However, social influence can also play an important role in information cascading. In financial markets this phenomenon is well documented, manifested as herding behavior. Friend et al. (1970) note the significant tendency of groups of mutual funds to adopt the investment choices of their more successful counterparts, which they call follow-the-leader behavior. Jiao and Ye (2013) find strong evidence that mutual funds collectively enter or exit stocks, following the herd of hedge funds. According to Brown et al. (2013), mutual fund managers follow analyst recommendation revisions when they trade stocks, and these analyst-motivated trades move stock prices. Mutual funds herd into stocks following consensus analyst upgrades and even more evidently herd out of stocks with consensus downgrades.

Moreover, related studies in the marketing and IS literature also highlight the importance of identifying influential users on social media. As Trusov et al. (2010) observe, community members differ in the frequency, volume, type, and quality of digital content they generate and consume. Influential people, such as opinion leaders, have disproportionate influence on others (Godes and Mayzlin 2009; Goldenberg et al. 2009), largely because they have greater exposure to mass media than their followers, They are more cosmopolitan, engage in more social participation, have a higher socioeconomic status, and are more innovative (Rogers 2010). The effectiveness and functioning of an online community strongly depend on the presence and activities of a vocal minority of opinion leaders, who can induce effects in various ways (Mehra et al. 2006). Following this stream of literature, we naturally expect to observe some degree of follow-the-influencer behavior in the bitcoin market too. Therefore, we hypothesize:

H2. The vocal minority has stronger impacts than the silent majority on the bitcoin market.

The Distinct Impact of UGC from Internet Forum versus Microblogging

In addition to user-level influence differences, we predict that different social media platforms affect financial markets differently. First, Internet forums are generally geared toward creating a collaborative environment that answers different users’ similar queries about a topic. Usually, seeking opinion consensus is not a primary objective. In contrast, on a microblogging site such as Twitter, communications move from the sender to his or her followers, who can spread the information further by retweeting. Limited by length restrictions, these followers might add brief, general sentiments, but they cannot engage in a thorough discussion of the original content. Second, a discussion forum is not a typical social network application, because it enables users to engage in discrete, transitory exchanges, without respect to proximity, social relations, or flows. But as a representative social networking site, Twitter supports a richer range of possible relational ties, including social relations in which users establish persistent connections with (i.e., follow) other users. Third, most users access an Internet forum through Web browsers. Twitter offers both web and mobile access, but users mainly engage through mobile devices[2]. These distinctive features in turn may have significant moderating effect on the impacts of UGC.

Accordingly, the link between the nature of a social media platform and the adoption of UGC has received attention. A one-sided message presents positively or negatively valenced information; a two-sided message can include both (Cheung and Thadani 2012). Kamins and Assael (1987) observe that two-sided information enhances information completeness, invoking greater credibility perceptions. Finance scholars also note that because investors have limited attention capacities, they respond asymmetrically to more visible information (Barber and Odean 2008; Hirshleifer and Teoh 2009): when information is more visible and accessible, investors are more likely to respond to it. The relationship between discussion

[2] Fitzgerald (2014) reports that 85% of time Twitter users spent on Twitter happened on a mobile device.

patterns and bitcoin prices thus should appear at a daily level, whereas the responses of the bitcoin market to the spread of news on Twitter may occur at an intra-day level, due to Twitter's mobile nature (Tafti et al. 2013). We hypothesize:

H3. User-generated content from Internet forum and microblogging site have different impacts for bitcoin returns.

Data and Variables

Bitcoin Market Variables

The data set comprises daily market prices (i.e., USD exchange rate) and trading volume series (in USD) from BitStamp Ltd., the top bitcoin exchange by volume. We also collected bitcoin-to-bitcoin transaction volume, defined as the total value of all transaction outputs per day, from bitcoincharts.com. The exchange trading volume measure refers to the amount of bitcoin traded for other currencies; transaction volume indicates the amount spent in the bitcoin economy. Because the total output volume includes coins returned to the sender as change, we use the adjusted transaction volume, net of change, which should offer a more accurate reflection of the true transaction volume. We denote the trading volume and transaction volume of day t as V_t and V_t^{TX}, respectively.

In addition, we define S_t as the market price of bitcoin at the end of day (hour) t. The continuously compounded return in bitcoin is the first difference of the log price:

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln(S_t) - \ln(S_{t-1})\ . \tag{1}$$

To measure the volatility of the daily/hour return we apply the exponentially weighted moving average (EWMA) model. The EWMA model tracks changes in the volatility, with the formula:

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)u_{t-1}^2. \tag{2}$$

The estimate of volatility on day t, σ_t^2 (obtained at the end of day $t-1$), is calculated from σ_{t-1}^2 (i.e., the estimate at the end of day $t-2$ of volatility for day $t-1$) and u_{t-1}^2 (most recent daily percentage change). The value of λ governs the responsiveness of the estimate to the most recent daily percentage change. We choose λ = .94, the value used by RiskMetrics[3]. In addition, at an hourly frequency, we measure the absolute value of the hourly return $|r_t|$.

Social Media Metrics

We implemented a Python-based Web crawler to collect discussion content from bitcointalk.org between November 22, 2009, and August 18, 2014. We chose this forum for two reasons: It was rated the most popular Bitcoin community in a recent survey (Smyth 2013), and it appears first in the community section of the official Bitcoin website. We limit our data collection to the Bitcoin discussion board, to which users contribute general news, community developments, innovations, and so forth. After filtering out content beyond our study period, we gathered 119,847 posts and 51,269 topics to retain for further analysis. Each post contained textual content, an author, and a timestamp. Among the 69,671 unique users who posted, the most active 5% of users generated 62.6% of the content. The average number of posts generated by a single user in the sample period was 12.75; the median was 3. As Figure 1 reveals, the distribution of the number of messages by users has a very long tail, such that most users are in the silent majority, and a small proportion of the vocal minority generated the most contents.

Next, we collected tweets (i.e., microblogging text messages of no more than 140 characters) that contained the hashtag "#Bitcoin" from the public application program interfaces (API) of Twitter. The

[3] The RiskMetrics database, originally created by JPMorgan and made publicly available in 1994, uses a EWMA model and λ = .94 to update daily volatility in its database. The company demonstrated that, across a range of market variables, this value of λ results in variance rate forecasts that come closest to the realized variance rate.

Twitter platform offers two different APIs, i.e. search API and streaming API, to access the data. The Twitter streaming API gives developers low latency access to Twitter's global stream of Tweet data, while the Twitter search APIs allow queries against the indices of recent or popular Tweets. The search API offers more powerful queries, and can collect a wider range of data, while the streaming API usually returns a much higher flow of Tweets. Based on our tests, the streaming API ejects newest and unique Tweets, and search API provides both latest and updated (favorited or retweeted) ones. We collected data from both APIs (at its highest frequency for the search API, which limits to 180 queries per 15-minute window) between April 18th and August 18th, 2014, during which we gathered 3,348,965 unique tweets from 339,295 unique users. In average 21,910 users tweeted 27,227 messages per day. Each tweet contains textual content, author information, a timestamp, and propagation-related flag and counting data.

For the sentiment analysis of forum posts and tweets, we applied a finance sentiment dictionary (Loughran and McDonald 2014), which includes 2,329 negative and 297 positive sentiment words. We used Natural Language Toolkit 3.0 (Bird 2006) for the language processing tasks, such as sentence segmentation, word tokenization, and lemmatization. For each post/tweet, we counted the number of positive and negative words. If a post contains more positive than negative words, it constitutes a positive post, and vice versa.

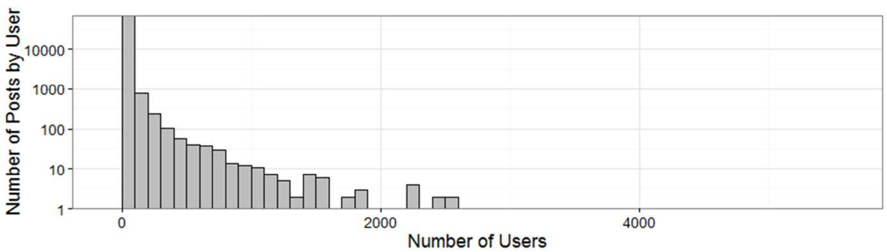

Figure 1: Distribution of posts by users

Finally, we considered the distinct influence of each tweet, according to its characteristics. In our data set, for each tweet, we collected *favorite_count*, to assess how many times a tweet had been listed as a "favorite" by Twitter users; *retweet_count*, or the number of times a tweet had been retweeted; and *followers_count* for the tweet's author, equivalent to the number of followers the account had at that moment. However, *favorite_count* and *retweet_count* are continuously updated along the tweet lifecycle, making it difficult to track their updates at hourly or daily frequencies. Therefore, we identified *follower-count* as the measure that offers insightful influence information. For each day, we ranked all tweets according to this measure, chose the top 20 tweets in the ranked list (following the practice in (Shi et al. 2014)), and calculated their sentiment scores. Table 1 summarizes our key measures.

Other Variables

We included a set of traditional Internet activity measures and stock market returns (S&P500 and NASDAQ composite) as exogenous control variables. To measure search interest related to Bitcoin, we collected data from Google Trends (www.google.com/trends/). The measure of *interest over time* indicated the popularity of a given keyword (in our case, *bitcoin*) in Google's search engine. We also gathered Web traffic data from the Alexa Web Information Service (aws.amazon.com/awis/). A Python program fetched traffic data related to bitcoin.org, including *reach* (number of unique visitors), *page views per user* (average number of webpages a user visits), and *traffic rank* (estimated daily ranking of the website).

Table 1: Key Measures

Variable	Meaning
r_t	Bitcoin returns, continuously compounded
σ_t^2	Volatility of bitcoin returns, updated daily
V_t	Daily trading volume (logged, liner trend removed) from top exchanges
V_t^{TX}	Estimated daily transaction volume (logged, liner trend removed)
POS^F	Number of positive forum posts in a day
NEG^F	Number of negative forum posts in a day
POS^T	Number of positive tweets in a day
NEG^T	Number of negative tweets in a day

Empirical Methodology

We are interested in both the contemporaneous and the dynamic relationships between social media and the bitcoin market. With our contemporaneous analysis, we seek to determine whether variation in social media activities is just noise or is associated with underlying market activities. Therefore, we examined four fundamental measures of bitcoin market activities: bitcoin returns, volatility of bitcoin returns, trading volume on major bitcoin exchanges, and bitcoin transaction volume. If, as we hypothesized, social media contain new information about the fundamental value of bitcoins, the bullishness measure should relate positively to (future) bitcoin returns. Even if the messages contain no new information, they may capture a general market sentiment and thus be positively correlated with bitcoin returns. Similar to Antweiler and Frank (2004), we consider the simple pairwise correlations of bitcoin market variables with social media variables.

For the dynamic analysis, we adopt a vector autoregression (VAR) system to capture linear inter-dependencies across time series. We choose the VAR approach rather than a more traditional multiple regression (cf. Antweiler and Frank (2004); Wysocki (1999)) for several reasons. First, with a VAR model, we can treat all of the key variables as jointly endogenous, without creating ad hoc model restrictions. Nor do we need the extensive knowledge about the forces influencing a variable, as required by structural models with simultaneous equations. Second, the model allows for both autocorrelation and cross-correlation, so we can better understand the dynamic relationships among the variables. Third, we can interpret the estimated model using Granger causality, impulse response functions, and forecast error variance decomposition. In prior finance literature, VAR models often support portfolio analyses of various risky assets or assessments of domestic and foreign interest rates jointly with foreign exchange rates. For example, Tumarkin and Whitelaw (2001) use a linear VAR model with one lag to examine the dynamic relationship of daily stock returns, trading volume, Internet message posting volume, and changes in opinions expressed in the messages.

In our empirical study, we examine models in which the endogenous variables include bitcoin market activities, namely, returns (r_t), volatility (σ_t^2), transaction volume (V_t^{TX}) and trading volume (V_t); the models also include measures of relevant social media activities, namely, number of forum posts or tweets expressing positive/bullish opinions (POS^F, POS^T) and negative/bearish opinions (NEG^F, NEG^T). In addition, we consider a model that combines the level and bullishness of Internet forum activities similar to Antweiler and Frank (2004)'s proposal for aggregating message classifications. For example, let M be the total number of new messages in the forum, and R=POS^F/NEG^F be the ratio of bullish to bearish messages. Then our bullishness measure is:

$$Bullishness \ = M * \frac{R-1}{R+1}. \tag{3}$$

To measure disagreement among forum contributors, we constructed an agreement index (Antweiler and Frank (2004):

$$AI = 1 - \sqrt{1 - bullishness\ I^2}. \tag{4}$$

This index is bound between 0 and 1 and decreases with greater disagreement levels.

Our models examine bitcoin market activities and social media activities over both long and short time intervals. The first set of models (1a and 1b) use aggregated forum data at daily frequency. Model 2 estimates the impact of Twitter activities at hourly frequency. The last two models 3 and 4 compare the impacts of both Internet forum and Twitter activities at daily and hourly frequency respectively. We summarize all the models in Table 2.

To determine an appropriate VAR system, we first test the stationarity of the variables. Conventional regression estimators, including VAR, encounter problems when applied to nonstationary processes, such that the regression of two independent random walk processes would yield a spurious significant coefficient, even if they were not related (Granger and Newbold 1974). We used an augmented Dickey-Fuller unit root test of each variables, with lag numbers chosen according to the Schwert (1989) rule. Among the time series in the model, the number of positive posts, S&P500 returns, and NASDAQ returns are stationary; the others have one order of integration.

Table 2: Model Summary

Model	Endogenous Variable	Sample Period	Sampling Frequency
1a	$r_t, \sigma_t^2, V_t, V_t^{TX}, POS^F, NEG^F$	1/1/2011-8/18/2014	Daily
1b	$r_t, \sigma_t^2, V_t, V_t^{TX}, M, bullishness, AI$	1/1/2011-8/18/2014	Daily
2	$r_t, \lvert r_t \rvert, V_t, POS^T, NEG^T$	4/18/2014-8/18/2014	Hourly
3	$r_t, \sigma_t^2, V_t, V_t^{TX}, POS^F, NEG^F, POS^T, NEG^T$	4/18/2014-8/18/2014	Daily
4	$r_t, \lvert r_t \rvert, V_t, POS^F, NEG^F, POS^T, NEG^T,$	4/18/2014-8/18/2014	Hourly

Next, we determined the appropriate lag length p using Akaike's information criterion (AIC) and the Bayesian information criterion (BIC), as is standard in the VAR literature (Love and Zicchino (2006)). For each model, we calculated the AIC and BIC values for the sample period and chose the lag length that minimized both criteria. If they indicated conflicting optimal lag lengths, we chose the length that minimized the BIC.

Although in a VAR system, we could model the interrelationship of the variables by taking first differences of each non-stationary series and including the differences in a VAR, this approach can suffer misspecification biases if cointegration is present. In that case, VAR expresses only the short-run responses between variables, without providing information about the long-run equilibrium in the case of cointegration between two or more series. We performed a Johansen test (Johansen and Juselius 1990) and confirmed the presence of cointegration in our daily frequency data. Therefore, in Models 1a, 1b, and 3, we extended the VAR model to an vector error correction model (VECM), which can fit the first differences of the non-stationary variables, using a vector of error correction terms that is equal in length to the number of cointegrating relationships added to the relationship (see Johansen (1995)). By taking potential long-term relationships into account, the VECM model with p variables, k lags, and cointegration rank r has the following form:

$$\Delta \boldsymbol{Y}_t = \sum_{j=1}^{k-1} \Gamma_j \Delta \boldsymbol{Y}_{t-k} + \alpha\beta' \boldsymbol{Y}_{t-1} + \boldsymbol{\mu} + \epsilon_t \tag{5}$$

where Δ is the first difference operator, $\boldsymbol{Y}_t$ is a $p \times 1$ vector with order of integration 1 , $\boldsymbol{\mu}$ is a $p \times 1$ constant vector representing the linear trend, k is the a lag structure, and ϵ is the residual vector. Γ_j is a $p \times p$ matrix that indicates short-term relationships among variables, β is a $p \times r$ matrix that represents the long-term relationships between the cointegrating vectors, and α is a $p \times r$ matrix denoting the speed of variables adjusting to the long term equilibriums. The difference between the VECM model and the VAR model with first differenced variables is the additional $\beta' Y_{t-1}$, known as the error correction term. In addition, in the VECM, the non-differenced variables (i.e., price (P_t) instead of return (r_t), number of posts instead of increase of posts) are used in the estimation because the model itself has first differences built in. The VECM model can be considered as a special case of the general VAR system as it can be expressed as an equivalent VAR.

Results

Contemporaneous Relation

The correlations between the bitcoin market measures and the Internet forum measures are large, with magnitudes comparable to the strong correlations of different market measures. For example, the number of relevant messages correlates at .312 with volatility, at .556 with transaction volume, and at .445 with trading volume. Because the number of posted messages is highly autocorrelated, its one-day lag value also correlates strongly with bitcoin volatility, transaction volume, and trading volume. This result confirms Wysocki (1999) finding regarding the relationship between message board activities and stock market activities: Daily variation in real market activity and information events can explain much of the variation in daily change in message postings. Because the magnitude of the relationships between bitcoin market measures and message board posting volume are not trivial, we conclude that social media reflect market information rapidly.

Row 6-7 of Table 3 contains the correlations of various bullishness measures with the bitcoin market measures. Bullishness correlates positively with contemporaneous returns and negatively with volatility and volume. According to the agreement index, disagreement among Internet forum contributors is associated with more trading. Finally, the bottom section of Table 3 reveals the strong correlations between traditional Internet activity measures and bitcoin market measures; they are also highly autocorrelated over time.

Table 3: Pairwise Correlations

	Return	Volatility	Transaction Volume	Trading Volume
Volatility	-0.002	1.000	**0.420**	**0.369**
Transaction volume	0.039	**0.420**	1.000	**0.732**
Trading volume	**0.084**	**0.369**	**0.732**	1.000
# Post	**-0.066**	**0.312**	**0.556**	**0.445**
# Post, 1d lag	-0.037	**0.311**	**0.531**	**0.425**
Bullishness	**0.078**	**-0.252**	**-0.424**	**-0.312**
Bullishness, 1d lag	0.022	**-0.248**	**-0.402**	**-0.280**
Agreement	-0.053	0.000	-0.039	**-0.056**
Agreement, 1d lag	0.004	-0.008	-0.034	**-0.072**
Rank	-0.026	**-0.251**	**-0.399**	**-0.268**
Rank, 1d lag	-0.018	**-0.245**	**-0.245**	**-0.399**
Reach	0.015	**0.329**	**0.543**	**0.408**
Reach, 1d lag	0.008	**0.334**	**0.540**	**0.401**
Page view	**0.068**	**0.477**	**0.106**	**-0.076**
Page view, 1 d lag	0.051	**0.472**	**0.119**	**-0.066**
Google trend	-0.023	**0.165**	**0.318**	**0.276**
Google trend, 1d lag	-0.016	**0.163**	**0.323**	**0.272**

Notes: Correlations that are significantly different from 0 at 0.05 level are bolded. The time period is one day. The four market variables are the log difference in bitcoin exchange rate (in USD) from the end of the previous day to the current day (return); daily volatility estimated with the EWMA model (volatility); the natural log of bitcoin daily transaction volume in USD, with the linear time trend removed (transaction volume); and the natural of log bitcoin daily trading volume in USD, with the linear time trend removed (trading volume).

Dynamic Relationships among Bitcoin Activity and Forum Activity

We examine the effects of the number and bullishness of forum messages posted during the whole sample period using Models 1a (Table 4). The two forum metrics generally work as we anticipated. Days with unexpected increases in the number of positive (bullish) posts tend to precede days with high bitcoin returns and high transaction volume. Days with unexpected increases in the number of negative (bearish) posts tend to precede days with lower bitcoin returns and lower transaction volume. All these relationships are statistically significant. Therefore, we can conclude that forum posts contain new information about the value of bitcoin or provide a better indication of general market sentiment than what is already contained in the trading record.

Panel b shows the results from the same model, using only messages posted by the silent majority of users (excluding top 5%)[4]. In this case, the impact of posts on bitcoin returns becomes stronger, despite the fact that posts from these users only account for less than 40% of the total posting volume. As a robustness check, Panel c contains the subsample results by time periods, which reveal that the signs of the t-statistics of POS^F and NEG^F are consistent across three years, but the relative effect varies from year to year.

With Panel d, we examine the effects of the posts from the vocal minority (top 5%). In contrast with the results for the silent majority, these posts are only strong indicators of future transaction activities but not returns. Specifically, an increase in the number of bearish posts by the vocal minority indicates lower future transaction volume, and an increase in the number of bullish posts indicates higher future transaction volume. These users are likely to be active traders on the bitcoin market. Moreover, the number of messages posted by the vocal minority depend more on market activity: Higher volatility indicates fewer posts on the next day, whereas higher transaction volume indicates more posts on the next day.

Table 4: t-Statistics from VECM Analysis (Model 1a)

Panel a: Full Sample Period, All Users

Dependent Variable	Independent Variable					
	r_{t-1}	$\Delta\sigma^2_{t-1}$	ΔV_{t-1}	ΔV^{TX}_{t-1}	ΔPOS^F_{t-1}	ΔNEG^F_{t-1}
r_t	4.16***	-0.89	-2.76***	0.50	2.60***	-1.77*
$\Delta\sigma^2_t$	11.75***	5.10***	2.24**	3.56***	-0.31	0.19
ΔV_t	-0.35	-4.77***	-7.26***	-0.31	0.61	-1.05
ΔV^{TX}_t	1.24	-4.85***	-0.22	-8.12***	2.97***	-3.07***
ΔPOS^F_t	0.33	-1.88**	0.30	3.58***	-4.80***	0.59
ΔNEG^F_t	0.70	-2.93***	-0.66	2.89***	-0.78	-0.86

Panel b: Full Sample Period, Silent Majority (Excluding Top 5% of Users)

Dependent Variable	Independent Variable					
	r_{t-1}	$\Delta\sigma^2_{t-1}$	ΔV_{t-1}	ΔV^{TX}_{t-1}	ΔPOS^F_{t-1}	ΔNEG^F_{t-1}
r_t	4.09***	-0.85	-2.69***	0.53	3.48***	-2.14**
$\Delta\sigma^2_t$	11.95***	5.10***	2.20**	3.53***	-1.54	1.09
ΔV_t	-0.30	-4.77***	-7.28***	-0.27	0.46	-0.82
ΔV^{TX}_t	1.37	-4.85***	-0.20	-8.12***	2.30**	-2.27**
ΔPOS^F_t	0.69	-0.22	1.38	2.27**	-3.65***	0.08
ΔNEG^F_t	0.74	-2.37**	-0.54	2.21**	-1.07	-1.05

[4] We performed robustness checks using cut-off values of 2.5% and 10%. The alternative cut-off values did not change our results materially.

Panel c: Subsample Periods, Silent Majority (Excluding Top 5% of Users)

Dependent Variable	Independent Variable 2011-2012		2013-2014	
	ΔPOS_{t-1}^F	ΔNEG_{t-1}^F	ΔPOS_{t-1}^F	ΔNEG_{t-1}^F
r_t	1.36	-2.46**	3.01***	-1.28
$\Delta\sigma_t^2$	-0.05	-0.09	-1.72*	0.06

Panel d: Full Sample Period, Vocal Minority (Top 5% of Users)

Dependent Variable	Independent Variable r_{t-1}	$\Delta\sigma_{t-1}^2$	ΔV_{t-1}	ΔV_{t-1}^{TX}	ΔPOS_{t-1}^F	ΔNEG_{t-1}^F
r_t	4.29***	-0.89	2.75***	0.49	1.24	-0.64
$\Delta\sigma_t^2$	11.62***	5.11***	2.24**	3.59***	0.68	-0.63
ΔV_t	-0.32	-4.70***	-7.22***	-0.32	0.22	-0.76
ΔV_t^{TX}	1.30	-4.84***	-0.24	-8.16***	2.51**	-2.62***
ΔPOS_t^F	-0.25	-2.39***	-0.39	3.88***	-5.22***	-0.52
ΔNEG_t^F	0.29	-2.81***	-0.52	3.27***	-1.42	-1.82*

Notes: The full sample period is 1/1/2011–8/18/2014. ***, **, * denote significance at the 0.01, 0.05, and 0.1 levels, respectively.

As a further robustness check, Model 1b contains an alternative set of Internet forum measures that capture posting volume and bullishness, as well as disagreement among contributors. As Table 5 shows, bitcoin returns are positively associated with the previous day's bullishness. In addition, the impacts of vocal minority and silent majority show a similar disparity as does Table 4 – the posts from less active users carry more weight in indicating future returns. As we predicted, disagreement induces trading - the agreement index is negatively associated with future trading volume. The agreement index and volatility measure emerge as the variables derived from the forum activities that exhibit significant effect on future trading volume.

Table 5: t-Statistics from VECM Analysis (Model 1b)

Dependent Variable	Independent Variable r_{t-1}	$\Delta\sigma_{t-1}^2$	ΔV_{t-1}	ΔV_{t-1}^{TX}	Number of Messages	Bullishness	AI
All Users							
r_t	4.15***	-0.31	-1.70*	-0.37	2.41**	1.93*	0.43
ΔV_t	-0.23	-3.60***	-8.27***	0.13	-0.94	-1.08	-1.66*
Vocal Minority (Top 5% of Users)							
r_t	4.37***	-0.32	-1.74*	-0.29	0.97	0.53	-0.25
ΔV_t	-0.20	-3.51***	-8.23***	0.09	-1.15	-1.21	-1.34
Silent Majority (Excluding Top 5% of Users)							
r_t	4.25***	-0.86	-2.74***	0.50	2.96***	2.93***	0.73
ΔV_t	-0.24	-3.55***	-7.22***	-0.20	-0.31	-0.15	-1.84*

Notes: This table reports on the full sample period, 1/1/2011–8/18/2014, excluding the top 5% of users. ***, **, * denote significance at the 0.01, 0.05, and 0.1 levels, respectively.

Dynamic Relationships among Bitcoin Activity and Twitter Activity

We examine the effects of the number and bullishness of all tweets over a four-month sample period (April 18–August 18, 2014), as summarized in Table 6 (Panel a), using hourly sampling frequency. The magnitudes of the *t*-statistics among the bitcoin market variables are similar to what we reported using daily-frequency data. Changes in POS^T and NEG^T have strong negative autoregressive relationships.

However, in terms of the impact on future bitcoin returns, volatility, and trading volume, effects of both variables are mostly nonexistent when using all the tweets.

Table 6: t-Statistics from VAR Analysis (Model 2)

Panel a: All Tweets, Hourly Frequency (April 18–August 18)

Dependent Variable	Independent Variable				
	r_{t-1}	$\|r_t\|$	ΔV_{t-1}	ΔPOS_{t-1}^T	ΔNEG_{t-1}^T
r_t,	8.57***	-0.31	2.50**	0.69	0.27
$\|r_t\|$	-0.72	11.76***	4.30***	0.69	0.77
ΔV_t	1.85*	-3.49***	-23.00***	1.34	1.81*
ΔPOS_t^T	-1.76*	2.37**	1.65	-30.17***	2.74***
ΔNEG_t^T	-0.25	-0.24	1.91*	0.51	-25.45***

Panel b: Tweets by Opinion Leaders, Hourly Frequency

Dependent Variable	Independent Variable				
	r_{t-1}	$\|r_t\|$	ΔV_{t-1}	$\Delta top\ 20\_POS_{t-1}^T$	$\Delta top\ 20\_NEG_{t-1}^T$
r_t,	8.58***	-0.36	2.53**	2.05**	-0.44
$\|r_t\|$	-0.71	11.71***	4.47***	0.33	1.88*
ΔV_t	1.92*	-3.58***	-22.80***	0.34	0.47
$top\ 20\_POS_t^T$	-1.30	-1.15	1.16	-35.31***	-1.11
$top\ 20\_NEG_t^T$	1.03	-0.20	-0.87	-0.36	-1.48

Notes: ***, **, * denote significance at the 0.01, 0.05, and 0.1 levels, respectively.

To determine if incorporating a social influence analysis improves the predictive power of the tweet metrics, we replaced POS^T and NEG^T (all tweets) with the $top\_20\_POS^T$ and $top\_20\_NEG^T$ (tweets by people who rank in the top 20 in terms of follower counts (Shi et al. 2014)), which is similar to using weighted opinions instead of simple aggregated opinions. The results for these opinion leaders appear in Panel b. Hours with many bullish (bearish) tweets by opinion leaders precede hours with high (low) bitcoin returns, and the relationship between bullish tweets and bitcoin returns is statistically significant at the 5% level. We define opinion leaders as those who rank among the top 20 in terms of follower counts, following the practice in (Shi et al. 2014). In a robustness check, we ran analyses with the 50 and 100 top ranked opinion leaders and found quantitatively similar results.

The Relative Impacts of Internet Forum versus Twitter Variables

To examine the relative impacts of the Internet forum and Twitter variables, we used a VECM model with eight endogenous variables (Model 3) for daily data and a VAR model with seven endogenous variables (Model 4) for hourly data, and the results are reported in Table 7. At a daily frequency (Panel a), the forum variables help predict bitcoin returns one day in the future, and the relations are statistically significant. Days with many bullish and few bearish forum posts precede days marked by higher bitcoin returns. In contrast, the Twitter variables have negligible influence on future bitcoin returns. No social media variables exhibit significant effect on volatility or trading volume during the sample period.

Panel b of Table 7 presents the results with hourly data. For the Twitter data, we limit our sample to tweets by top 20 opinion leaders following the findings reported in Table 6. Hours with a greater number of bullish tweets precede hours with higher bitcoin returns; this relationship is statistically significant. Hours marked by increases in the number of bearish tweets precede hours with lower bitcoin returns, but this relationship is insignificant. Forum variables instead have no impact on future hourly bitcoin returns. Changes in the hourly tweet volume and forum posting volume both imply future bitcoin volatility. In particular, more bearish tweets and more bullish messages each precede hours in which we find higher bitcoin absolute returns. These findings are consistent with our hypotheses.

Table 7: Forum versus Twitter

Panel a: Daily Frequency (April 18–August 18, VECM, Model 3)

Dependent Variable	Independent Variable			
	ΔPOS_{t-1}^{F}	ΔNEG_{t-1}^{F}	ΔPOS_{t-1}^{T}	ΔNEG_{t-1}^{T}
r_t	2.87***	-1.79*	-1.08	1.02
$\Delta\sigma_t^2$	-2.46***	-0.11	-0.07	1.24
ΔV_t	-0.82	0.03	-0.18	0.34
ΔV_t^{TX}	-0.71	1.52	0.17	0.92

Panel b: Hourly Frequency (April 18–August 18, VAR, Model 4)

Dependent Variable	Independent Variable			
	ΔPOS_{t-1}^{F}	ΔNEG_{t-1}^{F}	ΔPOS_{t-1}^{T}	ΔNEG_{t-1}^{T}
r_t,	-0.16	0.55	1.77*	-0.41
$\lvert r_t \rvert$	2.64***	-1.55	-0.03	2.15**
ΔV_t	2.41**	0.24	0.31	0.29

Notes: ***, **, * denote significance at the 0.01, 0.05, and 0.1 levels, respectively.

Discussion and Conclusion

In this study we examined the dynamic relationship between social media metrics and bitcoin market measures. The results suggest that social media is an important indicator of future bitcoin returns. First, a positive shock of bullish posting indicates positive bitcoin returns on the next day, and a positive shock of bearish posting affects returns negatively on the next day (H1a). Second, disagreement induces the trading of bitcoin: greater disagreement across messages precedes higher bitcoin exchange trading volume on the next day (H1b). Third, message posting has a significant impact on bitcoin transaction volume. From an investment perspective, Internet forum posts thus can be viewed as bitcoin popularity indicators, offering investors means to understand the price movements of bitcoin.

Our analysis of the dynamic relationships between social media and bitcoin returns reveals interesting differences at the social influence level and identifies the unique effects of different social media platforms. In line with H2, we find that integrating an influencer analysis and stratifying the sample by user characteristics can reveal new insights on the bitcoin market. For example, the numbers of bullish and bearish tweets by all users have negligible effects on bitcoin returns in the next hour. However, if we limit the sample to tweets from those users with the most followers, the relationship becomes significant. This finding suggests that follow-the-influencer behavior exists in the bitcoin market, extending prior finance and marketing findings to a new setting (Brown et al. 2009; Hinz et al. 2011; Jiao and Ye 2013; Libai et al. 2010). On the other hand, posts from the most active Internet forum contributors have weaker association with future bitcoin returns, as compared to posts from the silent, but salient majorities. Therefore, to leverage the effects of social media on bitcoin markets, investors should analyze empirical data and carefully select the most influential people in a social network to collect information.

The relative effects of forum messages and tweets, using both daily and hourly data, are also distinct (H3). Forum metrics have stronger effects on future bitcoin returns at a daily frequency; Twitter metrics, on the other hand, are more important at an hourly frequency. A potential explanation is that the two types of social media platforms differ in how information is shared and how users digest information. The predominantly mobile users of Twitter make tweets more visible to investors at an intraday level, and the market therefore responds in a more timely fashion. Internet forum members instead engage in more thorough discussions on positive or negative news. This in turn can lead to consolidated market reactions at the daily level.

With these findings, our study makes several unique contributions. First, we provide empirical evidence of social media's impacts on future bitcoin returns. Social media offer substantial information about bitcoin's acceptance among the general public, as well as daily fluctuations in its market sentiments. Investors thus can gain insights into bitcoin's value from this rich information source. We investgate how UGC from different users and platforms affect bitcoin performance, leading to our recommendation that investors choose appropriate information channels to gather intelligence for their specific investment decisions.

Second, this study contributes to research on the impact of social media more generally. UGC available through social media provides valuable and timely information for investors and business stakeholders. Prior studies note the relationship between digital user metrics and product sales (Chevalier and Mayzlin 2006; Dellarocas et al. 2007; Dhar and Chang 2009; Ghose and Yang 2009; Moe and Fader 2004), link social media to firm equity (Luo et al. 2013), or stock market performance (Tirunillai and Tellis 2012). We extend this stream by examining the relationship between UGC and decentralized virtual currency. Bitcoin's exclusively online presence provides an ideal setting to investigate the impacts of social media on this new market.

Third, while many studies rely on a single source, we employ multiple sources (Internet forum and microblogging site), various data formats (numeric and textual), and different facets of social media metrics (sentiment and influence). Our findings highlight the fact that all social media messages are not created equally and should not be treated equally, prompting calls for new measures to capture the distinct economic values of social media.

The cryptocurrency market is gaining momentum. An estimated 275 digital currencies are now in play, many of which cater to specific products or services (Mayer 2014). Whereas most prior research addresses technical issues (e.g., mining, security, privacy), the success of cryptocurrency relies on the credible demonstration of their business, economic, and societal value. We believe the IS community should take the lead in this emerging, interdisciplinary research area. This study provides an initial examination of the relationship between social media and bitcoin and sheds some new light on understanding the complex dynamics.

References

Antweiler, W., and Frank, M. Z. 2004. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *The Journal of Finance* (59:3), pp. 1259-1294.

Barber, B. M., and Odean, T. 2008. "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors," *Review of Financial Studies* (21:2), pp. 785-818.

Bird, S. 2006. "Nltk: The Natural Language Toolkit," in: *The COLING/ACL on Interactive Presentation Sessions*. Sydney, Australia: pp. 69-72.

Brito, J., and Castillo, A. 2013. *Bitcoin: A Primer for Policymakers*. Mercatus Center at George Mason University.

Brown, J. R., Fazzari, S. M., and Petersen, B. C. 2009. "Financing Innovation and Growth: Cash Flow, External Equity, and the 1990s R&D Boom," *The Journal of Finance* (64:1), pp. 151-185.

Brown, N. C., Wei, K. D., and Wermers, R. 2013. "Analyst Recommendations, Mutual Fund Herding, and Overreaction in Stock Prices," *Management Science* (60:1), pp. 1-20.

Chen, H., De, P., Hu, Y. J., and Hwang, B.-H. 2014. "Wisdom of Crowds: The Value of Stock Opinions Transmitted through Social Media," *Review of Financial Studies* (27:5), pp. 1367-1403.

Chen, L., Wang, W., and Sheth, A. P. 2012. "Are Twitter Users Equal in Predicting Elections? A Study of User Groups in Predicting 2012 Us Republican Presidential Primaries," in *Social Informatics*. Springer, pp. 379-392.

Cheung, C. M., and Thadani, D. R. 2012. "The Impact of Electronic Word-of-Mouth Communication: A Literature Analysis and Integrative Model," *Decision Support Systems* (54:1), pp. 461-470.

Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of marketing research* (43:3), pp. 345-354.

Das, S. R., and Chen, M. Y. 2007. "Yahoo! For Amazon: Sentiment Extraction from Small Talk on the Web," *Management Science* (53:9), pp. 1375-1388.

Dellarocas, C., Zhang, X. M., and Awad, N. F. 2007. "Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures," *Journal of Interactive Marketing* (21:4), pp. 23-45.

Dhar, V., and Chang, E. A. 2009. "Does Chatter Matter? The Impact of User-Generated Content on Music Sales," *Journal of Interactive Marketing* (23:4), pp. 300-307.

Duggan, M., and Brenner, J. 2013. "The Demographics of Social Media Users, 2012," Pew Research Center's Internet & American Life Project, Washington, DC. Available at http://www.lateledipenelope.it/public/513cbff2daf54.pdf

European Central Bank. 2012. "Virtual Currency Schemes." from http://www.ecb.europa.eu/pub/pdf/other/virtualcurrencyschemes201210en.pdf

Fitzgerald, B. R. 2014. "Data Point: Social Networking Is Moving on from the Desktop." Wall Street Journal. Retrieved on Sep 1, 2015, from http://blogs.wsj.com/digits/2014/04/03/data-point-social-networking-is-moving-on-from-the-desktop/

Friend, I., Blume, M., Crockett, J., and Fund, T. C. 1970. *Mutual Funds and Other Institutional Investors: A New Perspective*. McGraw-Hill.

Ghose, A., and Yang, S. 2009. "An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets," *Management Science* (55:10), pp. 1605-1622.

Godes, D., and Mayzlin, D. 2009. "Firm-Created Word-of-Mouth Communication: Evidence from a Field Test," *Marketing Science* (28:4), pp. 721-739.

Goldenberg, J., Han, S., Lehmann, D. R., and Hong, J. W. 2009. "The Role of Hubs in the Adoption Process," *Journal of Marketing* (73:2), pp. 1-13.

Granger, C. W., and Newbold, P. 1974. "Spurious Regressions in Econometrics," *Journal of econometrics* (2:2), pp. 111-120.

Hinz, O., Skiera, B., Barrot, C., and Becker, J. U. 2011. "Seeding Strategies for Viral Marketing: An Empirical Comparison," *Journal of Marketing* (75:6), pp. 55-71.

Hirshleifer, D., and Teoh, S. H. 2009. "Thought and Behavior Contagion in Capital Markets," in *Handbook of Financial Markets: Dynamics and Evolution*. pp. 1-46.

Jiao, Y., and Ye, P. 2013. "Mutual Fund Herding in Response to Hedge Fund Herding and the Impacts on Stock Prices." *Working Paper*, Available at http://faculty.ucr.edu/~yawenj/mfhf.pdf

Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Johansen, S., and Juselius, K. 1990. "Maximum Likelihood Estimation and Inference on Cointegration—with Applications to the Demand for Money," *Oxford Bulletin of Economics and statistics* (52:2), pp. 169-210.

Kamins, M. A., and Assael, H. 1987. "Two-Sided Versus One-Sided Appeals: A Cognitive Perspective on Argumentation, Source Derogation, and the Effect of Disconfirming Trial on Belief Change," *Journal of Marketing Research* (24:1), pp. 29-39.

Kristoufek, L. 2014. "What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis," in: *arXiv* 1406.0268v1. Available at http://arxiv.org/pdf/1406.0268v1.pdf

Libai, B., Bolton, R., Bügel, M. S., De Ruyter, K., Götz, O., Risselada, H., and Stephen, A. T. 2010. "Customer-to-Customer Interactions: Broadening the Scope of Word of Mouth Research," *Journal of Service Research* (13:3), pp. 267-282.

Loughran, T., and McDonald, B. 2014. "Measuring Readability in Financial Disclosures," *The Journal of Finance* (69:4), pp. 1643-1671.

Love, I., and Zicchino, L. 2006. "Financial Development and Dynamic Investment Behavior: Evidence from Panel Var," *The Quarterly Review of Economics and Finance* (46:2), pp. 190-210.

Luo, X., Zhang, J., and Duan, W. 2013. "Social Media and Firm Equity Value," *Information Systems Research* (24:1), pp. 146-163.

Mayer, A. 2014. "Bitcoin Has a Future, but Maybe Not as a Currency." CBCNews. Retrieved on Sep 14, 2014, from http://www.cbc.ca/news/technology/bitcoin-has-a-future-but-maybe-not-as-a-currency-1.2686045

Mehra, A., Dixon, A. L., Brass, D. J., and Robertson, B. 2006. "The Social Network Ties of Group Leaders: Implications for Group Performance and Leader Reputation," *Organization science* (17:1), pp. 64-79.

Metaxas, P. T., and Mustafaraj, E. 2012. "Trails of Trustworthiness in Real-Time Streams (Extended Summary)," in: *Design, Influence and Social Technologies Workshop (DIST 2012) at the 2012 ACM Conference on Computer Supported Cooperative Work*. Seattle, WA.

Moe, W. W., and Fader, P. S. 2004. "Dynamic Conversion Behavior at E-Commerce Sites," *Management Science* (50:3), pp. 326-335.

Mustafaraj, E., Finn, S., Whitlock, C., and Metaxas, P. T. 2011. "Vocal Minority Versus Silent Majority: Discovering the Opionions of the Long Tail," in: *2011 IEEE International Conference on Privacy, Security, Risk and trust and 2011 IEEE International Conference on Social Computing (PASSAT/SocialCom 2011)*. Boston, MA: IEEE, pp. 103-110.

PricewaterhouseCoopers LLP. 2014. "Digital Disruptor: How Bitcoin Is Driving Digital Innovation in Entertainment, Media and Communications ". Available at http://www.pwc.com/en_SG/sg/tmt/assets/tmtnews201402/digital_disruptor.pdf

Rogers, E. M. 2010. *Diffusion of Innovations*. Simon and Schuster.

Schwert, G. W. 1989. "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business & Economic Statistics* (7:2), pp. 147-159.

Shi, Z., Rui, H., and Whinston, A. B. 2014. "Content Sharing in a Social Broadcasting Environment: Evidence from Twitter," *MIS Quarterly* (38:1), pp. 123-A126.

Smyth, L. 2013. "The Demographics of Bitcoin." Sumulacrum. Retrieved on Sep 14, 2014, from http://simulacrum.cc/2013/03/04/the-demographics-of-bitcoin-part-1-updated/

Tafti, A., Zotti, R., and Jank, W. 2013. "Real-Time Diffusion of Information on Twitter and the Financial Markets," in: *The 2013 Conferece on Information Systems and Technology (CIST 2013)*. Minneapolis, MN.

Tirunillai, S., and Tellis, G. J. 2012. "Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance," *Marketing Science* (31:2), pp. 198-215.

Trusov, M., Bodapati, A. V., and Bucklin, R. E. 2010. "Determining Influential Users in Internet Social Networks," *Journal of Marketing Research* (47:4), pp. 643-658.

Tumarkin, R., and Whitelaw, R. F. 2001. "News or Noise? Internet Postings and Stock Prices," *Financial Analysts Journal* (57:3), pp. 41-51.

Wysocki, P. D. 1999. "Cheap Talk on the Web: The Determinants of Postings on Stock Message Boards." *University of Michigan Business School Working Paper*, No. 98025. Available at SSRN: http://ssrn.com/abstract=160170

Yermack, D. 2013. "Is Bitcoin a Real Currency? An Economic Appraisal." *National Bureau of Economic Research Working Paper Series*, No. 19747. Available at http://www.nber.org/papers/w19747