

Characteristics of Bitcoin users: an analysis of Google search data

Aaron Yelowitz* and Matthew Wilson

*Department of Economics, University of Kentucky, Lexington, KY
 40506-0034, USA*

The anonymity of Bitcoin prevents analysis of its users. We collect Google Trends data to examine determinants of interest in Bitcoin. Based on anecdotal evidence regarding Bitcoin users, we construct proxies for four possible clientele: computer programming enthusiasts, speculative investors, Libertarians and criminals. Computer programming and illegal activity search terms are positively correlated with Bitcoin interest, while Libertarian and investment terms are not.

Keywords: Bitcoin; digital currency; Google search data; Libertarians; illegal activity

JEL Classification: E42; F33; K42; K49

I. Introduction

Bitcoin, a virtual global currency, has been the topic of much media, Internet and policy discussion. Over 13.4 million Bitcoins are in circulation and have a total market value of \$4.6 billion.¹ Little is known about the characteristics of Bitcoin users, even though thousands of businesses accept Bitcoins as payment. Transactions with Bitcoin are near anonymous due to the cost associated with identifying a user's electronic signature. Although some convenience sampling exists of Bitcoin enthusiasts, no systematic data collection has been done.

We use Google Trends (hereafter, 'GT') data to study the clientele driving interest in Bitcoin, with the caveat that search query interest need not imply active participation. Based on anecdotal evidence about Bitcoin users, we construct proxies for four possible clientele: computer programming enthusiasts,

speculative investors, Libertarians and criminals. Illegal activity and computer programming are both positively associated with Bitcoin use, while no association exists for Libertarian ideology or investment motives in most specifications.

II. The Bitcoin Market

Bitcoin was created in 2009 as an unregulated, alternative method of exchange for online payments. Upon signing up for an account, an individual receives an electronic signature that secures transactions and disallows double spending (enforced by a diverse computer network). This process circumvents conventional methods that involve trust in and fees to a third party. Conventional methods involve third-party fees, deterring small transactions (Nakamoto, 2008).² Anonymity is theoretically

*Corresponding author. E-mail: aaron@uky.edu

¹ <https://blockchain.info/charts/total-bitcoins>

² <https://bitcoin.org/bitcoin.pdf>

50 achieved due to Bitcoin’s encryption, with the sole link being the electronic signature. Meiklejohn *et al.* (2013) find that anonymity is nearly impossible with large-scale transactions, but there are high costs to identifying users.

55 III. Who Might be Bitcoin Users?

Profit and politically charged aspirations coincide with the basic design of the Bitcoin market. Prices for Bitcoins have fluctuated enormously over time, which might prove tempting for a speculative investor. The unregulated set-up makes it appealing to Libertarians who philosophically oppose ‘inflationary central-bank meddling.’³ Other clientele appreciate Bitcoin’s market structure for different reasons. For example, Bitcoin has appeal among computer programmers; ‘miners’ (the term for those seeking to discover new Bitcoins) can earn the currency in exchange for utilizing special software to authenticate real-time Bitcoin transactions.⁴ The anonymity of Bitcoin is attractive for criminal activity. The 2 October 2013 FBI takedown of the Silk Road website – an online marketplace ‘for everything from heroin to forged passports’ where transactions took place in Bitcoins – highlighted the importance of Bitcoin’s perceived anonymity and led to a 22% reduction in Bitcoin’s price.⁵

80 In order to understand the underlying rationale for Bitcoin use, Lui (2013) surveyed 1133 members of the Bitcoin community (by posting links on Bitcoin websites).⁶ The survey identified three key motives: curiosity, profit and political. Respondents (which included both owners and nonowners of Bitcoin)

³ <http://www.economist.com/news/finance-and-economics/21599053-chronic-deflation-may-keep-bitcoin-displacing-its-fiat-rivals-money>

⁴ <http://www.bitcoinmining.com/>

⁵ <http://online.wsj.com/articles/SB10001424052702303722604579115692946177328> and <https://www.tradingview.com/v/4xVX2cFq/>

⁶ <http://simulacrum.cc/2013/04/13/overview-of-bitcoin-community-survey-feb-mar-2013/>

⁷ We started in January 2011 because GT better measures state-level search activity from that point. We ended in July 2013 because the ‘Silk Road’ website – unknown to most of the public – was shut down soon after and made front-page headlines in national publications.

⁸ GT data have been predictive of behaviour in diverse economic markets including entertainment, labour and housing (Askitas and Zimmerman, 2009; Varian and Choi, 2009; Hand and Judge, 2012; Wu and Brynjolfsson, 2013). It has also been used for detecting health patterns, including influenza outbreaks and Lyme disease cycles (Ginsberg *et al.*, 2009; Carneiro and Mylonakis, 2009; Seifter *et al.*, 2010).

⁹ He shows that cross-sectional state variation in GT is highly correlated with other data sources; for example the search rate for the word ‘God’ explains 65% of the variation in the percentage of a state’s residents believing in God.

¹⁰ <https://support.google.com/trends/answer/4355000?hl=en>

¹¹ We attempted to use alternative terms for these concepts (such as ‘Libertarian’ or ‘Ron Paul’ for Libertarianism), but search interest was either too sparse or had a strong political cycle.

are likely unrepresentative of the larger community; for example, those using Bitcoin for illegal activity are unlikely to participate.

IV. GT Data

We collected GT search query data from January 2011 to July 2013 for all US states and Washington DC.⁷ We looked for terms related to Bitcoin and its possible clientele.⁸ Some of these correlations are inherently difficult to measure, due to the sensitivity of the activity; Stephens-Davidowitz (2013, 2014) argues, however, that Google data are unlikely to suffer from major social censoring, and uses GT to explore child abuse and racial animus.⁹ Although it is conceivable that higher Bitcoin search volume need not translate into increased market participation, Kristoufek (2013) demonstrates a strong positive correlation between Bitcoin searches and exchange prices.

GT can be used to extract data for precise search terms and more general topics (see Fig. 1). Search terms will return data for the exact query while topics count related searches too.¹⁰ For instance, the topic ‘Bitcoin (Currency)’ includes the terms ‘Bitcoin’, ‘Bitcoins’, ‘Bitcoin Mining’, ‘Bit Coin’, ‘Bitcoin exchange’, ‘Bitcoin price’ and ‘Bitcoin value’. We use search topics for Bitcoin (under ‘Currency’) and Computer Science (under ‘Discipline’). For other clienteles – Illegal Activity, Libertarians and Speculative Investors – we use the search terms ‘Silk Road’, ‘Free Market’ and ‘Make Money’, respectively.¹¹

GT does not report raw search counts for a topic; such counts would be misleading because Google’s

85

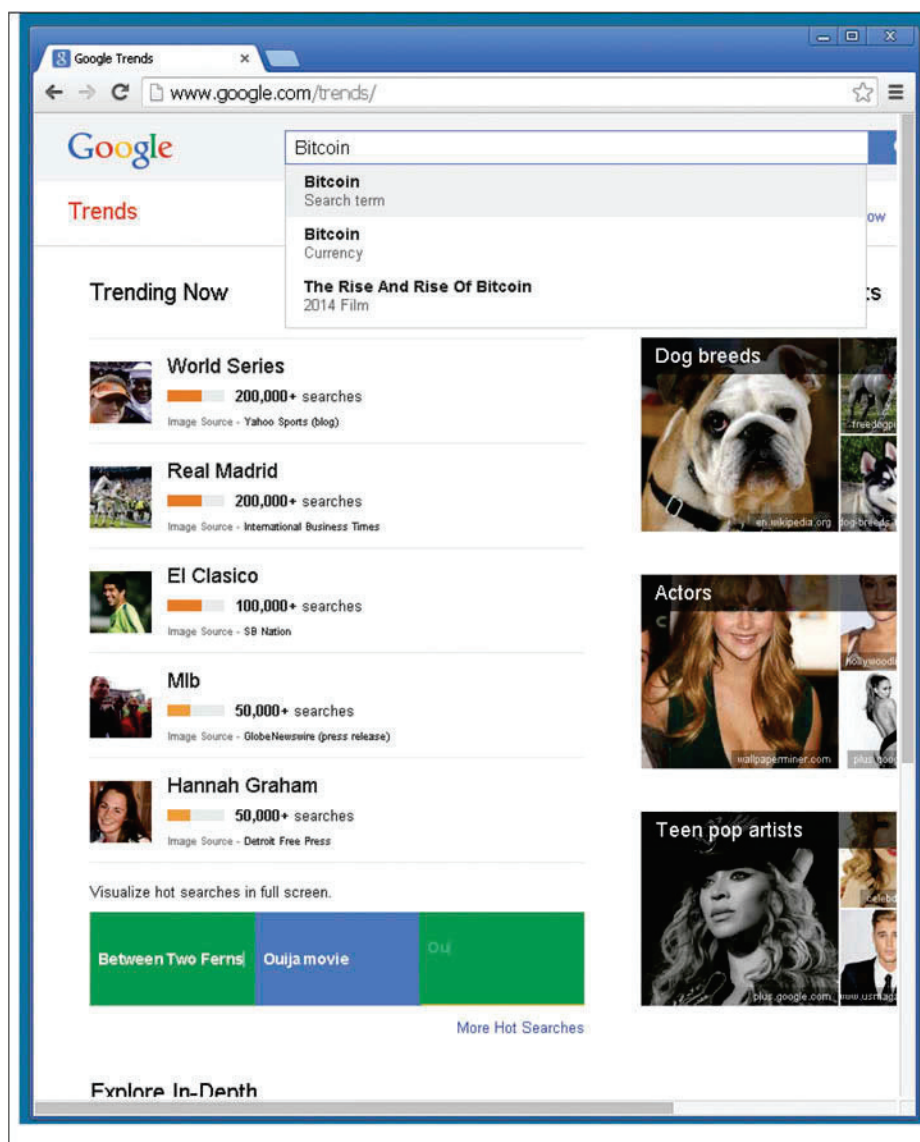
90

95

100

105

110



COLOUR ONLINE
BLACK & WHITE
IN PRINT

AQ1 **Fig. 1. Google ‘search term’ versus ‘topic (Currency)’**
Source: Google Trends (www.google.com/trends).

115 popularity (and search queries) grow over time.¹²
 Instead GT computes the number of topic searches
 relative to all searches, normalizes the series so the
 highest value is 100, and scales all other values
 relative to the highest. Figure 2 illustrates the
 Bitcoin time series in California, where popularity
 120 peaked in April 2013. For each state, we initially
 compute a 31-month time series for the relative
 popularity of Bitcoin and each clientele grouping.¹³

We then use GT to measure relative state-level popu-
 larity of each search term for the full period and scale
 each state-series relative to the most popular state. 125
 During the observed time frame, the states with the
 highest interest in Bitcoin were Utah, Oregon,
 California, Washington, Nevada, New Hampshire
 and Vermont (see Fig. 3). We then rescale each
 state-specific time series by its geographic 130
 popularity. Thus, using California’s value of 94

¹² <https://support.google.com/trends/answer/4365533?hl=en>

¹³ Some states and search terms had weekly activity (such as California’s Bitcoin activity in Fig. 2). In such cases, we computed monthly averages for all nonmissing values and then rescaled the series with a maximum value of 100.

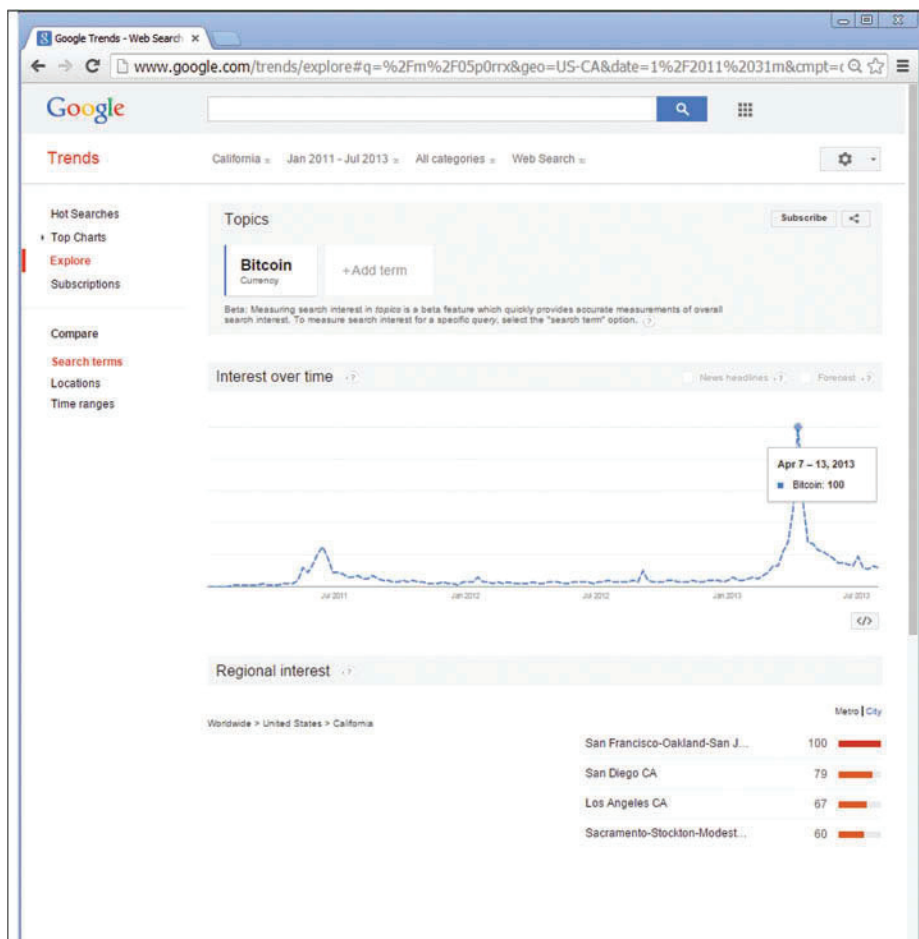


Fig. 2. Index for Bitcoin topic search California time series, January 2011–July 2013

from the geographic Bitcoin comparison, the entire California time series would be rescaled to 0.94 of its original value.

135 Our outlined methodology presents us with two limitations. First, GT samples its database and computes the index based on that sample.¹⁴ We observed slightly different values for the index by refreshing the web page, even with the same restrictions. Although the overall conclusions are unlikely to change from sampling, this prohibits exact replication. Second, GT gives a value of zero if it cannot gather enough data.¹⁵ We exclude state-month observations with missing values.
140 While every index has missing values for particular months, some states returned a missing value in the cross-sectional analysis, which prevents rescaling of the state-specific time series. Delaware, North

Dakota, and Wyoming were excluded as they had missing values for ‘Free Market’ and/or ‘Silk Road.’ Out of 1488 (48 states × 31 months) potential observations, our analysis uses 794 with non-missing values on Bitcoin, Computer Science, Free Market, Silk Road, and Make Money. The most populous states tend to have the fewest missing state-month observations.
155

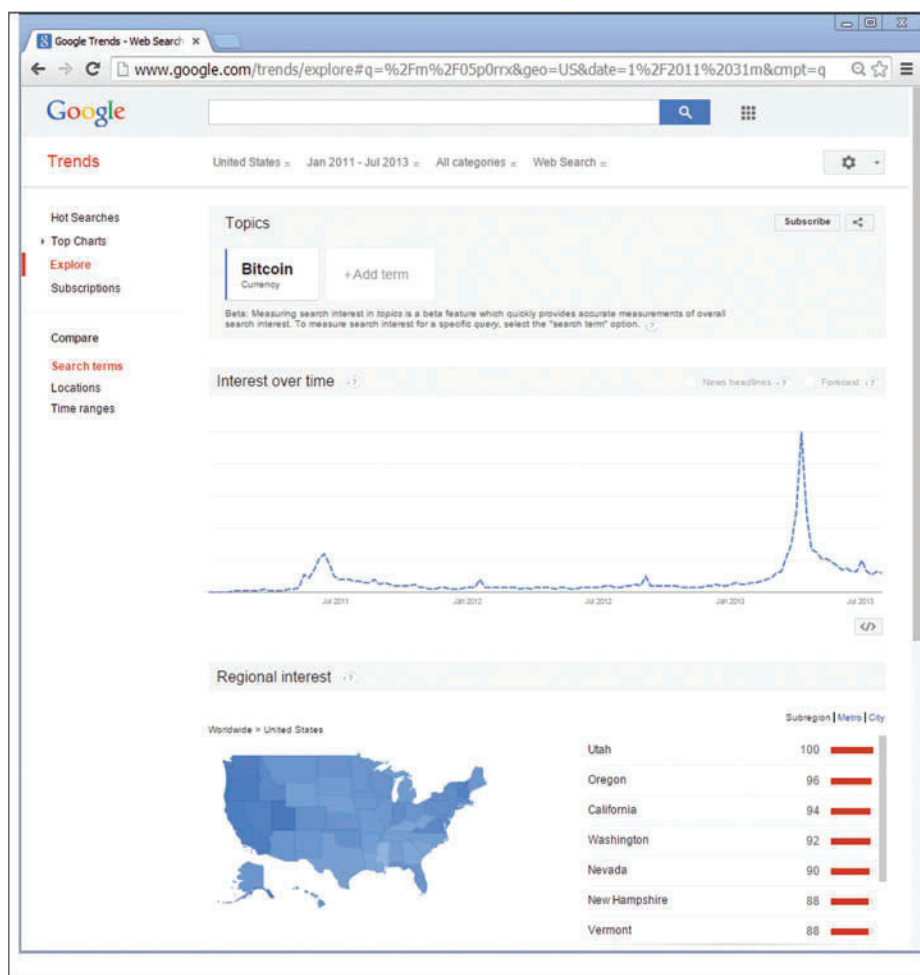
V. Empirical Results

Following Stephens-Davidowitz (2014), we normalize each search rate to its z-score and estimate the following specification:
160

$$\text{BITCOIN}_{jt} = \beta_0 + \beta_1 X_{jt} + \delta_j + \delta_t + \varepsilon_{jt} \quad (1)$$

¹⁴ https://support.google.com/trends/answer/4355213?hl=en&ref_topic=4365599

¹⁵ https://support.google.com/trends/answer/4355164?hl=en&ref_topic=4365531



COLOUR ONLINE
BLACK & WHITE
IN PRINT

Fig. 3. Index for Bitcoin topic search cross-sectional popularity, January 2011–July 2013

where BITCOIN_{jt} is Bitcoin interest in state j in month t , X_{jt} is clientele interest, and δ_j and δ_t are state and time fixed effects. Each state-month is weighted by state population in July 2011, and SEs are corrected for non-nested two-way clustering at the state and time levels (Cameron *et al.*, 2011). By including fixed effects in our fully saturated specification, the impact of clientele association on Bitcoin is measured through differential within-state changes over time (Yelowitz, 1995).

Results for a variety of specifications are presented in Table 1, Columns (1)–(3) progressively include additional controls for state and time. The inclusion of both state and time fixed effects identifies interest in Bitcoin by exploiting within-state changes over time. In this specification, interest in computer science and Silk Road is both positively associated with interest in Bitcoin and is

statistically significant at the 10% level. The interpretation of the specification in column (3) is the following: a one-SD increase in computer science interest leads to a 0.13 SD increase in Bitcoin interest, while a one-SD increase in Silk Road interest leads to a 0.09 SD increase in Bitcoin interest. Column (4) adds a ‘placebo clientele’ – searches for the singer Miley Cyrus. Reassuringly, inclusion of this placebo variable neither changes any of the inferences on the other clientele, nor is the variable itself significant.

Columns (5)–(6) interact each clientele search term with average monthly Bitcoin prices. Profit-motivated clientele – such as speculative investors – may find Bitcoin more intriguing when prices are high. However, we again observe a positive association between Bitcoin interest and our two clientele groups of computer programming enthusiasts and

165

170

175

180

185

190

195

Table 1. Determinants of Bitcoin search interest

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Computer Science	0.083 (0.066)	0.143 (0.173)	0.125 (0.073)	0.124 (0.073)	0.009 (0.029)	0.008 (0.028)	0.121 (0.059)	0.121 (0.059)	0.011 (0.027)	0.131 (0.064)	0.014 (0.030)	0.125 (0.065)
Computer Science X PRICE/100			0.208 (0.068)		0.208 (0.068)	0.209 (0.068)			0.205 (0.064)		0.202 (0.062)	
Silk Road	0.948 (0.374)	1.080 (0.408)	0.093 (0.051)	0.093 (0.052)	-0.007 (0.040)	-0.007 (0.040)	0.076 (0.039)	0.076 (0.040)	-0.012 (0.036)	0.105 (0.066)	0.010 (0.038)	0.088 (0.044)
Silk Road X PRICE/100			0.193 (0.101)		0.193 (0.101)	0.192 (0.100)			0.185 (0.097)		0.141 (0.082)	
Free Market	0.211 (0.076)	-0.172 (0.058)	0.023 (0.022)	0.023 (0.022)	-0.006 (0.022)	-0.005 (0.021)	0.031 (0.019)	0.031 (0.019)	0.003 (0.019)	0.036 (0.025)	0.004 (0.021)	0.021 (0.020)
Free Market X PRICE/100			0.043 (0.068)		0.043 (0.068)	0.047 (0.080)			0.030 (0.077)		-0.011 (0.073)	
Make Money	0.052 (0.089)	0.085 (0.121)	-0.004 (0.026)	-0.004 (0.026)	0.004 (0.026)	0.004 (0.025)	0.005 (0.030)	0.005 (0.029)	0.016 (0.026)	-0.041 (0.047)	0.003 (0.029)	0.006 (0.030)
Make Money X PRICE/100			-0.039 (0.070)		-0.039 (0.070)	-0.045 (0.075)			-0.069 (0.076)		-0.095 (0.075)	
Miley Cyrus				0.021 (0.040)		0.031 (0.080)		0.015 (0.040)	0.034 (0.075)			
Miley Cyrus X PRICE/100						0.010 (0.115)			0.007 (0.101)			
Unemp. Rate							-0.121 (0.064)	-0.121 (0.064)	-0.080 (0.051)	-0.281 (0.097)	-0.203 (0.072)	-0.105 (0.063)

Notes: Sample size is 794 in columns (1)–(9), 591 in columns (10) and (11) (2012 onward) and 580 in column (12) (states with ≥ 20 observations). SEs corrected for non-nested, two-way clustering at the STATE and MONTH levels. Observations weighted by population. State and time fixed effects included in columns (3)–(12). State fixed effects and a time trend included in column (2).

those possibly engaged in illegal activity (in the interaction term, not the main effect). The other clientele groups remain insignificant.

Columns (7)–(9) include the state-level monthly unemployment rate. Columns (7)–(8) show that the inferences on computer science and illegal activity are unchanged, but there is some evidence that Libertarian activity also drives interest in Bitcoin (although the specification including interactions with Bitcoin prices is insignificant). Higher unemployment rates are negatively associated with Bitcoin interest. Columns (10)–(11) estimate the model from 2012 onwards (when Bitcoin was more popular), while column (12) estimates it for the 24 states with at least 20 monthly observations. In all cases, fluctuations in computer science and illegal activity continue to drive Bitcoin interest, as well as the business cycle.

VI. Discussion

Although many commentators have speculated about motives for using Bitcoin, our study is the first to systematically analyse Bitcoin interest, including the interest of hard-to-observe clientele. We find robust evidence that computer programming enthusiasts and illegal activity drive interest in Bitcoin and find limited or no support for political and investment motives.

References

- Askitas, N. and Zimmerman, K. F. (2009) Google econometrics and unemployment forecasting, *Applied Economics Quarterly*, **55**, 107–20. doi:10.3790/aeq.55.2.107
- Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2011) Robust inference with multiway clustering, *Journal of Business & Economic Statistics*, **29**, 238–49. doi:10.1198/jbes.2010.07136
- Carneiro, H. A. and Mylonakis, E. (2009) Google Trends: a web-based tool for real-time surveillance of disease outbreaks, *Clinical Infectious Diseases*, **49**, 1557–64. doi:10.1086/630200

- Ginsberg, J., Mohebbi, M. H., Patel, R. S., *et al.* (2009) Detecting influenza epidemics using search engine query data, *Nature*, **457**, 1012–14. doi:10.1038/nature07634
- Hand, C. and Judge, G. (2012) Searching for the picture: forecasting UK cinema admissions using Google Trends data, *Applied Economics Letters*, **19**, 1051–55. doi:10.1080/13504851.2011.613744
- Kristoufek, L. (2013) BitCoin meets Google Trends and Wikipedia: quantifying the relationship between phenomena of the Internet era, *Scientific Reports*, **3**, 1–7. doi:10.1038/srep03415
- Lui, S. (2013) The demographics of Bitcoin (Part 1 Updated). Available at <http://simulacrum.cc/2013/03/04/the-demographics-of-bitcoin-part-1-updated/> (accessed 20 December 2014).
- Meiklejohn, S., Pomarole, M., Jordan, G., *et al.* (2013) A fistful of Bitcoins: characterizing payments among men with no names. In *Proceedings of the 2013 conference on Internet measurement conference*, IMC'13, 127–40.
- Nakamoto, S. (2008) Bitcoin: a peer-to-peer electronic cash system. Available at <https://bitcoin.org/bitcoin.pdf> (accessed 20 December 2014).
- Seifter, A., Schwarzwalder, A., Geis, K. *et al.* (2010) The utility of 'Google Trends' for epidemiological research: Lyme disease as an example, *Geospatial Health*, **4**, 135–37.
- Stephens-Davidowitz, S. (2013) Unreported victims of an economic downturn, Working Paper. Available at <http://static.squarespace.com/static/51d894bee4b01caf88ccb4f3/t/51e22f38e4b0502fe211fab7/1373777720363/childabusepaper13.pdf>.
- Stephens-Davidowitz, S. (2014) The cost of racial animus on a black candidate: evidence using Google search data, *Journal of Public Economics*, **118**, 26–40. doi:10.1016/j.jpubeco.2014.04.010
- Varian, H. R. and Choi, H. (2009) Predicting the present with Google Trends. Google Research Blog Available at <http://googleresearch.blogspot.com/2009/04/predicting-present-with-google-trends.html>. Available at SSRN: <http://ssrn.com/abstract=1659302> (accessed 20 December 2014).
- Wu, L. and Brynjolfsson, E. (2013). The future of prediction: how Google searches foreshadow housing prices and sales. Available at SSRN: <http://ssrn.com/abstract=2022293>.
- Yelowitz, A. (1995) The Medicaid notch, labor supply and welfare participation: evidence from eligibility expansions, *The Quarterly Journal of Economics*, **110**, 909–39. doi:10.2307/2946644