ARTICLE IN PRESS

International Journal of Forecasting ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Can we vote with our tweet? On the perennial difficulty of election forecasting with social media

Mark Huberty

San Jose, CA, USA

ARTICLE INFO

Keywords:
Evaluating forecasts
Surveys
Model selection
Elections
Social media

ABSTRACT

Social media and other "big" data promise new sources of information for tracking and forecasting electoral contests in democratic societies. This paper discusses the use of social media, and Twitter in particular, for forecasting elections in the United States, Germany, and other democracies. All known forecasting methods based on social media have failed when subjected to the demands of true forward-looking electoral forecasting. These failures appear to be due to fundamental properties of social media, rather than to methodological or algorithmic difficulties. In short, social media do not, and probably never will, offer a stable, unbiased, representative picture of the electorate; and convenience samples of social media lack sufficient data to fix these problems post hoc. Hence, while these services may, as others in this volume discuss, offer new ways of *reaching* prospective voters, the data that they generate will not replace polling as a means of assessing the sentiment or intentions of the electorate.

1. Introduction

The virtualization of human behavior has generated enormous quantities of data about what people do or say, to whom, and when. The presence of such data has naturally given rise to the desire to study past behavior, and ultimately, to forecast the future. High-profile commercial successes in advertising, sales, logistics, and other fields hint that it may be possible to fulfill this desire. If so, this offers a potential solution to political professionals and scholars who are struggling with the increasing cost and difficulty of traditional survey research and electoral polling.

However, our experience to date in forecasting important events, political or otherwise, has not always borne out the promise of new predictive powers. Success in predicting what people will want to buy, or what ads they might click on, has not translated into reliable predictions in sectors such as public health or finance. Research into other means of online opinion discovery, such as consumer rating systems, has shown them to be vulnerable to systematic bias. These difficulties have cast doubts on the potential for social media data to provide a new means of forecasting high-profile real-world events.

This paper discusses the history of electoral forecasting from social media and other online data. I note that all known examples of apparently successful forecasts have very quickly encountered problems that have undermined their predictive power. Their initial success at beating simple heuristics for electoral success, such as incumbency, has faded quickly when faced with the challenge of forward-looking election prediction. These difficulties have led to high-profile claims that social media data will soon supplant polling having a short shelf life. Hence, while social media and other online data may be useful for studying how citizens behave online, they have proven rather less useful for forecasting what those same citizens – and, more importantly, their offline fellows – will do back in the real world.

I conclude by discussing whether these problems can be overcome. I agree with earlier critiques of the election forecasting literature, which have recommended fundamental changes in the transparency, disclosure, method, and

E-mail address: mark.huberty@gmail.com.

http://dx.doi.org/10.1016/j.ijforecast.2014.08.005
0169-2070/

Please cite this article in press as: Huberty, M., Can we vote with our tweet? On the perennial difficulty of election forecasting with social media. International Journal of Forecasting (2015), http://dx.doi.org/10.1016/j.ijforecast.2014.08.005

publication of electoral forecasts that rely on social media. However, I raise the question of whether we should ever expect reliable election forecasts even after such changes. Because social media forecasting lacks the control over the data-generating process which is enjoyed by polling, it faces enormous instability in the origin, sample, structure, and content of the data on which it relies. This instability should undermine easy claims that social media-based forecasts will replace polls any time soon. Instead, social media may only prove one more channel for the application of traditional polling methods, rather than a revolution in social prediction.

2. Forecasting the political

From what began as a catalogue of human knowledge, the internet has become an instrument for monitoring human behavior. Using the data from that instrument, new companies have successfully measured and predicted a range of online and offline behaviors. The most notable successes include radically more effective ad targeting, translation, and email spam detection (Google), the discovery and recommendation of professional (LinkedIn) and social (Facebook) contacts, early detection of and marketing to people with specific medical conditions like pregnancy (Hill, 2012), and the real-time measurement of geographically-specific rates of influenza (Ginsberg et al., 2008). Some of these early successes (notably Google Flu, see Butler, 2013; Copeland et al., 2013) have later encountered difficulties. However, the potential for the future to be predicted through traces of the past left behind on servers around the world remains.

Political forecasting from social and online data gained popularity at the time of, and perhaps prompted by, these successes. Apart from its sheer novelty, several forces conspired to encourage scholars and political professionals to pursue this problem. The importance of online media grew after 2008: President Barack Obama's digital effort was hailed widely in both 2008 and 2012 (Levenshus, 2010; Scherer, 2012; Smith, 2009; Wallsten, 2010), and the Tea Party insurgency that brought the Republican party back into control of the US House of Representatives relied initially on online social media to circumvent traditional party channels (Williamson, Skocpol, & Coggin, 2011). At the same time, traditional public opinion surveys began to struggle with rising non-response rates and increased cell phone usage, which prompted a search for new means of measuring and forecasting political attitudes and intent (Boyle, Fleeman, Kennedy, Lewis, & Weiss, 2013; Christian, Keeter, Purcell, & Smith, 2010; Keeter, Kennedy, Dimock, Best, & Craighill, 2006; Kohut, Keeter, Doherty, Dimock, & Christian, 2012; Viera, Medway, Turner, & Marsh, 2013).

Social media services were among the most tempting targets. The year 2010 saw a wave of papers of the form "using ⟨social media source X⟩ to predict political contest Y" scattered across computer science, machine learning, and social science journals. With a few notable exceptions (Bond et al., 2012; Broockman & Green, 2013), Twitter has dominated this stream of studies. Unlike the other major online data companies, Twitter has made its data relatively freely available (albeit with heavy restrictions on their subsequent distribution). Among other options, Facebook appears to be open to social science research, but requires a greater up-front investment in relationship building. Google is largely closed to outside research collaboration, though aggregated data from Google Trends are available freely. Hence, Twitter-based forecasting faces far lower startup costs than the alternatives. The reasons for these differences, and their consequences for patterns of social science research, deserve a separate study of their own.

Social media studies often promised more than simply an alternative to traditional polling. Social media users appeared oddly insensitive to the usual social mores that inhibit survey respondents from expressing controversial or taboo opinions. If this insensitivity held at scale, then social media also promised a solution to the longstanding difficulty of measuring individuals' true attitudes towards taboo subjects like racism or homophobia.[1] While someone might not tell a phone survey worker that they voted against Barack Obama because he was black, many internet users appeared to have no qualms about broadcasting the same.[2] Whether that openness reflects the medium, or is an artifact of a selection process whereby political content on Twitter is dominated by more opinionated users, is unclear.

Unfortunately, political forecasting has had far less success in predicting the offline world from online behavior than its commercial predecessors. All known attempts at election forecasting with social media have failed. Some have failed simply because models that tested well on contemporaneous out-of-sample data performed poorly on future elections. Others have failed because their initial success relied on undisclosed data manipulation on the part of researchers.

Here, we present a selection of attempts and failures to forecast elections using the Twitter social media service. We discuss in detail our own experiences in what we believe to be the first attempt at multi-cycle forecasting of US House of Representatives elections. Again, as in other attempts, promising forecasting methods failed to beat simple heuristics like incumbency for predicting election outcomes. A close examination of the underlying data suggests the reason: changes to both the political landscape and the demographics and behavior of the Twitter user community altered the data-generating process in ways that the models could not account for. By implication, successful forecasting models must have the capability to check and update their assumptions about the real world dynamically.

[1] Whether social media are really any less prone to this is an open question. We might all be surprised at the blatant racism displayed by many individuals online, but of course this raises the question: are these individuals in fact this racist in real life? Or do they merely express such racism as a matter of conforming to the social norms of their online community? Wilson, Gosling, and Graham (2012) review research suggesting that users' behaviors on Facebook may in fact reflect their true selves. On the other hand, internet comment threads are famously divisive unless heavily policed (Binns, 2012).

[2] Stephens-Davidowitz (2013) exploits this hypothesis to estimate that racial animus cost Barack Obama 3%–5% of the national vote. Hard verification of this estimate remains difficult, for the very reason that this estimate is valuable: the inability to benchmark it against survey data of actual human subjects.

Please cite this article in press as: Huberty, M., Can we vote with our tweet? On the perennial difficulty of election forecasting with social media. International Journal of Forecasting (2015), http://dx.doi.org/10.1016/j.ijforecast.2014.08.005