Justin Stewart

A Description of Consumer Activity in Twitter

At least for the astute economist, the introduction of techniques from computational science into economics has and is continuing to change how economists understand some of the fundamental aspects of the discipline. Computation has enabled economists to gain new insights from data sources that have been previously ignored. In fact, novel economic results via computational methods are forcing economists to reconsider the following fundamental question: what is valid economic data? This question captures one of the frontiers of economics research, and computational methods offer a new approach for investigating this foreign territory. In this paper, I outline some novel economic results found by use of computational analysis on twitter data and then use them as a framework to present a further result about aggregate *consumer activity*. Finally, this result stimulates a discussion of the broader philosophical question: does the introduction of computation into economics yield insight into the possible limitations of classical economic analysis?

What is valid economic data anyway? A brief overview of this question will motivate the important implications and justify the significance of the following results. Classically, economists solely attribute meaning to 'choice' data. That is to say, any data that does not represent actual choices made by economic actors is classified as insignificant. For example, if a person goes to the store for breakfast and buys Cheerios cereal, then the agent has made an actual choice. If one asked the same person what would they buy if they were to go to the store for breakfast, and the agent replied Cheerios, this does not constitute an actual choice and hence is meaningless information. This distinction is crucial for economics. In short, it is essentially what enabled the axiomatization of economic theory making the science more amenable to mathematical analysis. Thus, I invite the reader to consider the possibly significant information lost by making this distinction.

For the most part, the vast amounts of information produced by social media falls into the category of non-choice data. Facebook and Twitter users are continually posting their opinions and updating their statuses, effectively creating social data, but are nonetheless not making any actual choices beyond deciding to use some particular social media site. Thus, classical economists do not view this as a valuable source of information. However, with the help of computation, this is slowly starting to change. In the last few years, as the online presence and importance of social media has dramatically increased, more have started to take it seriously as a source of data. The first studies done using this data mostly pertained to the dynamics of the data itself (e.g. how many people use twitter, what are popular twitter trends, etc). Also, twitter data has spurred new methods of textual analysis that enables one to answer other interesting questions.

However, it is still relatively unclear what the most important information that this new source of data contains. The few studies that connect Twitter data to economic phenomena consistently have preformed large-scale sentiment analysis.

Thus, the general consensus agrees that the most important feature of the data relevant to economics is how people feel as represented by their Twitter posts. Although there have been few other similar studies, I consider "Twitter mood predicts the stock market" by Johan Bollen et al. (2011) to be the representative study of this perspective. Bollen's study preformed textual analysis of Twitter feed using mood-tracking tools and created an aggregate measure of the national mood along six different mood dimensions. He then compared these mood time series to stock market indexes like Dow Jones Industrial Average closing values and the S&P 500. Bollen found that the calm mood index was predictive of the up or down movement of the DJIA closing values three days in advance with an accuracy of 87.6%.

Though the results shown in Bollen's study were surprising enough the get a story in the New York Times, the final result is not the most essential point to take away from their research. What Bollen's group did was successfully connect twitter data to an aggregate economic indicator. This is a deep and meaningful connection to establish, in fact far more meaningful than between twitter and specifically the DJIA. To successfully show this result opens the door of possibilities for further connections to other economic aggregates. In fact, the space of possibility is large enough to entice one to reconsider the established central feature of Twitter data, the possibility of capturing the emergent national mood. Is there other significant information contained in Twitter feed that is relevant to other economic phenomena?

Clearly, the notion that the aggregate mood affects important aspects of our society is an intuitively pleasing one, however, Twitter constantly is producing much more specific information that just how people feel. In contrast, consider the possibility that some tweets describe not just how a person feels but rather what a person is actually doing. Asur and Huberman (2010) actually implemented this idea by demonstrating that the rate of tweets produced about certain movies could be used to forecast box office revenues. The twitter measure was shown to be 97.3% accurate in predicting opening weekend box office sales, and actually out performed all the standard measures. The fundamental realization that motivates this result is that Asur and Huberman clearly realized that the essential feature indicating consumption of movies is simply captured by tweets that mention the movie. Thus, Twitter data has been shown to also contain a meaningful emergent description of actual economic activity in addition to an emergent mood.

Given the door opening result established already in Bollen's paper, the obvious question to now ask is if Asur and Huberman's result about movie consumption can be generalized to a larger scale. That is, does Twitter feed also contain an emergent description of aggregate or national consumer activity? If it did, certainly there is a large potential that the measure would be related in some way to other economic aggregates like the consumer spending and economic confidence indexes.[1] Thus, one considers how the essential features of aggregate consumer

---

[1] Note the important distinction between economic confidence and actual consumer activity.

activity would manifest in Twitter data. In other words, movie mentions is to box office sales as *X* is to general consumer activity.

“For one day at least, you could almost imagine the recession never happened. […] the busiest Black Friday ever” (CBS/news). I claim that if the essential features of consumer activity do indeed manifest in Twitter data, then the information would certainly be present during period of time around Thanksgiving and Black Friday. Thus, using text analysis tools, I closely examined the tweets during this period of extreme consumerism and came up with a list of words and short phrases that best indicated consumer activity. The top few words on the list are ‘buy’, ‘shopping’, ‘Black Friday’, and ‘gift’. Each of these words used within a tweet more than likely described a person’s intent to consume. For example, a tweet might read: “going Black Friday shopping!” Clearly, considering the use of ‘Black Friday’ in all tweets would not necessarily be indicative of consumerism all year round. Further analysis suggested that the word ‘buy’ best captured enough of the essential features of consumption activity, and was used consistently enough to analyze longer time interval.

Thus, I created what I termed the Intention to Consume Index. I implemented a simplistic model to count the number of times ‘buy’ is present in tweets on a given day. Of the specification possibilities (e.g. counts, frequencies, term freq. inverse document frequencies) the counts model worked best. This data was then used to create a time series that spanned from September 6th 2011 to December 6, 2011. This allowed for a comparison of the Intention to Consume Index with other economic indicators over the same time frame. The graph below depicts this comparison for the ITC index, Consumer Spending (3-day rolling average), and Economic Confidence index (3-day rolling average).

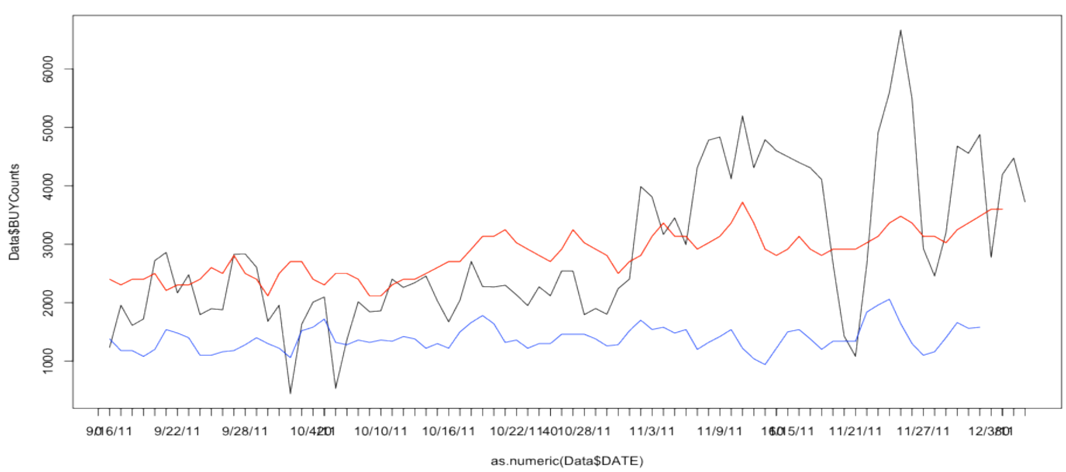

Figure 1: A comparison of the Intention to Consume index (black), Economic Confidence index with a two-day lag (red), and Consumer Spending with a four-day lag (blue).

It is important to note that each different time series originally had three vary different scales on the Y-axis and had to be translated to fit on the same graph.[2] Thus one best interprets this graph by only comparing the three indexes by how often they are moving in the same direction. By following the time series' over the interval, one will find that in many cases the peaks and troughs align. It suggests there is a nontrivial amount of correlation. In particular, observe the large spike in all three indexes at Thanksgiving and Black Friday (11/24-25/11). The crucial point to pay attention to when interpreting this graph is that indexes have different lags. This means that when all the spikes align on the graph, in actuality the ITC spike happened two days before the EC3 spike, and four days before the CS3 spike. This suggests that in certain cases the Intention to Consume index is predictive of the direction movement of the other two indices.

	EC3 (2-lag)	CS3 (4-lag)	CS14	ITC
EC3 (2-lag)	1	.289	.689	.640
CS3 (4-lag)	.289	1	.422	.305
CS14	.689	.422	1	.525
ITC	.640	.305	.525	1

Table 1: A correlation table that lists the correlations between the Economic Confidence Index with a 2-day lag, the 3-day rolling average of Consumer Spending with a 4-day lag, the 14-day rolling average of Consumer Spending, and the Intention to Consume Index.

Above, Table 1 displays the actual correlations between each index. I have include in this correlation table the 14-day rolling average of consumer spending in addition to the others for more in depth comparison. The Intention to Consume index is most highly correlated with the Economic Confidence index with the 2-day lag at a correlation of .640 and less correlated with Consumer Spending with 2-day lag at .305. Further, I examined the correlations using other variations of the time series that may be more strongly correlated, for example first differences and positive or negative movement in the following period. However the above table displays the strongest correlations found. Given the somewhat low correlations with the Intention to Consume Index, I urge the reader to remember the simplistic nature of the model. This model has shown that a time series composed of the number of times people tweet the word 'buy' per day has non-trivial correlation with a significant economic index. This is not a disappointment but rather demonstration of great potential.

The next step is to develop the model. Although the word 'buy' certainly captured some of the essential features of national consumer activity, it is by no means the whole picture. I expect that the results would dramatically improve by expanding the lexicon used in the ITC index. For example, one should also include the other terms found to be significant indicators of intentional consumption, like

[2] The Economic Confidence and Consumer Spending indices were translated linearly to match the scale of the ITC index. Note that linear transformations of the data preserve the relevant directional movement information being considered.

‘shopping’ and ‘gift’. In addition to other words that capture positive consumer activity, it would be optimal to also include words that suggest *negative* consumer activity. For example, words that suggest frugal consumer behavior, like ‘saving’, could strengthen the ITC index’s ability to capture times of low consumer activity. Further, sophistication of the model would include weighting the terms differently, potentially by co-occurrences in tweets and other methods.

However, examining another potential way to expand the lexicon both suggests direction for further exploration of the model and drives at some of the deeper questions posed at the outset of this paper. Consider that one could construct a genetic algorithm that searches the twitter feed for emergent word usage and is specifically programmed to optimize the correlation relationship between the ITC and other economic indexes. Thus, one could construct the lexicon for the ITC index on the basis that the words yield extremely highly correlation with lagged values of any index you wish. Further, another GA could be programmed to find the optimal weighting arrangements of terms in the lexicon to magnify this result. These possibilities seem particularly exciting, however I am hesitant to pursue this route after considering some of the greater implications involved in the construction of the Intention to Consume Index. Namely, the ITC index is intended to capture a qualitatively different economic phenomenon and not just predict other economic indicators.

Thus, reconsidering the questions concerning valid economic data and how the use of computation has affected our conceptions about economics in this new context proves insightful. Clearly, the use of twitter data has produced novel economic results as discussed in the earlier studies. Is this enough to qualify twitter data as economic? In addition, consider the possibility that certain economic phenomena are best described by intrinsically emergent properties. This paper has argued that the intention to consume in the aggregate is an important macroeconomic phenomenon, and moreover, the intrinsically non-choice nature of the data necessary to reveal its emergence cannot be described by classical economic analysis. This suggests probable limitations to what classical economic analysis can in principle describe. Clearly, as seen by the results discussed in this paper, the introduction of computational methods can help circumvent these limitations.

References

Asur, S & Huberman, B. A. (2010). Predicting the Future with Social Media. http://arxiv.org/abs/1003.5699.

Bollen, Johan, Huina Mao, and Xiaojun Zeng. "Twitter mood predicts the stock market. "*Journal of Computational Science*. 2. (2011): n. page. Web. 16 Dec. 2011. <www.elsevier.com/locate/jocs>.

E. Gilbert, K. Karahalios, Widespread worry and the stock market, in: Fourth International AAAI Conference on Weblogs and Social Media, Washington, DC, 2010, pp. 58-65, http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/download/1513/1833

A. Pak, P. Paroubek, Twitter as a corpus for sentiment analysis and opinion finding, in: N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, D. Tapias (Eds.), Proceedings of the Seventh conference on international Language Resources Association (ELRA), Valletta, Malta, May 2010, pp. 19-21

X. Zhang, H. Furhres, P.A. Gloor, Predicting the Stock Market through Twitter 'I hope it is not as bad as I fear', Collaborative Innovation Networks (COINs), Savannah, GA, 2010, http://www.elsevier.com/locate/procedia.