

PS11 McGuire: Using Geolocated Tweets to Predict Savings Rates

Joel McGuire

Donald Li

April 2018

1 Introduction

1.1 Future Orientation

Future orientation is a dimension of intrapersonal and intercultural psychology with ties to health and economic metrics. Evidence suggests it has the capacity to predict growth and savings rates on a national level as well as economic and health behaviors on an individual scale.

Time perspective is a fundamental psychological dimension humans use to frame experienced reality as belonging to the past, present, or future [16]. Whether one is more oriented towards the future or present is deeply related to self regulation [7]. Individuals who are more future orientated have been shown to be more likely to engage in numerous positive health behaviors, including moderation of ones diet, limited intake of mind altering substances, increased seat belt use, etc. Contrarily, individuals who are characterized as being more present orientated have been shown to be have a higher likelihood of engaging in risky health behaviors such as excessive consumption of alcohol, drugs and unprotected sex [4, 2]. These effects persist even when factors such as being a rapscallion youth are accounted for.

Among other factors such as financial literacy and risk preference, future orientation affects personal savings rate. Future orientated person's are more like to save both on an individual level [9], and as a country [6, 5]. Read, cited by Hofstede, found that a country scoring higher on an index of long term orientation is positively related to the marginal propensity to save [13].

1.2 Twitter as a Social Science Tool

The use of twitter for social science research has grown in recent years. Mitchel and colleagues used sentiment analysis to create a new index of happiness by city in the USA [10]. Most closely related in methodology is the the work done to predict HIV rates on a county level using an index of geolocated future oriented tweeted [8]. Twitter has also been used to predict instances of crime

[3], elections results [15], and of course future movements of the stock market [1].

1.3 Time Perspective and the Internet

Prior research has created a future orientation index from Google search terms of a countrywide level [12]. No one has yet pursued any research on seeing if an index of future orientation can be constructed from geolocated tweets for a more granular geographic region. We seek to find if such an index has any ability to predict factors related to future orientation on a state level, or metropolitan area.

In the vein of Chen’s work on how language itself changes inter-temporal choice, finding that ”languages that grammatically associate the future and the present foster future-oriented behavior” (2013), we hypothesize that areas with relatively future oriented tweets will have higher savings rates. A shortcoming of previous metrics of future orientation are an inability to capture changes in culture, primarily by the assumption of cultural being relatively invariant [14]. Benefits to this method are that it opens an avenue to the possibility of creating a dynamic understanding of cultural change through language, and an ability to compare that relationship with macroeconomic indicators.

2 Methodology

2.1 Data Source

Tweets were streamed using the Twitter API ($N = 2,007,230$) from February 2018 to April 2018. Collecting millions of geolocated tweets is a non trivial endeavor.

The dictionary of future oriented words was retrieved from LIWC.

GDP growth rate was retrieved from the Bureau of Economic Analysis. Most problematic is the current inability to match any macroeconomic indicator to this small window of twitter sentiment retrieved.

2.2 Analysis Strategy

A simple approach to natural language processing was used to create our future orientation index for the United States. Implementing the 2015 dictionary of the Linguistic Inquiry and Word Count (LIWC), an extensively validated tool for text analysis, was used to identify a set of future oriented words such as ”gonna”, ”will”, and ”hope” [11]. These words that have been previously identified as relating to a longer time perspective were aggregated per state. The index was created by counting the amount of future oriented words as a percent of total words written in all geolocated tweets in a state.

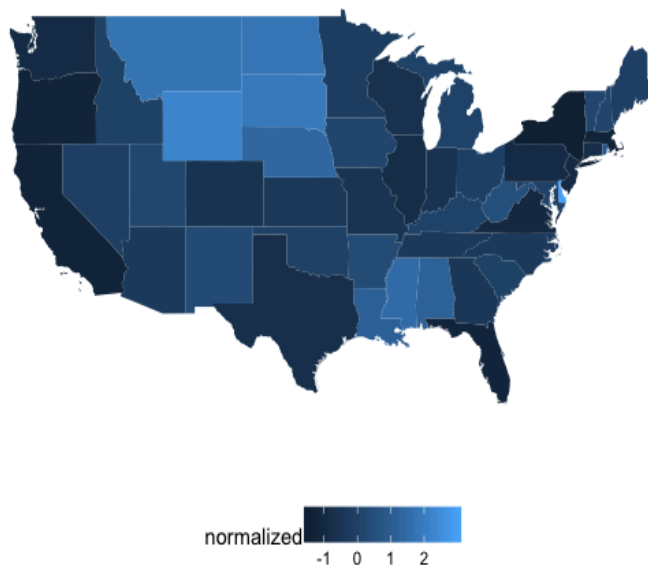
We need to create a better metric that weights each word according to its frequency without allowing an outlier to dominate the score for that state.

Even though we do naively regress the index of future orientation on the nearest data we have on GDP growth rate, we acknowledge there is little reason to suggest they would be related. We must wait for the data to catch up. Perhaps there would be more flexibility on a metropolitan level.

2.3 PROBLEMS

1. There is not any data for savings rate on the state level. Alternatives: Growth in investment related industry, financial services perhaps, what about change in personal consumption? Venture capital on the metropolitan level?
2. Each word is not weighted in accordance with its frequency.
3. There is no apparent relationship between the growth rate of the 3rd Quarter of 2017 and the future sentiment index of (mostly) the 2nd quarter of 2018. I am shocked.

2.4 Results



Above you can see that there is something. What it means we aren't exactly sure. States with large metropolitan populations appear less future oriented, other than that, who knows.

2.4.1 Real Results

So far the merits of this project have been in the creation of an easy to use set of wrapper functions for automating twitter collection in R.

References

- [1] Johan Bollen, Huina Mao, and Xiaojun Zeng. Twitter mood predicts the stock market. *Journal of computational science*, 2(1):1–8, 2011.
- [2] James R Daugherty and Gary L Brase. Taking time to be healthy: Predicting health behaviors with delay discounting and time perspective. *Personality and Individual differences*, 48(2):202–207, 2010.
- [3] Matthew S Gerber. Predicting crime using twitter and kernel density estimation. *Decision Support Systems*, 61:115–125, 2014.
- [4] James M Henson, Michael P Carey, Kate B Carey, and Stephen A Maisto. Associations among health behaviors and time perspective in young adults: Model testing with boot-strapping replication. *Journal of behavioral medicine*, 29(2):127–137, 2006.
- [5] Geert Hofstede. Cultures and organizations. intercultural cooperation and its importance for survival. software of the mind. *London: McGraw-Hill*, 2010.
- [6] Geert Hofstede and Michael Harris Bond. The confucius connection: From cultural roots to economic growth. *Organizational dynamics*, 16(4):5–21, 1988.
- [7] Elizabeth Howlett, Jeremy Kees, and Elyria Kemp. The role of self-regulation, future orientation, and financial knowledge in long-term financial decisions. *Journal of Consumer Affairs*, 42(2):223–242, 2008.
- [8] Molly E Ireland, H Andrew Schwartz, Qijia Chen, Lyle H Ungar, and Dolores Albarracín. Future-oriented tweets predict lower county-level hiv prevalence in the united states. *Health Psychology*, 34(S):1252, 2015.
- [9] Joy M Jacobs-Lawson and Douglas A Hershey. Influence of future time perspective, financial knowledge, and financial risk tolerance on retirement saving behaviors. *Financial Services Review*, 14(4):331, 2005.
- [10] Lewis Mitchell, Morgan R Frank, Kameron Decker Harris, Peter Sheridan Dodds, and Christopher M Danforth. The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PloS one*, 8(5):e64417, 2013.

- [11] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. The development and psychometric properties of liwc2015. Technical report, 2015.
- [12] Tobias Preis, Helen Susannah Moat, H Eugene Stanley, and Steven R Bishop. Quantifying the advantage of looking forward. *Scientific reports*, 2:350, 2012.
- [13] Russell Read. *Politics and policies of national economic growth*. Stanford University, 1993.
- [14] Linghui Tang and Peter E Koveos. A framework to update hofstede’s cultural value indices: economic dynamics and institutional stability. *Journal of International Business Studies*, 39(6):1045–1063, 2008.
- [15] Andranik Tumasjan, Timm O Sprenger, Philipp G Sandner, and Isabell M Welpe. Election forecasts with twitter: How 140 characters reflect the political landscape. *Social science computer review*, 29(4):402–418, 2011.
- [16] Philip G Zimbardo and John N Boyd. Putting time in perspective: A valid, reliable individual-differences metric. In *Time perspective theory; review, research and application*, pages 17–55. Springer, 2015.