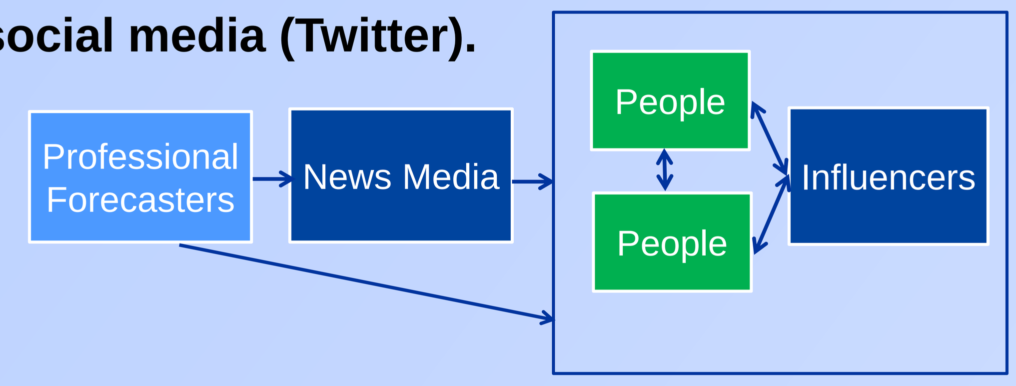
Modeling Social Learning as Epidemics Using Twitter Data

Shivi Kalra, PhD Candidate in Economics, SUNY Binghamton
Wei Xiao, Professor, SUNY Binghamton

Introduction

Empirical research that questions the micro-foundations of sticky information approach. Social learning is made more realistic by assuming that the agents in the economy receive their forecasts from professional forecasters and news media who are econometricians and also learn from their neighbors which is modeled as learning from social media (Twitter).



Question and Contribution

Research Question:

- **How quickly or effectively news spreads? Fast transmission of news implies that people adjust their expectations quickly vs. slow transmission that implies sticky information.**
- **Whether the spread of news can be represented by epidemic models using Twitter data?**
- **What kind of epidemics represent spread of economic news on social media?**
- **How does an economic news about a specific event spread via Twitter? Do different news spread differently? Example – non economic news like hurricanes, earthquakes etc. vs. an economic news like the Fed interest rate decision.**

Contribution to Literature:

- **Support the micro-foundations of sticky information approach through empirical analysis**
- **Extend the common source model by Carroll(QJE,2003) and use a more complex epidemiological model to capture learning on social media.**
- **Empirically derive estimates on the speed of transmission and whether a certain information like diseases can be classified as epidemics**
- **Gather Twitter data and use corresponding definitions to epidemiological model for estimation.**

Model and Data Estimation

Model: Susceptible-Infected-Recovered

$$S \xrightarrow{\beta} I \xrightarrow{\Upsilon} R$$

$$S(t) = -\beta S\ I/N \qquad (1)$$
$$I(t) = +\beta S\ I/N - \Upsilon I \qquad (2)$$
$$R(t) = \Upsilon I \qquad (3)$$

S(t): Set of susceptible individuals at time t. Set of users who have received tweets or retweets from infectious individuals at time t.
I(t): Set of infectious individuals at time t. Set of individuals who tweeted or retweeted about that topic at time t.
R(t): Set of individuals who have recovered at time t. Set of infectious individuals who have been inactive for a predefined period of time by not tweeting about that topic.
β: Force of infection. Speed of Transmission.
γ: Recovery rate.
Data : Keyword Search on Twitter API

- Initial values of susceptibles, infected and recovered are taken from the Twitter data
- Runge Kutta method to solve ode's.
- Least squares used to estimate unknown parameters.

Results

The potential of infection in a population depends on the basic reproduction number R0 that is defined as the average number of persons directly infected by an infectious individual during his entire infectious period when he enters a totally susceptible population.
E.g. Zika Virus in Columbia: R0 ~ 3.9

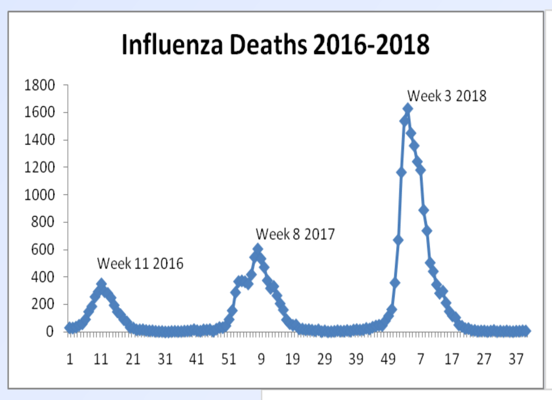
Economics News	Beta	R0
Fed Rate Hike	6.9	1.24
Is Bitcoin a Bubble?	18.04	1.74
President Lopez and Mexico's Economy	11.05	3
Bond Market and the Inverted Yield Curve	40.30	7.31

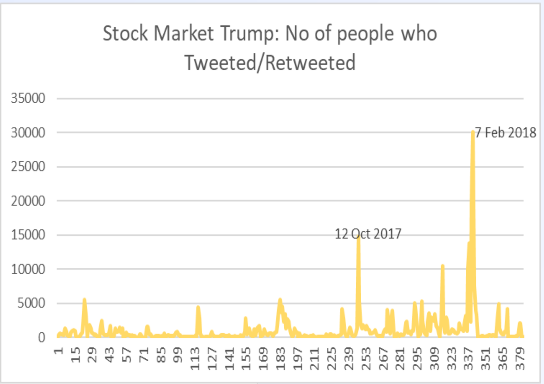
Non Economics News	Beta	R0
#Hurricane Florence	45.90	4.5
Florida Mass Shooting	16.87	5.48
#midtermelections	40.76	10.14
Trump Impeachment	124.97	51.62

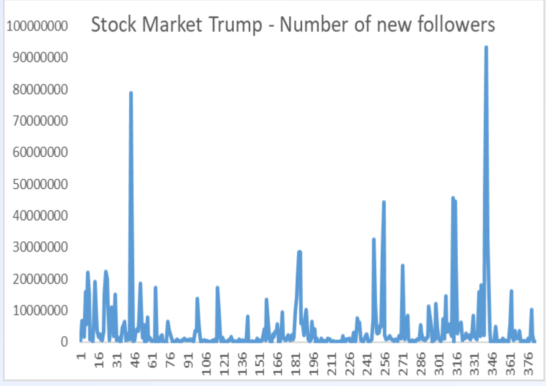
News on Twitter spreads in a similar manner to mixed epidemics that spreads from both common source like news media and person to person i.e. ordinary users to each other.

Results (Continued)

- **Influencers on Twitter play a huge role in spreading the information at a faster rate.**
- **The estimates obtained for the speed of transmission (beta) of economic news like Bitcoin being a bubble are quite low as compared to that of non-economic news like mass shootings.**
- **News spread on Twitter represents epidemics like influenza that changes its course over the year as environmental conditions become favorable or more crowding takes places. It follows a wave like pattern.**







Data on Influenza obtained from Centers of Disease and Controls

Conclusion

- **This empirical analysis supports Carroll's (2003) finding where he attributed the constant term obtained by estimating regression equations on aggregate inflation and unemployment expectations, to person to person learning.**
- **Each person has a different probability of spreading information as well as to update their expectations. Introducing heterogeneity about expectations in macro models is a step ahead to derive more accurate forecasts.**
- **The theory of rational inattention still holds as news update by people is limited by a circle of friends and followers on social media. Not everyone pays attention to every news item and hence update their expectations.**
- **Limitation of Data – Data not segregated for spread of news from news media vs. ordinary users communicating with each other. It would give us important insights for calibrating and formulating assumptions in standard macroeconomic models. This is left for future work.**

References

Christopher D. Carroll. 2003. The Epidemiology of Macroeconomic Expectations. Quarterly Journal of Economics, Volume 118, Number 1, February 2003