Journal of Economic Behavior and Organization 149 (2018) 46–57

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

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

Finite languages, persuasion bias, and opinion fluctuations

Manuel Foerster

Department of Economics, University of Hamburg, Von-Melle-Park 5, Hamburg, 20146 Germany

ARTICLE INFO

Article history:
Received 20 January 2017
Revised 1 February 2018
Accepted 1 March 2018

JEL classification:
D83
D85
Z13

Keywords:
Bounded rationality
Finite languages
Opinion fluctuations
Persistent disagreement
Persuasion bias
Social networks

ABSTRACT

We propose a boundedly rational model of opinion formation in which agents are subject to persuasion bias and communicate via finite languages. Agents are organized in a social network and repeatedly update their beliefs based on coarse messages about their neighbors' beliefs. We show that agents do not reach a consensus; instead, their beliefs keep fluctuating forever if different languages are present in their neighborhoods. In particular, we recover the classical result that under persuasion bias agents typically reach a consensus if there is a unique language in society, while small perturbations lead to fluctuations. Our approach provides and formalizes a possible mechanism to account for theories according to which storytelling may generate excessive confidence swings.

1. Introduction

Individuals form their beliefs and opinions on various economic, political and social issues based on information they receive from their social environment. The way we use language to transmit this information differs between individuals, e.g., some people may express themselves more clearly than others, or attach different meanings to words. Our aim is to analyze how such differences in language use affect long-run beliefs when agents are boundedly rational.

We develop a model of opinion formation in which agents are subject to *persuasion bias* or *correlation neglect*, i.e., fail to adjust for possible repetitions of information they receive. We can see persuasion bias "as a simple, boundedly rational heuristic for dealing with a very complicated inference problem" (DeMarzo et al., 2003). Correctly adjusting for repetitions would be extremely complicated when individuals interact repeatedly on complex social networks, see Battiston and Stanca (2015), Brandts et al. (2015) and Corazzini et al. (2012) for experimental evidence.[1] The novelty of our model is that we drop the typical assumption that agents can precisely communicate their beliefs; instead, they use *finite languages* that may *vary* within society. Hence, meanings of *messages* are necessarily *coarse* and can differ both with respect to the beliefs that induce them and with respect to their interpretations. This may be due to conflicts between individuals similar to the cheap talk framework of Crawford and Sobel (1982).[2] But it may as well be simply because people only understand finitely many words and use them differently depending on their cultural background, vocabulary knowledge and personal taste,

E-mail address: manuel.foerster@wiso.uni-hamburg.de

[1] We refer to DeMarzo et al. (2003) for a detailed discussion of this assumption, including references on psychological evidence.

[2] For instance, a sender who would like to bias the receiver upwards can use messages indicating a high belief for a larger range of beliefs than messages indicating a low belief; the receiver could choose any elements from the respective ranges as interpretations. The resulting finite language is qualitatively

https://doi.org/10.1016/j.jebo.2018.03.001
0167-2681/

while the information to communicate is a much more complex object, see Blume and Board (2013), Cremer et al. (2007), and Jäger et al. (2011). A more concrete example are financial markets where agents often report stock ratings (evaluations of a stock's expected performance and risk level) according to a finite rating scale and may use the scale differently. Additionally, given the complexity of financial markets, it is likely that agents suffer from persuasion bias.[3]

We consider a society where each agent starts with an *initial belief (or opinion)* about some unknown state. Agents are sequentially selected (in a random order) to "meet" their *neighbors* in the *social network*. When an agent is selected, she (the *receiver*) receives information from one of her neighbors (the *sender*) according to a stochastic process that forms the social network. We assume that each pair of agents uses a *finite language* to communicate, which is exogenously given and may vary between pairs. The sender sends one of finite *messages* that provides coarse information about her belief to the receiver. In turn, the receiver interprets the message. Finally, the receiver updates her belief to the average of the interpretation and her pre-meeting belief, reflecting that she is subject to persuasion bias.

We show that the opinion dynamics fails to converge to a steady state if agents communicate via different languages. Moreover, each agent's belief eventually fluctuates on a subset of the belief space if different languages are present in her neighborhood. By contrast, if there is a unique language in society, we recover the classical result that under persuasion bias a connected social network (and typically some weak regularity condition) makes society reach a consensus, see, e.g., DeGroot (1974), DeMarzo et al. (2003), and Golub and Jackson (2010). Thus, it follows that convergence to a consensus is not stable with respect to perturbations in the languages. Finally, we show that beliefs fluctuate in an ergodic way, that is, their empirical time averages converge to their long-run expectations. In particular, the fluctuations are shaped by the languages and the social network.

Disagreement among individuals and constantly changing opinions are the norm on many central issues in our societies, and agreement is the rare exception even though these issues have been debated for decades, see, e.g., Kramer (1971) and DiMaggio et al. (1996). Akerlof and Shiller (2010) argue that storytelling may generate excessive confidence swings that could drive aggregate distortions.[4] Recent experimental evidence by Enke and Zimmermann (2017) lends support to this theory. We demonstrate that persuasion bias in conjunction with finite languages constitutes a possible mechanism to explain how storytelling may generate confidence swings.

Our results are surprising in view of the literature on opinion formation under persuasion bias, where agents typically reach a consensus in the long-run.[5] To this respect, Acemoglu et al. (2013) are the closest to our work. These authors introduce stubborn agents that never change their opinions, which also leads to fluctuations if the other agents update regularly from different stubborn agents. Furthermore, several authors have proposed models to explain non-convergence of beliefs, usually incorporating some kind of homophily that leads to opinion polarization, see, e.g., Axelrod (1997), Golub and Jackson (2012), and Hegselmann and Krause (2002). The literature on rational learning in networks mostly focuses on observational learning,[6] with two recent exceptions being Acemoglu et al. (2014) and Foerster (2016), where agents communicate their information via cheap talk.

The paper is organized as follows. In Section 2 we introduce the model and notation. Section 3 contains the analysis of the long-run opinion dynamics. In Section 4 we conclude. The proofs are presented in the Appendix.

2. Model and notation

We consider a set $N = \{1, 2, \ldots, n\}$, with $n \geq 2$, of *agents* who repeatedly communicate with their neighbors in a social network. Each agent $i \in N$ starts with an *initial belief* or *initial opinion* $x_i(0) \in [0, 1]$ about the unknown state $\theta \in [0, 1]$.[7] At time $t \geq 0$, we write agent *i*'s belief as $x_i(t)$. The *social network* is given by a stochastic matrix $P = (p_{ij})_{i,j \in N}$, that is, $p_{ij} \geq 0$ for all $i, j \in N$ and $\sum_{j \in N} p_{ij} = 1$ for all $i \in N$. For agent *i*, p_{ij} is the probability to meet agent *j*, and $N_i = \{j \in N \mid p_{ij} > 0\}$ denotes *i*'s *neighborhood*. Let $g = \{(i, j) \mid p_{ij} > 0\}$ denote the directed graph induced by meeting probabilities $p_{ij} > 0$. We make the following assumptions about the social network.

Assumption 1.

(i) (No self-communication) Agents do not communicate with themselves, i.e., $p_{ii} = 0$ for all $i \in N$.
(ii) (Connectedness) The graph g is strongly connected, i.e., for all $i, j \in N$ there exists a directed path connecting i to j with links in g.[8]

similar to the equilibria in Crawford and Sobel (1982). An earlier draft of this paper was built on this idea and has circulated under the title "Strategic communication under persuasion bias in social networks" (Foerster, 2015).

3 See Enke and Zimmermann (2017) for experimental evidence.

4 See also Farmer (2013) who proposes a rational expectations model in which beliefs about future asset prices are subject to exogenous shocks to explain financial crises.

5 This finding persists in presence of certain strategic elements, e.g., misreporting of beliefs due to preferences for conformity (Buechel et al., 2015) or manipulation of the social network (Foerster et al., 2016), and the introduction of forceful agents who (almost) do not update their beliefs (Acemoglu et al., 2010).

6 See, e.g., Acemoglu et al. (2011), Bala and Goyal (1998), Gale and Kariv (2003), and Mueller-Frank (2013).

7 Initial beliefs could be noisy signals about the state, but may in general be any distribution-free observations, see also DeMarzo et al. (2003).

8 We say that there is a *directed path* from i to j in g if there is a sequence of distinct agents $i = i_1, i_2, \ldots, i_K = j$ such that $(i_k, i_{k+1}) \in g$ for all $k = 1, 2, \ldots, K-1$.