Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 167 (2017) 229–268

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

# Bayesian persuasion by a privately informed sender

Jonas Hedlund

*Department of Economics, University of Heidelberg, Bergheimer str. 58, 69115 Heidelberg, Germany*

Received 16 June 2015; final version received 1 November 2016; accepted 11 November 2016

Available online 17 November 2016

**Abstract**

This paper introduces private sender information in a game of Bayesian persuasion with monotonic sender preferences. I derive properties of increasing differences related to the precision of signals and use these to characterize the set of equilibria selected by the D1 criterion. These equilibria are either separating (i.e., the sender's *choice* of signal reveals his private information to the receiver) or fully disclosing (i.e., the *outcome* of the sender's chosen signal fully reveals the payoff-relevant state). Which of these two cases occurs is determined by the optimality properties of fully disclosing signals. If full disclosure is optimal for all sender types, then the equilibrium is fully disclosing. Otherwise, the equilibrium is fully separating, and incentive compatibility requires the sender to use signals that are strictly more informative than the ones that would be used under symmetric information. Therefore, when full disclosure is suboptimal, the sender incurs a cost in comparison to the symmetric information case, whereas the receiver benefits from better information.



*JEL classification:* D82; D83; D86

*Keywords:* Bayesian persuasion; Signaling; Information transmission

## 1. Introduction

The recent literature on *Bayesian persuasion*, pioneered by Kamenica and Gentzkow (2011), studies the strategic design of experiments by a *sender* whose objective is to influence the beliefs of a *receiver*. In contrast to most established models on information transmission (e.g.,

*E-mail address:* jonas.hedlund@awi.uni-heidelberg.de.

http://dx.doi.org/10.1016/j.jet.2016.11.003

Crawford and Sobel, 1982; Milgrom, 1981; Spence, 1973), the sender and the receiver share the same prior information, and the sender influences the receiver by committing to an informative experiment. The experiment may consist of a public signal, a protocol of information acquisition, or independent research sponsored by the sender. Situations fitting the description include firms specifying the terms of free trials of new products or interest groups funding research for lobbying purposes. In both examples, the sender (firm or interest group) controls the type of information generated and is able to commit to reveal the generated information to the receiver. Applications of the framework have led to novel insights in, for example, collective decision making (Alonso and Câmara, 2016c) and industrial organization (Bergemann et al., 2015).

In some situations, however, it is unlikely that the sender and the receiver would share the same prior information. For example, a firm specifying the terms of free trials (e.g., their length or functionality) upon launching new software is likely better informed than a typical consumer about the user-friendliness of the software. An interest group promoting a product may have private information regarding health concerns prior to funding further independent research on the subject. How would such private sender information impact the equilibrium design of experiments? To which extent is it possible to infer the sender's private information from the type of experiment he conducts?[1] The present paper finds that, in a natural class of environments, private sender information unravels and can be inferred from the nature of the evidence the sender collects, the tests he conducts, or the trials he offers, even if the private information is unverifiable and not subject to standard unraveling arguments (as in Milgrom, 1981).

The framework consists of a simple model of Bayesian persuasion in which the sender has unverifiable and imperfect private information about a binary payoff-relevant state prior to generating further information about this state. Contingent on his private information, the sender chooses a random signal (i.e., an experiment), which can be informative about the payoff-relevant state. The receiver observes an outcome of the signal, updates her beliefs, and the sender collects a payoff that is continuous and strictly increasing in the receiver's updated belief.[2]

The first formal result reveals that the ensuing game of privately informed Bayesian persuasion is structured by properties of increasing differences, which arise endogenously in equilibrium and roughly state that a sender with more favorable private information has stronger preferences for more precise signals (Lemmata 1 and 2). As a consequence, much of the standard signaling logic applies, and by confining attention to equilibria selected by Cho and Kreps' (1987) D1 criterion (hereafter "equilibria"), it is possible to make a number of predictions about the sender's behavior. The first main result demonstrates that private information leads to a form of unraveling (Proposition 1). In particular, the sender's equilibrium strategy consists of signals that are either separating or fully reveal the payoff-relevant state. That is, either the sender's *choice* of signal reveals his private information to the receiver, or the *outcome* of the chosen signal fully reveals the payoff-relevant state. Roughly, the increasing differences in precision combined with the D1 criterion require the receiver to attribute deviations to sufficiently precise signals to sender types with favorable private information, making it possible to find profitable deviations for such types from most pooling strategies.

[1] That is, will the terms of the free trials signal the firm's private information about user-friendliness? Can the interest group's private information about health risks be inferred from the design of the research project it funds?

[2] Such monotonic preferences play an important role in the literature on transmission of verifiable information (Milgrom, 1981; Milgrom and Roberts, 1986), and in several signaling models (Cho and Sobel, 1990; Mailath, 1987; Spence, 1973).