Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 159 (2015) 531–551

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Strategic experimentation with private payoffs [☆]

Paul Heidhues [a], Sven Rady [b,*], Philipp Strack [c]

[a] *Lufthansa Chair in Competition and Regulation, ESMT European School of Management and Technology, Schlossplatz 1, D-10178 Berlin, Germany*

[b] *Department of Economics and Hausdorff Center for Mathematics, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany*

[c] *Department of Economics, University of California at Berkeley, 513 Evans Hall, Berkeley, CA 94720, USA*

Received 3 October 2014; final version received 21 May 2015; accepted 27 July 2015

Available online 30 July 2015

Abstract

We consider a game of strategic experimentation in which players face identical discrete-time bandit problems with a safe and a risky arm. In any period, the risky arm yields either a success or a failure, and the first success reveals the risky arm to dominate the safe one. When payoffs are public information, the ensuing free-rider problem is so severe that equilibrium experimentation ceases at the same threshold belief at which a single agent would stop, even if players can coordinate their actions through mediated communication. When payoffs are private information and the success probability on the risky arm is not too high, however, the socially optimal symmetric experimentation profile can be supported as a sequential equilibrium for sufficiently optimistic prior beliefs, even if players can only communicate via binary cheap-talk messages.

JEL classification: C73; D83

☆ Our thanks for helpful comments and suggestions are owed to Simon Board, Patrick Bolton, Susanne Goldlücke, Johannes Hörner, Navin Kartik, Nicolas Klein, Moritz Meyer-ter-Vehn, Johannes Münster, Nicolas Vieille, an associate editor, two anonymous referees, and to seminar participants at UCLA, UC Berkeley, Yale, Frankfurt, the 2011 North American Winter Meeting of the Econometric Society, the 14th SFB/TR 15 Conference (Mannheim, April 2012), and the Workshop on Advances in Experimentation (Paris, December 2013). We thank the Economics Department at UC Berkeley and the Cowles Foundation for Research in Economics at Yale University for their hospitality. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

* Corresponding author.
E-mail address: rady@hcm.uni-bonn.de (S. Rady).

http://dx.doi.org/10.1016/j.jet.2015.07.017

Keywords: Strategic experimentation; Two-armed bandit; Bayesian learning; Information externality; Mediated communication; Cheap talk

1. Introduction

In many real-life situations, economic agents face a trade-off between exploring new options and exploiting their knowledge about which option is likely to be best. A stylized model allowing one to analyze such experimentation problems is the two-armed bandit setting in which a gambler repeatedly decides which of two different slot machines to play with the ultimate goal of maximizing his monetary reward. Consequently, starting with Rothschild (1974), variants of this bandit problem have been studied in a wide variety of economic set-ups; see Bergemann and Välimäki (2008) for an overview, and the references in Section 2 for specific applications.

In this paper, we analyze strategic experimentation problems in which players can learn not only from their own past experiences but also from those of others. In contrast to the previous literature on multi-agent bandit problems, we allow players to communicate with each other. We consider both the case in which they observe each other's actions and payoffs and the case in which they only observe each other's actions. To fix ideas, think of experimental consumption. A consumer can learn by trying different products herself, of course. But she can also learn from observing others' consumption choices and from communicating with others about their experiences. Typically, however, it will be impossible for her to directly observe other agents' payoffs. In contrast, farmers experimenting with various crops may be able to observe not only the crop planted by their neighbors but also whether the crop thrives or not.

Section 3 introduces our strategic experimentation environment with N identical bandit machines, each of them consisting of a safe and a risky arm. The payoff distribution of the risky arm is the same for all machines, and can be either "good" or "bad". In each period, a good arm produces either a "success" or a "failure", while a bad arm always produces failures. Studying the scenario where all these machines are operated by one and the same agent yields the efficient benchmark; the special case where $N = 1$ leads to the autarky solution.

Section 4 turns to the case in which each machine is controlled by a different player, and players' actions *and* payoffs are publicly observable, so that all players always share a common belief about the state of the world. Focusing on continuous-time set-ups, Bolton and Harris (1999, 2000), Keller et al. (2005) and Keller and Rady (2010) show that if the players condition their actions on their common belief only, i.e. if they use Markov strategies, it is impossible to achieve the social optimum. Furthermore, if a single success on the risky arm fully reveals the good state of the world (and players are not allowed to switch actions infinitely often in finite time), then in any Markov perfect equilibrium players stop experimenting once the common belief reaches the single-agent cut-off (Keller et al., 2005).

Maintaining the assumption of fully revealing successes, we work in a discrete-time set-up and introduce a mediator in the spirit of Forges (1986) and Myerson (1986)—thereby allowing for general forms of communication between players. This enables us to considerably strengthen the existing results by showing that in *any* communication equilibrium, players stop experimenting once the common belief falls beneath the single-agent cut-off. It is worth emphasizing that this impossibility result holds even though our equilibrium concept (Nash equilibrium in the game induced by the mediator's strategy) is permissive in two ways: it allows the players to coordinate through mediation, and it abstains from imposing sequential rationality.