Behavioral Implementation*

Geoffroy de Clippel†

This Version: June 2012

Abstract

Implementation theory assumes that participants' choices are rational, in the sense of being derived from the maximization of a context-independent preference. The paper investigates implementation under complete information when the mechanism designer is aware that individuals suffer from cognitive biases that lead to violations of IIA, or cannot exclude the possibility of such "irrational" behavior.

1 Introduction

Implementation under complete information is a classic problem in mechanism design. The designer would like to implement a rule that selects acceptable outcomes as a function of a problem's characteristics. Unfortunately, while commonly known among participants, these characteristics are unknown to him. He must thus rely on their reports to tailor his selection of outcomes. Taking into account the participants' incentives to misrepresent their information, what are the rules that the designer can effectively implement?

Characteristics encode participants' preferences in standard implementation models. However, there is ample evidence in marketing, psychology and behavioral economics that people's choices need not be consistent with the maximization of a preference relation. Classic examples, which have played

*The paper benefited from insightful conversations with Kfir Eliaz, Eric Maskin, Kareen Rozen, and Roberto Serrano. I also wish to thank Chris Chambers, George Mailath, Rene Saran, and Rajiv Vohra for useful comments. Financial support from the Deutsche Bank through the Institute for Advanced Study is gratefully acknowledged.

†Brown University, Department of Economics, Providence, Rhode Island - declippel@brown.edu.