Temporal stability and psychological foundations of cooperation preferences

Stefan Volk [a],[*], Christian Thöni [b], Winfried Ruigrok [c]

[a] University of Tübingen, Faculty of Business Administration and Economics, Melanchthonstrasse 30, 72074 Tübingen, Germany
[b] University of St. Gallen, Department of Economics, Varnbüelstrasse 14, 9000 St. Gallen, Switzerland
[c] University of St. Gallen, Research Institute for International Management, Dufourstrasse 40a, 9000 St. Gallen, Switzerland

ARTICLE INFO

ABSTRACT

A core element of economic theory is the assumption of stable preferences. We test this assumption in public goods games by repeatedly eliciting cooperation preferences in a fixed subject pool over a period of five months. We find that cooperation preferences are very stable at the aggregate level, and, to a smaller degree, at the individual level, allowing us to predict future behavior fairly accurately. Furthermore, our results provide evidence on the psychological foundations of cooperation preferences. The personality dimension 'Agreeableness' is closely related to both the type and the stability of cooperation preferences.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

In his influential essay on "The Methodology of Positive Economics" (1953, p. 4), Friedman explains that the task of economic theory "is to provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances".

In brief, an important aspect of economic theory is to make predictions, which are enormously facilitated if peoples' tastes are stable across time and circumstances. The assumption of stable preferences belongs accordingly to the core elements of economic theory, or, as Becker (1976, p. 5) puts it: "The combined assumptions of maximizing behavior, market equilibrium, and stable preferences, used relentlessly and unflinchingly, form the heart of the economic approach as I see it".

On the same page Becker also argues that "preferences are assumed not to change substantially over time". He explains that the assumption of preference stability both across different circumstances and over time "provides a stable foundation for generating predictions about responses to various changes, and prevents the analyst from succumbing to the temptation of simply postulating the required shift in preferences to "explain" all apparent contradictions to his predictions". Whether

* Corresponding author. Tel.: +49 7071 29 78179; fax: +49 7071 29 5534.
E-mail addresses: stefan.volk@uni-tuebingen.de (S. Volk), christian.thoeni@unisg.ch (C. Thöni), winfried.ruigrok@unisg.ch (W. Ruigrok).

the assumption of preference stability is reasonable for real world preferences is ultimately an empirical question. However, empirical evidence on preference stability is surprisingly scarce.

In this paper we present experimental evidence on the stability of social preferences, more precisely on the stability of preferences for cooperation in a social dilemma. We measure cooperation preferences by Fischbacher et al.'s (2001, henceforth FGF) variant of the four player one-shot public goods game and repeat the measurement with the identical subject pool after two and a half and five months. The FGF method has the advantage of measuring cooperativeness to a large degree independent of subjects' beliefs, which are presumably more volatile than preferences.[1] Furthermore, we also present evidence on the psychological foundations of cooperation preferences, more specifically on the effects of individual differences in personality on these preferences. Our experimental setup enables us to investigate (i) aggregate-level and individual-level stability of cooperation preferences over time and (ii) the relationship between stable personality traits and cooperation preferences.

Our paper contributes to a nascent body of literature on the stability of social preferences. However, the existing literature has largely focused on the distribution of heterogeneous preference types across different experiments at the aggregate level.[2] The few studies that have examined individual-level stability can be divided into two categories.

The first category includes studies investigating the consistency and stability of social preferences across different games or variants of the same game in a given experimental session. Andreoni and Miller (2002) and Fisman et al. (2007) apply the axioms of revealed preferences to decisions observed in modified dictator games. They find that subjects' choices can be rationalized by a well-behaved utility function. Fischbacher and Gächter (2010) elicit individual cooperation preferences in a strategy method experiment and observe contributions in ten consecutive one-shot games in the direct response mode. They find that data on individual cooperation preferences allow to predict the development of contributions over time very accurately.[3] Blanco et al. (2011) investigate the stability of preferences across different games by observing the same sample of subjects in a number of different one-shot games. They find that about one third of their subjects exhibit stable preferences in the sense that they consistently follow a plausible behavioral norm.[4]

The second category of research includes a very small number of studies examining longitudinal stability of social preferences over time. Muller et al. (2008) elicit subjects' cooperation preferences repeatedly in a series of five consecutively played two-stage public goods experiments within one session. They find that 37 percent of their 60 subjects remain in the same preference category for all five measurements. Studies allowing for more time between the measurements are particularly scarce.[5] Brosig et al. (2007) implement sequential prisoner's dilemma games and repeat the initial experiment with the same subject pool two times with one month in between each repetition. They find that 43 percent of the subjects choose the same response in all three waves, and all of these subjects act consistently selfish. On average they observe a decay of cooperative behavior over time.

In contrast to Brosig et al. (2007), we observe highly stable aggregate results. The distribution of cooperation preferences is basically unchanged across time. At the individual level we observe a more diverse picture. If we classify our subjects into three categories (Conditional Cooperators, Free Riders and Others) we find that about two thirds of the subjects remain in the same category for two consecutive measurements. Half of the subjects remain in the same category for all three measurements.

While the temporal stability of cooperation preferences is the main focus of the present paper, we are also exploring the psychological foundations of preference heterogeneity and stability. The idea to include psychological variables into the analysis of cooperation preference stability came from a related paper (Volk et al., 2011), which is based on the same data as this paper, and which served to some extent as a preparatory analysis for the present paper. In Volk et al. (2011) the main focus is on how personal values and personality traits are related to one another and how they jointly impact cooperation preferences. This analysis is important because one of the problems inherent in the study of psychological correlates of economic behavior is the large number of conceivably important psychological constructs to pick from. Two particularly interesting constructs are personality traits and personal values as they represent the difference between innate characteristics representing the nature of an individual (i.e., personality traits) and socially learned characteristics resulting from the interaction of nature and nurture (i.e., personal values) (see for example Olver and Mooradian, 2003; Parks and

---

[1] There is of course no way to measure preferences directly. In a strict sense it is thus impossible to test preference stability. We also cannot identify the proximate mechanism which causes subjects to contribute, be it interdependent preferences or reciprocity (see Cox, 2004; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006; Fehr and Schmidt, 1999). The FGF design offers a way to measure contribution strategies in public goods games. A large variety of social preferences give rise to identical contribution strategies. Thus, in a strict sense, all we can show is the stability of contribution strategies over time. For many applications (e.g. predicting future behavior in similar situations) it is sufficient to know that contribution strategies are stable.

[2] For experiments using the FGF design see Kocher et al. (2008), Herrmann and Thöni (2009), and Thöni et al. (2009). Homo/heterogeneity between subject pools in repeated public goods games are studied in Gächter et al. (2010).

[3] Burlando and Guala (2005) and Gächter and Thöni (2005) investigate preference stability in a setting where they regroup subjects according to their cooperation preferences. They use initial public goods games to identify different types and form groups of alike subjects and observe contributions in repeated public goods games.

[4] See also De Oliveira et al. (2009). A related strand of literature investigates whether different elicitation methods (strategy method vs. direct response method) lead to identical inference about subjects' preferences, see Brandts and Charness (2011) for an overview or Fischbacher and Gächter (2009) for the FGF design.

[5] Horowitz (1992), Andersen et al. (2008) and Zeisberger et al. (in press) investigate the stability of risk preferences over time. Meier and Sprenger (2010) present data on the stability of time preferences.