



UvA-DARE (Digital Academic Repository)

Heterogeneity in response to incentives: Evidence from field data

Czibor, E.

[Link to publication](#)

Citation for published version (APA):

Czibor, E. (2015). *Heterogeneity in response to incentives: Evidence from field data.*

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 5

Summary

This dissertation consists of a series of studies that explore individual differences in response to incentives. Chapter 2 and 3 focus on gender differences in reaction to tournaments, while Chapter 4 concentrates on other dimensions of heterogeneity such as education and occupational categories in relation to sorting into teams.

The aim of Chapter 2 is to study in a natural yet controlled environment the impact of absolute vs. relative grading on student effort and performance, with a focus on potential gender differences in reaction to the grading schemes. To this end, a large-scale framed field experiment is conducted among students following a Bachelor course at a Dutch university. Students are randomly assigned to either absolute or relative grading at the midterm exam (grading schemes are reversed for the final exam). The random assignment of students to grading schemes allows causal inferences about the effect of grading schemes on student outcomes, and a clean estimation of the potential gender differences in response to competitive grade incentives. The data collected in the experiment include exam scores, different proxies for preparation effort as well as a rich set of individual characteristics such as demographic information, preferences, confidence and ability measures.

The results show no clear difference either in effort provision or exam performance under the two grading schemes. Even though the exams represent high stakes, and comprehension of the treatment conditions is adequate, students in this sample do not react differently to grading

on the curve than to absolute grading. There is no evidence for a heterogeneity in response by gender, ability or competitive preferences, either. These findings are likely explained by an overall lack of ambition and a general disinterest in obtaining high grades, a manifestation of the often criticized “just-pass” attitude of Dutch students. This explanation is supported by the analysis of those students who are conjectured to care most about the way their exam scores are mapped into grades, i.e. students close to the pass-fail threshold. While this subgroup is relatively small, so inferences should be made with caution, there is a strong indication for a gender gap in response to relative grading among such “marginal” students. Overall, the results of this chapter suggest that competitive grade incentives can not solve the issue of insufficient student motivation: a tournament is ineffective when the prizes (i.e. high grades) are not considered valuable by the majority of the contestants.

Chapter 3 exploits a large set of naturally occurring data from an online card-game community to study gender differences in risk taking and competitive choices. The data contain information from over 4 million games, generated by more than fifteen thousand individual players. While the setting is simplified, it nonetheless has several appealing features, such as self-selection into an uncertain and competitive environment as well as real, strategic and repeated interaction (with feedback) between the players. The data allow us to make a distinction between what we call ‘selection’ and ‘playing’ behavior: we observe players’ choices regarding the level of risk and competition they prefer to bear in each round, and also their subsequent playing performance (i.e. their ability to win tricks) in the resulting tournaments. Players’ performance is measured exactly and depends on skill, strategic behavior and luck.

The data show that even though players sort into the community to participate in a game that effectively consists of a series of tournaments, this self-selection does not close the gender gap in subsequent choices related to risk taking and competition. Female players are still less likely than males to increase the stakes of the given round or to actively initiate games. The gap in initiating is largest for the game types involving individual (as opposed to team) competition. As a result of their selection choices, women end up more often in the difficult opponent po-

sition, and even when they initiate and win games, their earnings are lower due to the smaller stakes. Consequently, women accumulate lower scores in the game than men do. This gender gap in scores is not a reflection of differences in ability in the playing stage. Controlling for the type of the game and their role, women are as good as men in winning games. This chapter thus demonstrates the negative consequence of “shying away”: despite no gender differences in on-task performance, women end up lagging behind men because of their lower propensity to take risk and to launch tournaments. The data further suggests that female players’ differential choices are to a large extent attributable to gender differences in risk and competitive preferences.

Chapter 4 analyzes the determinants of sorting into teams based on results from a large-scale lab-in-the-field (i.e. artefactual field) experiment: an incentivized online survey among Dutch entrepreneurs, managers and employees. The sample of respondents consists of more than one thousand individuals and is diverse in terms of gender, age, experience, level of education and income. The survey features a real-effort task (i.e. solving Raven matrices (Raven et al., 2003)) and measures participants’ risk preferences, confidence and beliefs. In the key part of the survey, participants’ willingness to form a team with another respondent is elicited by means of an incentive-compatible mechanism. Respondents are randomly assigned to one of the two treatments that affect the content of the team option. In the Baseline condition, the team option only entails joint production, while in the Joint Decision treatment it includes *both* joint production and a joint investment choice. In both treatments, earnings from the real effort task under the team option are determined by a production function that recognizes complementarities between the partners, and thus allows for but does not guarantee efficiency gains.

Results presented in Chapter 4 replicate several findings from related lab studies on sorting into team incentive schemes. Absolute and relative confidence are confirmed to be important determinants of the willingness to pay for the team option also in a setting where team production is not implemented as an equal split of the pooled output but entails potential synergy

gains. It is further shown that strategic risk resulting from uncertainties about the teammate's performance has a large impact on the selection into teams even when the possibility of free riding is eliminated by design. A novel finding presented in this chapter is the importance of education for individuals' participation decision in teams. Controlling for differences in task performance, confidence and risk preferences, higher education is associated with a greater willingness to pay for the team option. There is suggestive evidence that this heterogeneity is explained by differences in evaluating the team option: while participants with higher levels of education tend to primarily consider the potential gains from team pay, lower-educated respondents focus more on the risks associated with teams. As a consequence, lower educated people miss out on the sizable efficiency gains that the team option entails. Finally, Chapter 4 of this dissertation finds that participants are heterogeneous in their response to a potential compromise in decision making. In particular, entrepreneurs are shown to be averse to joint decision making when they predict that it moves them away from their individual optimal choice, while no such effect is observed for managers or employees.

Based on the three studies discussed in this dissertation we can conclude that there are systematic differences between individuals' reactions to incentives, and such heterogeneities have economically important consequences. The first two chapters discuss gender differences in response to tournament incentives. Chapter 2 presents suggestive evidence that among "marginal" students (who are conjectured to care about grade incentives) competitive grading increases the exam scores of male students but decreases the performance of female students. Chapter 3 demonstrates persistent differences between the choices of female and male card game players in relation to risk taking and competition, and also the negative consequences of these differences for women's relative success in terms of accumulated scores. Finally, Chapter 4 suggests that lower-educated individuals tend to focus more strongly on the risks associated with team incentive schemes than those with higher education, and as a result, miss out on the sizable potential gains from cooperation. Chapter 4 also confirms that entrepreneurs are different from the rest of the population in their approach to shared decision rights. Taken together, these

findings imply that observable individual characteristics can be helpful in predicting people's responses to certain incentives, and taking them into account when designing incentive schemes could increase efficiency. However, we should caution against interpreting our results as causal: observable individual characteristics do not necessarily induce choices, rather they have the potential to serve as proxies for the unobservable traits, attitudes and beliefs that are the real drivers of behavior.

It is interesting to note that some important results of the dissertation would not have been possible to obtain in a lab setting. In Chapter 3 the large dataset, both in terms of the number of individual players and the number of observations per player, allows us to detect small effect sizes and heterogeneous effects. It also enables us to evaluate the long-term consequences of seemingly small and innocent differences each round. This kind of precise measurement would be difficult to acquire in a typical laboratory experiment with a limited number of participants. Additionally, the findings of Chapter 4 that selection into teams is related to participants' level of education, and that entrepreneurs respond differently to joint decision making than managers and employees would not have been possible to observe with a sample of college students typically participating in laboratory experiments. This dissertation thus highlights the importance of complementing laboratory experiments with scientific projects based on field data.