

Measuring Strategic Uncertainty in Coordination Games

FRANK HEINEMANN

Technische Universität Berlin

ROSEMARIE NAGEL

Instituto Catalana de Recerca i Estudis Avançats and Universitat Pompeu Fabra

and

PETER OCKENFELS

Goethe-Universität Frankfurt am Main

First version received December 2004; final version accepted May 2008 (Eds.)

This paper proposes a method to measure strategic uncertainty by eliciting certainty equivalents analogous to measuring risk attitudes in lotteries. We apply this method by conducting experiments on a class of one-shot coordination games with strategic complementarities and choices between simple lotteries and sure payoff alternatives, both framed in a similar way. Despite the multiplicity of equilibria in the coordination games, aggregate behaviour is fairly predictable. The pure or mixed Nash equilibria cannot describe subjects' behaviour. We present two global games with private information about monetary pay-offs and about risk aversion. While previous literature treats the parameters of a global game as given, we estimate them and show that both models describe observed behaviour well. The global-game selection for vanishing noise of private signals offers a good recommendation for actual players, given the observed distribution of actions. We also deduce subjective beliefs and compare them with objective probabilities.

1. INTRODUCTION

Consider a coordination game in which an investment pays off if and only if a sufficient number of firms invest in the same industry. If nobody else invests, then the investment of a single firm fails and all others receive their outside option payoff. “No one invests” and “everyone invests” are Nash equilibria, with the latter being efficient.

Examples for this game structure are (i) networks that are sustainable only if the number of participants who invest in the network reaches a certain threshold, (ii) standards that are established only if they are adopted by a sufficiently high number of agents, (iii) devaluations that occur only if a sufficient amount of domestic currency is sold to attack a currency peg, (iv) liquidity crises and inefficient liquidations that can be avoided only if a sufficiently high number of creditors agree to extend lines of credit, and (v) asset price bubbles that collapse if a sufficiently high number of traders sell their assets.^{1,2}

In these games, the outcome depends on players' actions, which are affected by their expectations concerning other players' actions. Deductive equilibrium analysis based on common knowledge of rationality fails to determine a unique strategy profile. Nor can players predict the

1. For more detailed examples, see Morris and Shin (2003) and Cooper (1999).

2. Public good games with step-level requirements have a similar structure. In those games, however, there is usually an interior equilibrium in which the step level needs to be met; multiplicity arises from the problem *who* should contribute, but there is also a no-contribute equilibrium (Goeree and Holt, 2005).

behaviour of other players even if they assume that others are rational and that rationality is common knowledge. Following Brandenburger (1996), we define strategic uncertainty as uncertainty concerning the purposeful behaviour of players in an interactive decision situation.³

In this paper, we propose to measure strategic uncertainty by eliciting certainty equivalents for a situation in which the payoff depends on behaviour of other subjects, analogous to measuring risk attitudes in lotteries. We apply this method by conducting experiments on a class of coordination games with strategic complementarities and on lottery choices, both framed in a similar way. More specifically, each subject makes 40 choices: 10 lottery choices and 30 choices in binary coordination games. In each coordination game, the subject can choose between a safe amount $X \leq \text{€}15$ and an option, in which he or she earns $\text{€}15$ if at least a fraction $k \in (0, 1]$ of the other players makes the same choice, but zero otherwise. The safe amount X is varied; most subjects choose the safe option when X is large and the option with uncertain payoff when X is small. A subject's switching point can be interpreted as his or her certainty equivalent for strategic uncertainty in the coordination game. Similarly, choices between a lottery and varying safe payoffs give us certainty equivalents for risk that we use for measuring subjects' risk aversion.

We conduct the experiment in different locations and vary both the number of players in a group and the number of players needed for coordination. In two control sessions, we also ask for subjective beliefs regarding the behaviour of others. The experiment ends with an extended questionnaire containing Zuckerman's Sensation Seeking Scale-V (SSS-V) that psychologists use to characterize personalities.⁴

We find that aggregate behaviour is fairly stable and predictable. Naturally, the lower the safe payoff X or the higher the fraction of players k needed for coordination, the more likely it is that players choose the safe option. These aggregated results can be described neither by pure or mixed Nash equilibria of the game with risk neutrality nor by the equilibria of a Bayesian game in which subjects' types are defined by their risk attitudes. Morris and Shin (2002) propose to measure strategic uncertainty using the posterior distribution of actions in the equilibrium of a global game. Following this suggestion, we estimate two global games in which players are assumed to have private information about the payoff function. Here, we compare the original version of global games, introduced by Carlsson and van Damme (1993), in which players have asymmetric information about monetary payoffs and perfect information about a common utility function with a version, suggested by Hellwig (2002), in which players know monetary payoffs but have asymmetric information concerning a common degree of risk aversion and make small errors. Both versions deliver a good fit to aggregate data. While previous theoretical and experimental literature treat parameters of a global game as given exogenously, we are the first to estimate a global game.

Besides testing descriptive theories, we try to identify a simple strategy as a recommendation for an individual player in coordination games. Using lottery choices as measures of individual risk aversion, we estimate utility functions and compare expected utilities of various refinement strategies, given the observed distribution of choices. For most subjects, the global-game selection (GGS), that is the equilibrium of a global game for vanishing noise of private information, calculated under risk neutrality leads to a higher expected utility than any other considered strategy and thus can be recommended for play in a one-shot coordination game.

The literature distinguishes two kinds of uncertainty (Knight, 1921). Exogenous uncertainty (or risk) is characterized by the existence of *a priori* given and known probabilities for all possible states of the world. A lottery is the prototype of a risky situation. More generally, risky situations are games against nature. Endogenous uncertainty is characterized by the absence of

3. To our knowledge, the term "strategic uncertainty" was introduced by Van Huyck, Battalio and Beil (1990). They do not give a proper definition, but it seems clear that they mean the uncertainty arising from multiple equilibria.

4. For details, see Zuckerman (1994).

exogenously given probabilities. It arises, for example, in situations where the outcome depends on social interaction. We find that subjects' certainty equivalents of coordination games are positively related to certainty equivalents of lotteries and to the "experience seeking" subscale of the Zuckerman test. Subjects who avoid risk or new experience also avoid strategic uncertainty. This suggests that subjects have similar perceptions of exogenous and strategic uncertainty if both situations are framed in a similar way.

There is a long, ongoing discussion in economics as to whether and how strategic decision-making can be modelled as a game against nature. Luce and Raiffa (1957, p. 306) suggest (without further specifics regarding the method of modelling) assigning subjective probabilities to an adversary's choices in a game: "The problem of individual decision making under uncertainty can be considered as a one-person game against a neutral nature. Some of these ideas can be applied indirectly to individual decision making (...) where the adversary is not neutral but a true adversary". Manski (2004) gives an overview of methods for eliciting subjective beliefs and analyses whether they achieve the objective to "improve our ability to predict choice behavior" (p. 1365). Aumann and Dreze (2004) emphasize that measuring subjective probabilities requires a prize to be staked on the outcome. In strategic games, staking a prize changes the payoff function and thereby the "rules of the game". To solve this problem, they posit a preference order over lotteries whose prizes may be either a pure strategy in a strategic game or a certain outcome. If the preference order satisfies the von Neumann–Morgenstern axioms, it implies the existence of two functions that may be interpreted as a utility function on outcomes and a subjective probability distribution on the adversaries' strategies.

Our within-subject design of measuring attitudes towards risk and uncertainty allows this decomposition. We deduce agents' subjective probabilities for successful coordination at their switching points. We show that in games that require a low (high) fraction of players to get a reward, most subjects underestimate (overestimate) success probabilities. Control sessions with stated beliefs confirm this result and show that subjects often fail to best respond to stated beliefs.⁵ The distribution of stated subjective probabilities is about the same, whether we ask for the probability that another individual chooses the uncertain option or for the probability that at least a fraction k of the other players choose the uncertain option.

Currently, several experimental economists relate risk and uncertainty by comparing the behaviour of individual subjects in lottery-choice tasks and strategic games. Goree, Holt and Palfrey (2003) combine lottery choices with matching pennies games to show that the inclusion of risk aversion in a quantal response equilibrium can explain systematic deviations of subjects' behaviour from Nash equilibrium. Lange, List and Price (2004) relate lottery choices with resale auctions to estimate equilibria under risk aversion. Schechter (2007) and Bohnet and Zeckhauser (2004) combine lottery choices with trust games, finding that risk-averse subjects make lower contributions.

While theory lacks a clear prediction of behaviour in games with multiple equilibria, experiments on coordination games show clear patterns of behaviour. For repeated games, Van Huyck, Battalio and Beil (1990, 1991) show that subjects coordinate rather quickly on an equilibrium that depends on group size, the coordination requirement, and subjects' experience. A high degree of efficiency can be achieved by matching in pairs but not in situations that require the coordination of all members in a group of seven or more players. Berninghaus and Erhart (1998, 2001) show that disaggregate information and a longer time horizon enhance efficiency.

5. Nyarko and Schotter (2002) estimate and elicit subjective beliefs about opponents' behaviour in a repeated game to study social learning of beliefs. They compare the predictive power of stated beliefs, fictitious play beliefs, and Cournot best response beliefs in terms of best reply and prediction of other players' behaviour. Costa Gomez and Weizsäcker (2007) compare stated beliefs with actions in 2-person 3×3 games and observe that subjects perceive the game differently when choosing an action than when stating beliefs. On average, they fail to best respond to their stated beliefs.

Heinemann, Nagel and Ockenfels (2004) compare coordination games with public information and those with private information and find no significant difference in the predictability of aggregate behaviour, despite the fact that the latter has a unique equilibrium and the former multiple equilibria. Subjects coordinate on strategies that are fairly predictable and that vary (depending on the payoff function) from the payoff-dominant equilibria to the risk-dominant equilibria of the public information game. Public information reduces coordination failures and leads to more efficient strategies. Comparative statics with respect to the parameters of the payoff function are consistent with the risk-dominant equilibrium. Schmidt *et al.* (2003) use coordination games, in which they vary either risk dominance or the level of payoff dominance, holding the other constant. They show that changes in risk dominance affect behaviour, while changes in the level of payoff dominance do not. All these experiments use repeated games. Our experiment shows that aggregate behaviour is also highly predictable in one-shot coordination games.

Refinement concepts like risk dominance are characterized by assumptions on players' beliefs regarding other players' behaviour. While some concepts are rather *ad hoc*, risk dominance has an axiomatic justification as laid out in Harsanyi and Selten (1988). Similarly, the theory of global games, developed by Carlsson and van Damme (1993) and advanced by Morris and Shin (2003), assumes that players behave as if they have private information about payoffs. This approach predicts that agents do not coordinate their actions perfectly because they receive different private signals about the potential payoffs. Private signals about payoffs serve as a vehicle for modelling strategic uncertainty. In the limit, when the variance of private information vanishes, the limit of equilibria, called "global-game selection",⁶ predicts a pure strategy that can be used as a recommendation for players. In binary-action games with two players, the GGS coincides with the risk-dominant equilibrium (Carlsson and van Damme, 1993). In other games, both concepts give similar predictions. Risk dominance has a firm axiomatic foundation, while the GGS is easier to calculate.

In the laboratory, behaviour in repeated coordination games converges to an equilibrium with comparative statics in line with the GGS. This paper demonstrates that behaviour in a one-shot game can be described by a global game with positive variance of private information, while the best response to aggregate behaviour is close to the GGS.

In Section 2, we define coordination games of our experiment and provide a theoretical analysis. We describe Bayesian equilibria and apply the theory of global games. Section 3 describes the experimental design. Section 4 presents descriptive statistics and explores predictability of behaviour in coordination games. Section 5 estimates and compares probabilistic decision models, in particular two global games. Section 6 explores recommendations for a participant in a one-shot coordination game. In Section 7, we estimate subjective beliefs and compare them to stated beliefs and to objective probabilities. Section 8 concludes the paper.

2. COORDINATION GAME

We are interested in the following critical mass coordination game: N players simultaneously decide between two actions, A and B. Action A is associated with a fixed monetary payoff X , while action B leads to a monetary payoff $R \geq X$, if and only if at least K players choose B, and zero otherwise, where $1 < K \leq N$.

As an example, imagine that $N = 10$ players choose between A and B. For choosing A, a player receives $X = 9$. For choosing B, he or she receives $R = 15$ if at least $K = 7$ players

6. This refinement has also been called "Laplacian selection" by Morris and Shin (2003), and "global-game solution" by Heinemann, Nagel and Ockenfels (2004). In our game, it is a "noise-independent selection" as defined by Frankel, Morris and Pauzner (2003)

including himself or herself choose B and zero otherwise. A descriptive theory serves the purpose of predicting the proportion of players who choose B. A refinement attributes a pure strategy to such a game and may be used for advising players.

2.1. Complete information game

In the experiment, we always set $R = \text{€}15$. The other parameters are varied across games. The complete information game is characterized by N , K , and X and by utility functions on monetary payoffs. For non-decreasing utility functions, players' choices are strategic complements and the game has two equilibria in pure strategies: either all players choose A or all choose B.

If all players are risk-neutral, there is one mixed equilibrium, in which each player chooses B with some probability p , such that all players are indifferent. Denote the cumulative binomial distribution by "Bin". The expected payoff of an agent who chooses B is $15q$, where:

$$q = 1 - \text{Bin}(K - 2, N - 1, p),$$

is the probability that at least $K - 1$ of the other $N - 1$ agents choose B. In equilibrium, p solves:

$$15(1 - \text{Bin}(K - 2, N - 1, p)) = X.$$

Across games, the equilibrium probability p increases with increasing X and K or decreasing N . An example is shown in Figure 7 in Section 5.2.

2.2. Bayesian game with private degrees of risk aversion

In the experiment, we use lottery choices to elicit players' degrees of risk aversion. This allows us to employ the notion of a Bayesian game with a finite number of types, distinguished by their revealed risk aversion. In a Bayesian game, types are drawn randomly and each player knows his or her own type but not the types of the other players. The prior distribution of types is common knowledge.

Denote a risk type by α and assume, without loss of generality, that risk aversion increases with increasing α . Denote the probability for drawing a player of type α by $f(\alpha)$. In a Bayesian equilibrium, strategies may depend on players' types. In an interior Bayesian equilibrium, players with high risk aversion choose A, players with low risk aversion choose B, and there may be at most one risk type α^* (in each equilibrium) who is indifferent and chooses B with some probability π . Thus, a Bayesian equilibrium is characterized by a pair (α^*, π) . The prior probability that a randomly selected player chooses B is:

$$p(\alpha^*, \pi) = \sum_{\alpha < \alpha^*} f(\alpha) + \pi \cdot f(\alpha^*). \quad (1)$$

The resulting probability of success with choosing B is:

$$q = 1 - \text{Bin}(K - 2, N - 1, p),$$

and the resulting expected payoff from choosing B is $q U_\alpha(15)$, where U_α denotes the utility function of type α . A pair (α^*, π) establishes a Bayesian equilibrium if and only if:

$$\left. \begin{array}{l} 1. \quad U_\alpha(X) > q U_\alpha(15) \quad \text{for all } \alpha > \alpha^*, \\ 2. \quad U_\alpha(X) \leq q U_\alpha(15) \quad \text{for all } \alpha \leq \alpha^*, \text{ and} \\ 3. \quad U_{\alpha^*}(X) = q U_{\alpha^*}(15) \quad \text{if } \pi < 1. \end{array} \right\} \quad (2)$$

Interior Bayesian equilibria are similar to the mixed equilibrium of the complete information game described above: the equilibrium probability p increases with increasing X and K or decreasing N . In Section 5, we contrast the data from our experiment with pure and mixed

equilibria of the complete information game and of the Bayesian game with uncertainty about others' risk aversion. Since the equilibria of these games perform poorly, we suggest two alternative models.

2.3. Global game with private information about monetary payoffs

In the games mentioned above, strategies are assumed to be common knowledge among players. Subjective uncertainty arises from probabilistic strategies and from the selection of players' risk types but not from uncertainty about other players' strategies. The theory of global games, instead, provides an explicit model for strategic uncertainty. Players are assumed to behave *as if* they have private information about the (common) payoff function. Thereby, it introduces an artificial information asymmetry. Strategies depend on private signals, so that, even in equilibrium, one agent cannot perfectly predict the probability of another agent choosing B because he or she does not know the other's signal.

For a unique equilibrium, uncertainty must be modelled such that, for appropriate realizations of a random variable, either option (A or B) may be a dominant strategy. Here, we assume that payoffs from option B are $15 + y$ if at least K players decide for B and $0 + y$ otherwise, that is we add a random variable y to both potential payoffs from choosing B. The true game, with $y = 0$, is treated as being selected randomly out of a class of games distinguished by different realizations of y . Assume that y has an improper uniform distribution on the reals. Then, there are two dominance regions; if $y > X$, choosing B is a dominant strategy, and if $y < X - 15$, choosing A is a dominant strategy.

Players behave as if they receive private signals y_i that are independently and normally distributed around the true value $y = 0$ with variance σ^2 . Due to the improper prior, an agent's posterior belief about this payoff given signal y_i is normal with mean y_i and variance σ^2 . Given strategic complementarity, the two dominance regions, and private information about payoffs, the global game has a unique equilibrium.⁷

If all agents have the same utility function U , an equilibrium is characterized by a common threshold $Y^*(N, K, X)$, such that an agent with signal $y_i = Y^*$ is indifferent between both actions, while a player chooses A if $y_i < Y^*$, and B if $y_i > Y^*$.

The equilibrium condition formalizes the indifference of an agent with signal $y_i = Y^*$. Denote the probability for success with action B again by $q = 1 - \text{Bin}(K - 2, N - 1, p)$, where p is the probability that another agent receives a signal above Y^* and chooses B. The expected utility from choosing B is then:

$$\begin{aligned} \text{EU}(B | y_i) &= \int_{-\infty}^{\infty} [qU(y + 15) + (1 - q)U(y)]f(y|y_i) dy \\ &= \int_{-\infty}^{\infty} [U(y + 15) - (U(y + 15) - U(y)) \\ &\quad \times \text{Bin}(K - 2, N - 1, \text{prob}(y_j > Y^*|y))]f(y|y_i) dy \\ &= \int_{-\infty}^{\infty} \left[U(y + 15) - (U(y + 15) - U(y))\text{Bin} \left(K - 2, N - 1, 1 - \Phi \left(\frac{Y^* - y}{\sigma} \right) \right) \right] \\ &\quad \times \phi \left(\frac{y - y_i}{\sigma} \right) dy, \end{aligned}$$

7. For a general exposition of the theory of global games, see Morris and Shin (2003).

where ϕ is the standard normal distribution and Φ is the cumulative standard normal distribution. The equilibrium threshold Y^* solves:

$$EU(B|y_i = Y^*) = U(X). \quad (3)$$

A common threshold of Y^* implies that the probability with which a randomly chosen player opts for B equals the probability of receiving a signal above $Y^*(N, K, X)$. Thus, the prior probability that a randomly selected player chooses B is:

$$p(N, K, X; \sigma, U) = \text{prob}(y_i > Y^* | y = 0) = 1 - \Phi(Y^*/\sigma). \quad (4)$$

In Section 5, we use the data from our experiment to estimate the S.D. of signals σ . To our knowledge, we are the first to estimate a global game, whereas previous experimental and theoretic literature treats distribution parameters as given.

The theory of global games predicts a unique equilibrium in which behaviour is not perfectly coordinated because players receive different private signals. If uncertainty about payoffs and thus the variance of private signals vanishes, the theory predicts a Nash equilibrium with almost perfect coordination that can be used as a refinement. Let us call this solution GGS. Refinements can be used as advice to players in a game.

To derive the GGS, consider that, for a uniform prior and independent signals, it is equally likely that agent i receives the lowest, second lowest, ..., or highest of all N private signals. If $y_i = Y^*$, the expected probability of success for agent i is the probability that at least $K - 1$ signals are higher than i 's own signal. Thus, an agent's subjective probability for success is:

$$E(q|y_i = Y^*) = \int_0^1 (1 - \text{Bin}(K - 2, N - 1, p)) dp = 1 - \frac{K - 1}{N}.$$

If σ^2 converges to zero, players are almost perfectly informed about the payoffs associated with B. Combining these arguments, for $\sigma^2 \rightarrow 0$, the expected utility from choosing B can be simplified to:⁸

$$\begin{aligned} EU(B|y_i = Y^*) &= U(15)E(q, y_i = Y^*) + U(0)E(1 - q|y_i = Y^*) \\ &= U(15) \left(1 - \frac{K - 1}{N}\right) + U(0) \left(\frac{K - 1}{N}\right). \end{aligned}$$

Normalizing $U(0) = 0$, for $\sigma^2 \rightarrow 0$, an agent is indifferent if and only if:

$$U(X) = U(15) \left(1 - \frac{K - 1}{N}\right).$$

Solving this equation for X yields a threshold X^* for the safe option, such that players choose B with probability 1 in games with $X < X^*$ and with probability 0 if $X > X^*$. For $X = X^*$, both options yield the same expected payoff, and behaviour is undefined. If players are risk-neutral, $U(x)$ is linear, and the common threshold is:

$$X^* = 15 \left(1 - \frac{K - 1}{N}\right).$$

In Section 6, we analyse whether the GGS constitutes better advice for actual players than other refinements.

8. For a more general treatment of the GGS, see Frankel *et al.* (2003).

2.4. Global game with private information about risk aversion

Another way of modelling a global game has been introduced by Hellwig (2002). He also assumes that players have private information about a common payoff function. Instead of embedding *monetary payoffs* in a stochastic environment, he suggests introducing private information about a common degree of *risk aversion*. The true game is treated as being selected randomly out of a class of games distinguished by the common degrees of risk aversion.

Players are assumed to behave as if they have private information about risk aversion. As before, strategies depend on private signals, so that agents cannot perfectly predict others' behaviour. For a unique equilibrium, uncertainty must be modelled such that, for appropriate realizations of a random variable, either option (A or B) may be a dominant strategy. To achieve this, Hellwig assumes that each player makes mistakes with some probability ε . Thus, for a sufficiently high degree of risk aversion, the safe option A is a dominant strategy. Vice versa, for an agent with a signal indicating a high preference for risk, it is a dominant strategy to choose B because the probability of another agent choosing B is at least ε .

Hellwig assumes an unknown, common degree of constant ARA α that is drawn from an improper uniform distribution on the reals. Players receive private signals α_i that have a uniform distribution in $[\alpha - \delta, \alpha + \delta]$. Instead of a uniform, we use a normal distribution⁹ for private signals with mean α and variance σ^2 . Then, player i 's posterior belief about α is also normal with mean α_i and variance σ^2 . In equilibrium, for $X < 15$, there is a critical signal $\alpha^*(N, K, X, \sigma, \varepsilon)$, such that agents with lower signals prefer B and agents with higher signals prefer A. They choose their preferred option with probability $1 - \varepsilon$. For a given α and $X < 15$, the probability that a randomly selected player chooses B is:

$$p(N, K, X, \alpha, \sigma, \varepsilon) = (1 - \varepsilon)\Phi\left(\frac{\alpha^* - \alpha}{\sigma}\right) + \varepsilon\left(1 - \Phi\left(\frac{\alpha^* - \alpha}{\sigma}\right)\right). \quad (5)$$

For $X = 15$, the probability of a player choosing B is ε . The function $\alpha^*(N, K, X, \sigma, \varepsilon)$ is defined by equality of the expected utilities from both options for an agent with the marginal signal α^* , that is:

$$\begin{aligned} \text{EU}(B|\alpha_i = \alpha^*) &= \text{EU}(A|\alpha_i = \alpha^*) \\ \Leftrightarrow \int_{-\infty}^{\infty} \phi\left(\frac{a - \alpha^*}{\sigma}\right) &\left[\frac{1 - \exp(-a15)}{a} (1 - \text{Bin}(K - 2, N - 1, p(N, K, X, a, \sigma, \varepsilon))) \right. \\ &\left. - \frac{1 - \exp(-aX)}{a}\right] da = 0. \end{aligned} \quad (6)$$

In Section 5, we use the data from our experiment to estimate the parameters $(\alpha, \sigma, \varepsilon)$ of this model.

Both versions of the theory of global games are closely connected. In the limit, for vanishing uncertainty, their predictions coincide. Both require strategic complementarities and the existence of dominant strategies for appropriate realizations of random variables. There is one remarkable difference, though: the original version with monetary-payoff uncertainty predicts that each subject plays a threshold strategy, while Hellwig's model predicts that subjects deviate from a threshold strategy in each single decision with some probability ε .

9. With a normal distribution, we obtain a better fit of observations than with a uniform distribution.

3. DESIGN OF THE EXPERIMENT

Sessions were run at a computer laboratory in the Economics Department of the University of Frankfurt and in the LEEEX at University Pompeu Fabra, Barcelona, between May and July 2003. In both places, most of the participants were business and economics undergraduates. The procedure during the sessions was identical throughout all sessions at both places except for language (German and Spanish). In two additional sessions, run at the experimental laboratory at the University of Bonn in February 2006, we extended the standard treatment by explicitly asking subjects for their beliefs about others' behaviour. All sessions were computerized, using the program *z-tree* (Fischbacher, 2007). Students were seated in a random order at PCs. Instructions¹⁰ were then read aloud and questions answered in private. Throughout the sessions, students were not allowed to communicate with one another and could not see each other's screens. Each subject could only participate in one session.

Subjects were randomly assigned to groups of size N , where N was 4, 7, or 10, with the same size maintained throughout a session. There were at least two groups in each session to prevent members within a group from being able to identify each other. Before starting the experiment, subjects were required to answer a few questions to ascertain their understanding of the rules. The experiment started after all subjects had given the correct answers to these questions.

In the experiment, subjects face 40 independent decision situations, organized in 4 blocks of 10. One block contains lottery choices and three blocks contain coordination games. In each situation, each subject decides between two options, A and B. Option A gives a secure payoff that ranges from €1.50 to €15.00 in increments of €1.50 within each block. The payoff for a B-choice is either 0 or €15.00. In a lottery situation, the payoff for B is €15.00 with probability $2/3$.

Coordination blocks are framed as similarly as possible to lottery blocks. Option A gives a secure payoff that varies from €1.50 to €15.00 within each block. The payoff for option B is €15, provided that at least K out of N members of the subject's group choose B in this situation. If less than K members choose B, those who choose B receive 0. K was varied across sessions and blocks but kept constant throughout the 10 situations of a block. Figure 1 shows the screen of a coordination block, where parameters K and N are replaced by numbers.

We present the 10 situations in each block ordered by the safe payoff, as Heinemann *et al.* (2004) show, in a similar coordination game with randomly ordered situations, that most subjects play threshold strategies, which means here that they play B for low secure payoffs and A for high payoffs, switching between the two at most once. Presenting options in an ordered fashion strengthens the selection of thresholds and, thereby, increases the number of data sets that can be used for statistical analyses. This design is called a "multiple price list" and has previously been used by Holt and Laury (2002) for measuring risk aversion.

To measure subjective beliefs, one could let subjects decide directly between a lottery and an outcome subject to strategic uncertainty. However, measuring both certainty equivalents separately yields identified measures for risk aversion and strategic uncertainty that allow for the analysis of optimal behaviour and the estimation of subjective beliefs.

Eliciting beliefs requires an incentive. The game is an indirect mechanism that allows the deduction of subjective beliefs from behaviour. A direct mechanism asks for beliefs and rewards subjects according to the accuracy of their beliefs, a procedure that we did not apply in standard sessions. We had three reasons for not asking for beliefs in standard sessions: (i) rewarding

10. Instructions, programs, and data are available on the supplements page of the *Review of Economic Studies* website at <http://www.restud.org>.

Period		1 of 1		Time remaining[sec]: 93		
Decide between A or B for each of the 10 situations.						
Number of decision situation	Payment for A:	Your decision A or B:		Payment for B:		
11	1.50 Euro	A	<input type="radio"/>	<input type="radio"/>	B	0 Euro, if less than 3 members of your group choose B.
12	3.00 Euro	A	<input type="radio"/>	<input type="radio"/>	B	15 Euros, if at least 3 members of your group choose B.
13	4.50 Euro	A	<input type="radio"/>	<input type="radio"/>	B	dito.
14	6.00 Euro	A	<input type="radio"/>	<input type="radio"/>	B	dito.
15	7.50 Euro	A	<input type="radio"/>	<input type="radio"/>	B	dito.
16	9.00 Euro	A	<input type="radio"/>	<input type="radio"/>	B	dito.
17	10.50 Euro	A	<input type="radio"/>	<input type="radio"/>	B	dito.
18	12.00 Euro	A	<input type="radio"/>	<input type="radio"/>	B	dito.
19	13.50 Euro	A	<input type="radio"/>	<input type="radio"/>	B	dito.
20	15.00 Euro	A	<input type="radio"/>	<input type="radio"/>	B	dito.
<input type="button" value="OK"/>						
Help When you have decided please press the button OK.						

FIGURE 1

Sample screen of a coordination game setup with $K = 3$ and $N = 10$. “dito” indicates that in each situation, the payment for B follows the same structure. For two belief elicitation sessions in Bonn, described below, we added a fifth column in which the estimated probabilities had to be given for each situation

accuracy of beliefs in a game where appropriate actions are rewarded already leads to two pay-offs in the same game; (ii) we did not want to ask for point beliefs that give imprecise measures of strategic uncertainty; and (iii) asking for probabilities, instead, imposes beliefs to be probability distributions, while our design allows for a test whether beliefs are indeed probability distributions.

In two control sessions run in Bonn, we inserted a fifth column on the screen, asking subjects directly for their beliefs about other subjects' behaviour. In one session [treatment Individual Behaviour (IB)], we asked for beliefs about individual behaviour with instructions stating (in German) “Please estimate, for each of the 10 situations, the probability with which a randomly chosen member of your group other than yourself will decide for B”. In the second session [treatment Group Behaviour (GB)], we asked for beliefs about group behaviour or success probabilities in coordination games with the text “Please estimate, for each of the 10 situations, the probability, with which you would obtain 15 euros, if you were to select B. In other words: please estimate the probability with which at least $[K - 1]$ out of the other 9 members of your group choose B in this situation”. $[K - 1]$ was replaced by the appropriate number in the respective decision block. For each lottery situation in both treatments, we asked subjects to estimate “the probability with which a randomly chosen member of your group other than yourself chooses B”. Probabilities were given as percentages (from 0 to 100).

TABLE 1

Parameters of the games

	$k = 1/3$	$k = 2/3$	$k = 1$
$N = 4$	$K = 2$	$K = 3$	$K = 4$
$N = 7$	$K = 3$	$K = 5$	$K = 7$
$N = 10$	$K = 4$	$K = 7$	$K = 10$

Note: N is the group size and K is the coordination requirement.

To prevent learning, we gave no feedback between blocks. We are primarily interested in the initial coordination behaviour since it typically determines final outcomes in repeated (coordination) games, for example in Van Huyck *et al.* (1991) and Heinemann *et al.* (2004).

At the end of the session, 1 of the 40 situations is selected randomly to determine subjects' remuneration (plus €5 for participation). Subjects were notified about which situation was selected, the result of the die (if a lottery situation was selected) or the number of group members who chose B (if a coordination game was selected), and their own profit. We select a single situation for payments to avoid hedging. This gives us the highest possible impact of risk aversion on any decision. Two high-stake sessions in Heinemann *et al.* (2004) indicate that paying a randomly selected situation does indeed evoke more risk-averse behaviour in coordination games.

In the belief elicitation sessions, we additionally remunerated the belief of one randomly chosen situation using a quadratic scoring rule adopted from Nyarko and Schotter (2002). When the probability that a randomly selected player chooses B had been elicited, we selected another player randomly and paid the subject $3 - 3(1 - p/100)^2$ euros, if the other had actually chosen B, and $3 - 3(p/100)^2$ euros otherwise, where p is the stated probability. When aggregate behaviour had been elicited, we compensated according to these rules conditional on whether at least $K - 1$ of the other group members had chosen B or not. We explained these rules and emphasized that it is optimal to state one's true probability assignment.

In each session, we used one particular group size N , one lottery block, and three blocks with different coordination requirements K . Combinations of N and K were chosen in such a way that, within a block, a subject's success with a B-choice required that at least one-third, two-thirds, or all the *other* group members chose B. Thus, $k = (K - 1)/(N - 1)$ equals 1/3, 2/3, or 1 in the three coordination setups of each session. Table 1 shows the parameter combinations used in the experiment. As no feedback was given, the order of the four blocks should not matter much. However, to minimize systematic order effects, we changed the order between sessions with otherwise equal parameters.

At the end of a session, each player had to answer a questionnaire asking for personal data, questions concerning the experiment, questions about attitudes towards various kinds of risk, and the Zuckerman SSS-V. The duration of a session without belief elicitation was 40–60 minutes, with an average payoff of €16.68 per subject. Sessions with belief elicitation lasted somewhat longer than 60 minutes, with an average payoff of €17.20.

An additional, hand-run experiment combines our treatments 4C, 7C, or 10C (with payoffs scaled down by 0.4) with a guessing game, a trust game, an ultimatum game, and decisions testing the Allais paradox.¹¹ This experiment had no Zuckerman test. Subjects were 86 participants of a meeting for people with a high intelligence quotient¹² in Cologne. Table 2 gives an overview of all our sessions.

11. The treatments discussed in this paper were presented between the trust and ultimatum games. There was no feedback given between the different tasks.

12. Sixty-five members of MENSA in Germany and 21 interested visitors.

TABLE 2
Overview of sessions

Treatment	Group size N	Order of blocks (L = lottery, K = numbers)	Belief elicitation	Location	No. subjects
4A	4	L-4-3-2	No	Frankfurt	20
4B	4	L-2-3-4	No	Frankfurt	16
4C	4	4-3-2-L	No	Frankfurt	12
4D	4	2-3-4-L	No	Frankfurt	16
7A	7	L-7-5-3	No	Frankfurt	21
7B	7	L-3-5-7	No	Barcelona	14
7C	7	7-5-3-L	No	Frankfurt	21
7D	7	3-5-7-L	No	Barcelona	14
10C	10	10-7-4-L	No	Frankfurt	20
10D	10	4-7-10-L	No	Frankfurt	20
10C-IB	10	10-7-4-L	Beliefs about individual behaviour	Bonn	20
10C-GB	10	10-7-4-L	Beliefs about successful coordination	Bonn	20
Total number of subjects in the 12 computerized sessions:					214
4C	4	4-3-2-L	No	Cologne	28
7C	7	7-5-3-L	No	Cologne	28
10C	10	10-7-4-L	No	Cologne	30
Total number of subjects in hand-run sessions:					86

Note: L-3-5-7 with $N = 7$ means, for example that first the lottery is played, then the block $K = 3$ out of 7 required for success of B is played, then 5/7 and 7/7.

4. RESULTS: DESCRIPTIVE STATISTICS

Lottery and coordination game setups are framed as similarly as possible: within each setup, subjects choose, for 10 situations, between an option A with a safe payoff increasing over situations and an option B with a risky or uncertain payoff. In all setups, subjects typically choose B when the alternative safe payoff is low and opt for A when the safe payoff is high, with only one switch between the two. Thereby, we obtain approximate measures of certainty equivalents for the lottery up to an interval of €1.50 and comparable measures for coordination games that may be interpreted as certainty equivalents for the coordination games. Data are available on the supplements page of the *Review of Economic Studies* website at <http://www.restud.org>.

Result 1. A high majority of all subjects uses threshold strategies.

We say that a subject uses a “threshold strategy” if he or she never switches back from A to B for rising safe payoffs. In Frankfurt, 131 (90%) of 146 subjects used threshold strategies in all four setups (including those who chose the same action in all 10 situations of a setup); in Barcelona, 27 (96%) of 28 chose thresholds; in Cologne, 85 (99%) of 86; and in Bonn, 34 (85%) of 40. The widespread use of threshold strategies is in line with previous experiments that presented situations in random order.¹³

Some subjects (seven in Frankfurt, four in Barcelona, one in Cologne, and one in Bonn) chose the lottery in all 10 situations, even with a safe payoff of 15. This is inconsistent with expected utility maximization. Three subjects in Frankfurt and one in Barcelona did not complete the Zuckerman test. Two subjects in Cologne did not state their age and gender. Most statistical analyses will only consider data of subjects who used threshold strategies in all four setups, chose the safe payoff instead of the lottery when the former was equal to 15, and completed

13. For details, see Heinemann *et al.* (2004).

TABLE 3
Distribution of risk types in different subject pools

	Frankfurt (%)	Bonn (%)	Cologne (%)	Barcelona (%)	All (%)
Risk-averse	50	47	30	45	43
Risk-neutral	34	29	54	18	38
Risk-loving	16	24	16	36	19

the relevant part of the questionnaire. This yields 121 subjects in Frankfurt, 22 in Barcelona, 82 in Cologne, and 34 in Bonn. The results of these subjects are used throughout unless explicitly stated otherwise.

Result 2. There are some significant differences between subject pools.

In Frankfurt, individual thresholds in coordination games are significantly lower (at the 0.05 level) than in the other locations, using pair-wise Mann-Whitney U tests separately for each coordination requirement k . Thresholds in Barcelona are significantly higher than in Bonn and Cologne for $k = 1$ at the 0.05 level. There is no significant difference in behaviour between Barcelona, Cologne, and Bonn, for $k = 1/3$ and $k = 2/3$. The session in Cologne is not entirely comparable though, because it was performed in paper form and combined our treatments with four other games. Furthermore, there are differences in the subject pools: 76% of the participants in Cologne are members of MENSA, which requires from its members an IQ above 130. Some knew each other from previous meetings. With respect to profession and age, these subjects are more diverse than our student populations. With respect to nationality, subjects in Frankfurt are more diverse than those in Barcelona and Bonn. The latter have more experience with experiments.

For the lottery setup, sessions in Cologne show significantly higher numbers of B-choices than in Frankfurt, while the other pair-wise Mann-Whitney U tests of lottery choices show no significant differences (at $p = 5\%$). The lottery has an expected payoff of €10. Thus, a risk-neutral subject should choose the lottery six times. Subjects with certainty equivalents below €9 reveal some degree of risk aversion, those with thresholds above 10.50 can be viewed as risk-lovers. This leads to an observed distribution of risk types as stated in Table 3.

Given the low payoff, an expected utility maximizer should behave approximately risk-neutral if he or she considers the induced changes of their life-time income. But we know from other experiments that subjects exhibit a degree of risk aversion that is better explained by a utility function that separates experimental income from other income.¹⁴ Distributions in Frankfurt and Bonn are in line with these other experiments. The higher proportion of risk-neutral subjects in Cologne may be explained by the fact that payoffs were scaled down by 0.4 and that other games contributed about 60% of subjects' total payoffs. The large percentage of risk-lovers in Barcelona may be a random effect of the small sample.

Because of these differences, we present results for different subject pools separately throughout this paper and use the different groups for out-of-sample tests. Table 4 gives a summary statistic of the number of B-choices. Note that for a threshold strategy, the highest safe payment, at which B is chosen, equals the number of B-choices \times €1.50.

Table 5 presents results from linear regressions with the number of B-choices in coordination setups as the explained variable. Since a subject selects three thresholds for coordination games, a simple OLS regression overestimates significance levels. We correct this using OLS regressions with clusters defined by subjects. Besides the parameters of the game, we include individual risk

14. For an example and references, see Heinemann (2008).

TABLE 4
Average number of B-choices

Group size, location	No. subjects	Lottery	Coordination games		
			$k = 1/3$	$k = 2/3$	$k = 1$
$N = 4$, Frankfurt	56	5.04 (1.73)	6.18 (2.01)	4.41 (2.16)	3.25 (2.57)
$N = 7$, Frankfurt	34	5.12 (1.77)	5.91 (2.22)	4.24 (2.09)	3.00 (2.15)
$N = 10$, Frankfurt	31	4.74 (1.72)	5.62 (2.22)	4.06 (2.55)	2.53 (2.60)
$N = 7$, Barcelona	22	6.04 (2.13)	7.27 (2.57)	6.45 (2.91)	5.59 (3.63)
$N = 4$, Cologne	26	5.38 (1.27)	7.61 (1.68)	5.58 (2.32)	3.88 (2.85)
$N = 7$, Cologne	27	5.62 (2.27)	8.11 (1.60)	6.59 (2.62)	5.11 (3.39)
$N = 10$, Cologne	29	5.66 (1.47)	7.31 (1.97)	5.62 (2.13)	3.93 (2.46)
$N = 10$, Bonn	34	5.35 (1.86)	6.94 (2.20)	5.29 (2.62)	4.09 (3.09)
Average over all N and all locations	259	5.35 (1.79)	6.81 (2.17)	5.16 (2.51)	3.40 (2.84)

Note: S.D. are given in parentheses.

aversion calculated by the number of lottery choices (BL), and also age, gender, and the four subscales of Zuckerman's SSS-V as explanatory variables. In Cologne, we did not perform the Zuckerman test.

Results 3–7 summarize the insights from Tables 4 and 5.

Result 3. The number of B-choices in coordination games decreases with an increasing coordination requirement k . Group size N has no additional effect on choices.

The coordination requirement k has a strong negative effect on the number of B-choices, which means that most subjects lower their threshold when k increases (see also Figure 2).

Subjects do not respond to changes in N , when k is held constant. If, instead, K is held constant, N has a strong positive effect on coordination because then N determines the relative hurdle to coordination. To see this, compare the increasing average number of B-choices (in Table 4) with the same K but increasing N : for example, 4.41 for $N = 4$ and $k = 2/3$, and 5.91 for $N = 7$ and $k = 1/3$, both with $K = 2$. Thus, the larger group size eases coordination. While the group size N does not affect behaviour for constant k , it has a negative impact on the probability of success given actual behaviour, as shown in Table 6. Apparently, subjects do not seem to expect this effect.

Result 4. In coordination games, the dispersion of thresholds (S.D.) is higher than in lottery setups. It tends to increase in k .

Table 4 shows (with one exception: Cologne, $N = 7$, $k = 1/3$) that the S.D. of thresholds in coordination games is higher than in lottery setups. In lottery setups, different thresholds are an expression of different degrees of risk aversion. In coordination games, there is an additional source of diversity: subjective beliefs about successful coordination. With rising coordination requirement k , the S.D. of thresholds increases (with one exception: Frankfurt, $N = 7$). This indicates that subjective beliefs are more diverse when coordination becomes more difficult. The analysis of stated beliefs in Section 7 confirms this impression.

TABLE 5
Clustered OLS regressions on the number of B-choices in coordination setups

Explaining variables	Coefficients (<i>t</i> values) of regression with data from:				
	Frankfurt	Barcelona	Cologne	Bonn	All locations
Constant	2.79 (1.68)	1.73 (0.27)	5.38*** (4.50)	8.52 (1.98)	3.40*** (4.17)
Dummy: 1 = Barcelona					1.88*** (3.34)
Dummy: 1 = Cologne					1.03*** (3.75)
Dummy: 1 = Bonn					1.17** (2.47)
Group size <i>N</i>	-0.06 (-0.86)		-0.07 (-0.81)		-0.06 (1.20)
Coordination requirement <i>k</i>	-4.56*** (-14.04)	-2.52*** (-3.44)	-5.05*** (-12.76)	-4.28*** (6.15)	-4.51*** (-19.89)
No. B-choices in lottery (BL)	0.55*** (5.94)	1.18*** (5.03)	0.38*** (2.65)	0.01 (0.03)	0.52*** (6.53)
Age	0.04 (0.77)	0.11 (0.45)	0.07** (2.62)	-0.10 (0.66)	0.07*** (2.74)
Gender (0 = female, 1 = male)	0.57 (1.56)	0.52 (0.40)	0.37 (0.90)	0.76 (0.81)	0.34 (1.36)
Boredom	-0.09 (-1.28)	-0.70 (-1.29)		-0.06 (0.22)	
susceptibility (BS)					
Disinhibition (DIS)	-0.05 (-0.64)	0.36 (1.54)		-0.34 (1.95)	
Experience seeking (ES)	0.23*** (2.83)	-0.11 (-0.43)		0.71*** (3.12)	
Thrill and adventure seeking (TAS)	0.04 (0.59)	-0.25 (-0.98)		-0.08 (0.44)	
<i>R</i> ² (adjusted <i>R</i> ²)	0.44 (0.43)	0.50 (0.44)	0.35 (0.34)	0.33 (0.27)	0.38 (0.38)
No. subjects	121	22	82	34	263

Note: Significance levels are: *5%, **2.5%, ***1%.

Result 5. The number of B-choices in coordination games is positively related to the number of B-choices in the lottery setup.

The number of B-choices in the lottery setup is a measure of risk aversion. The higher this number, the lower is a subject's revealed risk aversion. Regression results indicate that risk-averse subjects choose B less often in coordination games. It is highly significant (*t* values above 2.6) in the samples from Frankfurt, Cologne, and Barcelona, but not in Bonn. We can only speculate as to whether the independence of coordination game and lottery choices in Bonn is a random sample effect or related to belief elicitation. Nyarko and Schotter (2002), Costa-Gomez and Weizsäcker (2007), and Grosskopf and Nagel (2008), among others, reject the hypothesis that belief elicitation affects actual behaviour, while, for example Croson (2000) states the contrary.

The generally positive relation between B-choices in coordination games and lotteries has several possible explanations. First, assume that subjective beliefs and risk aversion are distributed independently. If subjects have probabilistic beliefs about others' strategies, those with higher risk aversion should, on average, have lower certainty equivalents for coordination games, which leads to the observed correlation. Otherwise, we could reject the hypothesis that beliefs are probabilities independently distributed from risk aversion. Thus, Result 5 supports the idea that subjects behave as if they have probabilistic beliefs about the outcome of coordination games.

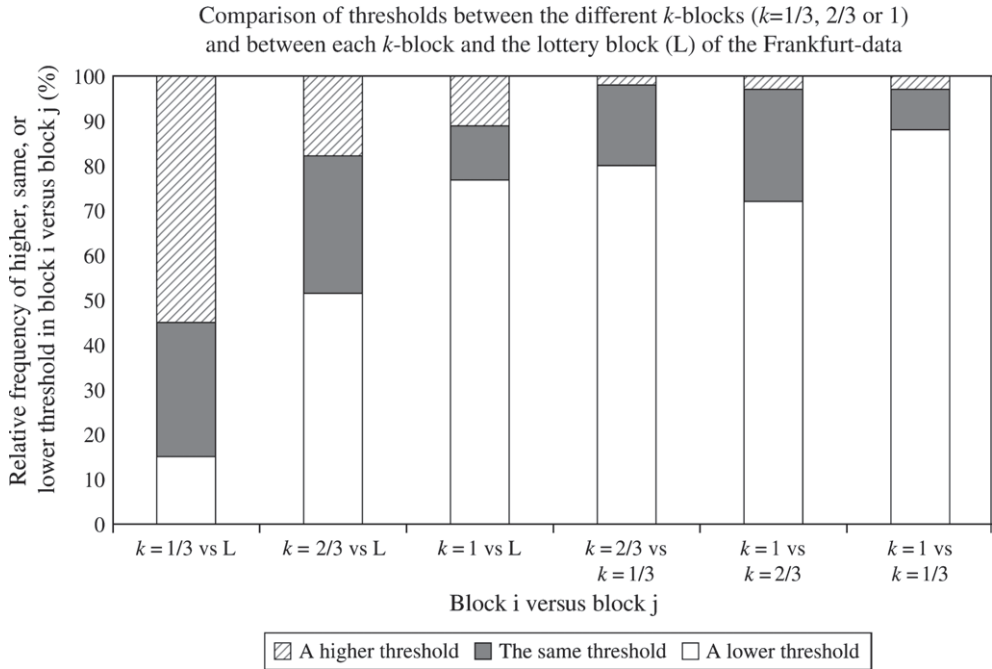


FIGURE 2

Pair-wise comparison of thresholds in different blocks. Data from Frankfurt. For example, in the first column, about 15% (white area) of all subjects have a lower threshold for $k = 1/3$ than in the lottery (L)-block, about 30% (dark area) have the same threshold in both blocks, and the remaining 55% show higher thresholds for $k = 1/3$ than in the L-block (shaded area). In the last column, 90% of players have a lower threshold for $k = 1$ than for $k = 1/3$.

This justifies the common approach to model beliefs as probability distributions, even without an exogenous random process.

Alternatively, Result 5 could be a framing effect, induced by the similarity of the design between lottery choices and coordination games. Or, additionally, if risk-averse subjects believe in a higher average risk aversion than risk-neutral or risk-loving subjects, then they may also choose a lower threshold in coordination games, provided that they believe in a correlation between risk aversion and behaviour in the coordination game.¹⁵ The results of the belief elicitation sessions in Bonn indicate that there is indeed a correlation between risk aversion and beliefs about others' risk aversion. Detailed results from these sessions are presented in Section 7 below.

Result 6. The number of B-choices in coordination games is positively related to the experience seeking subscale of the Zuckerman test and tends to be higher for older participants. Gender has no significant effect.

Combined, Results 5 and 6 tell us that subjects who avoid risk or new k experience tend to avoid strategic uncertainty. Only in Barcelona does experience seeking have no significant impact. We attribute this to the small sample. There is little age variation among subjects in Frankfurt, Barcelona, and Bonn. Participants of the MENSAs meeting in Cologne are more diverse with respect to age (16–51 years). Here, age is significant. Older subjects opt more often for

15. We are grateful to the editor, Juuso Välimäki, for pointing this out to us.

TABLE 6
Probability for successful coordination

<i>N</i>	<i>K</i>	<i>k</i>	<i>X</i> = 1-50	3-00	4-50	6-00	7-50	9-00	10-50	12-00	13-50	15-00
Frankfurt												
4	2	1/3	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>0.99</u>	<u>0.98</u>	0.86	0.59	0.29	0.16	0.06
7	3	1/3	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>0.98</u>	0.92	0.64	0.27	<i>0.10</i>	0.10
10	4	1/3	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	0.94	0.73	0.49	0.22	0.12
4	3	2/3	0.92	0.92	0.88	0.66	0.27	0.06	0.01	0.01	<u>0.00</u>	<u>0.00</u>
7	5	2/3	0.90	0.68	0.73	0.36	<i>0.13</i>	0.04	0.00	0.00	0.00	0.00
10	7	2/3	<u>0.95</u>	<u>0.95</u>	0.88	0.22	0.04	0.00	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
4	4	1	0.37	0.27	0.14	<u>0.04</u>	<u>0.01</u>	0.00	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
7	7	1	0.09	0.04	0.01	0.00	0.00	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
10	10	1	0.02	0.01	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
Barcelona												
7	3	1/3	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	1.00	0.98	0.87	<i>0.87</i>	0.32
7	5	2/3	0.99	0.99	0.97	0.89	<i>0.83</i>	0.52	0.29	0.29	0.29	<u>0.01</u>
7	7	1	0.45	0.25	0.18	0.18	0.07	<u>0.02</u>	<u>0.01</u>	<u>0.01</u>	<u>0.01</u>	<u>0.00</u>
Cologne												
4	2	1/3	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	0.98	0.97	0.83	0.45	0.01
7	3	1/3	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	1.00	1.00	0.94	<i>0.83</i>	0.07
10	4	1/3	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	1.00	0.99	0.92	0.35	0.00
4	3	2/3	0.99	0.99	0.94	0.85	0.55	0.17	0.13	0.07	<u>0.02</u>	<u>0.00</u>
7	5	2/3	1.00	1.00	0.99	0.89	<i>0.68</i>	0.52	0.23	0.13	0.09	<u>0.00</u>
10	7	2/3	<u>1.00</u>	<u>1.00</u>	1.00	<i>0.97</i>	0.65	0.30	<u>0.03</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
4	4	1	0.38	0.38	0.26	<u>0.03</u>	<u>0.01</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
7	7	1	0.45	0.34	0.13	0.03	0.00	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
10	10	1	0.24	0.16	<u>0.04</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
Bonn												
10	4	1/3	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>	1.00	0.98	0.78	0.19	0.00
10	7	2/3	<u>1.00</u>	<u>1.00</u>	0.92	<i>0.74</i>	0.51	0.13	<u>0.04</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>
10	10	1	0.10	0.05	<u>0.03</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>	<u>0.00</u>

Notes: Numbers indicate the probability for getting at least *K* B-choices from *N* randomly selected subjects within the respective subject pools (including non-threshold players). The underlined numbers indicate situations in which success or failure of coordination can be predicted with an error rate of less than 5% in *all* locations. The italicized numbers within a column are three examples for how much behaviour can differ across subject pools in the same situation.

coordination.^{16,17} Males tend to choose the uncertain action more frequently than females, but *p* values for significance are above 10%.

Thresholds tell us how risky a subject views a situation to be. The lower the certainty equivalent of a game is, the more risk seems to be associated with it. This allows for ranking situations of strategic uncertainty and comparing them with situations of exogenously given probabilities. Figure 2 shows how subjects in Frankfurt changed their thresholds across the four decision blocks.

Result 7. Most subjects have higher certainty equivalents for games with $k = 1/3$ than for the lottery, while they have lower certainty equivalents when $k = 1$ than for the lottery.

For $k = 1/3$ in Frankfurt, 55% chose a higher threshold than in the lottery, that is this coordination game is viewed as less risky than the lottery. Only 15% reveal the opposite view. For $k = 2/3$ ($k = 1$), more than half (76%) view the lottery as less risky, while 18% (11%) show

16. We did not ask subjects to state their personal income that may be related to age among the subjects in Cologne.

17. This result is similar to the findings in ultimatum games (Güth, Schmidt and Sutter, 2003) and in dictator games (Bosch-Domènech *et al.*, 2007), in which older participants give more.

the opposite. Subjects in Barcelona, Cologne, and Bonn have similar perceptions of strategic uncertainty as subjects in Frankfurt for $k = 1/3$ and $k = 1$. However, for $k = 2/3$, the median subjects in the former locations chose the same threshold as in the lottery.

The three right bars of Figure 2 show that a vast majority of all subjects lowered their thresholds with increasing k . For only 2% of subjects, thresholds are increasing in k . Thus, subjects view a situation as more risky, when the coordination requirement rises. This is related to Result 3.

Predictability of the aggregate outcome is one of our main concerns. The more accurately the outcome of a game can be predicted, the lower strategic uncertainty is, and the more effective the design of mechanisms used to establish efficient outcomes can be. For this purpose, we analyse the distribution of individual thresholds and its implications for the probability of successful coordination in the various coordination games. The experiment covers a range of 90 binary-choice games, distinguished by group size N , hurdle k , and safe payoff X . If a subject chooses B with probability p , then the probability for at least K subjects to choose B is $1 - \text{Bin}(K - 1, N, p)$, where Bin is the cumulative binomial distribution. Replacing p with the observed proportion of B-choices within a situation, we derive an objective probability for the success of option B for randomly drawn subjects. These probabilities are given in Table 6. For these calculations, we use data from all subjects, including non-threshold players. This is done separately for each situation and location, but combining data from sessions with a different order of decision blocks in Frankfurt and Barcelona.

Result 8. Success or failure of a B-choice can be predicted with an error rate of less than 5% in more than half (66%) of the different situations within a subject pool. Across different subject pools, 49% of the games can be predicted with an error rate below 5%.

While the proportion of B-choices is almost independent of group size (as reported in Result 3), the probabilities for successful coordination depend on N . This is most pronounced when $k = 1$ (see data from Frankfurt and Bonn). However, the major influence comes from the hurdle k . It is stunning that for most coordination situations (on average 66%, with 58 (64%) of 90 in Frankfurt, 16 (53%) of 30 in Barcelona, 61 (68%) of 90 in Cologne, and 23 (77%) of 30 in Bonn), success or failure can be predicted (in sample) with an error of less than 5%. Even across subject pools, 44 (49%) of 90 of all games are predictable with an error rate below 5% (see cells with underlined numbers in Table 6). So, even if we do not know the subject pool, we can predict the outcome in about half of all coordination games, despite there being two equilibria for all games with $X < 15$.

Coordination games with $k = 1/3$ are successful with a probability of at least 95% whenever the alternative safe payoff is 7.50 or lower. Games that require coordination of all group members are successful with a probability of at most 7% when the alternative safe payoff is 7.50 or higher. These results give an impression of some circumstances under which one may expect coordination or coordination failure. For some games, however, the subject pool has extreme effects on the probability of successful coordination. Consider, for example the game with $N = 7$, $k = 2/3$, and $X = 7.50$. In Frankfurt, the probability for success is only 13%, while it is 83% in Barcelona and 68% in Cologne (see italicized numbers in column $X = 7.50$). These differences present a challenge for the out-of-sample predictive quality of the estimated models that we analyse in the next section.

5. ESTIMATING DESCRIPTIVE THEORIES FOR COORDINATION GAMES

In this section, we estimate descriptive theories for coordination games. We compare their fit with our observations and analyse how useful they are for out-of-sample predictions.

5.1. *Logit models*

The observed relative frequency of B-choices can be nicely approximated by a logistic distribution function. The same holds for the objective probability of success, as defined in Section 4. This led us to construct and test two logit models, in which individual behaviour is described by logistic functions. These estimates serve two purposes: (i) they provide a benchmark for the fit of the more theoretic and general models discussed below; and (ii) they test whether including personal characteristics improves the fit.

Models 1 and 2 describe the probability that a person chooses B by logistic functions. Model 1 uses the two significant parameters of the game, X and k , as explanatory variables. Model 2 includes players' personal characteristics, such as risk aversion, gender, age, and the Zuckerman scales (if available), as explanatory variables. As a measure of risk aversion, we use the number of B-choices in the lottery block (BL). Including all variables leads to the probability of choosing B defined by:

$$\text{Prob(B)} = \frac{1}{1 + \exp(b_0 - b_X X - b_k k - b_N N - b_r \text{BL} - b_g \text{gender} - b_a \text{age} - b_{\text{BS}} \text{BS} - b_{\text{DIS}} \text{DIS} - b_{\text{ES}} \text{ES} - b_{\text{TAS}} \text{TAS})}$$

Results of parameter estimates are reported in Table 7. The likelihood function is constructed to describe the probability of observing actual choices.¹⁸ LL is the log likelihood of observed choices; avg.LL (= LL/(30n)) is the average log likelihood per observation; and avg.l.(= 1/(exp(avg.LL))) is a measure of the predictability of an individual's decision. The reference value is 0.5 for pure randomness. n is the number of subjects in the data set. Likelihood ratio tests reveal that including personal characteristics in Model 2 increases the log likelihood significantly (at 1%) in all four locations.¹⁹

The probabilities for a subject choosing B given by the estimated coefficients of Model 1 are presented as curves in Figures 3–6 to visualize the data fit. The figures show the proportion of subjects who chose B conditional on the alternative safe payoff, denoted by X for each block k . Dots indicate the observed relative frequency of B-choices. For comparison, we include choices in the lottery block. We also checked two other functional forms, in which beliefs were modelled as logistic functions and decisions were assumed following an error response function. Beliefs could also depend on personal characteristics, as in Model 2. It turned out that these models reduce the likelihood of observations, compared to the simple logistic Model 2. Therefore, we exclude the presentation of these results. They are available on the supplements page of the *Review of Economic Studies* website at <http://www.restud.org>.

5.2. *Mixed and Bayesian Nash equilibria*

For risk-neutral players, a game is characterized by (N, K, X) . For $0 < X < 15$, there are two pure and one symmetric mixed Nash equilibria. As laid out in Section 2.1, each player chooses

18. Estimates have been computed with R, provided by R Development Core Team (2005).

19. While it is common practice in probabilistic decision models to treat all subjects as being identical, Stahl and Wilson (1994), Ho, Camerer and Weigelt (1998), Bosch *et al.* (2002), and others have demonstrated that the data fit can be significantly better, if heterogeneity among agents is considered by allowing players to differ in their levels of reasoning concerning other players' behaviour. McKelvey, Palfrey and Weber (2000) introduce heterogeneity by a distribution of error rates and estimate the parameters of this distribution. Our data allow a more sophisticated treatment of individual differences because we actually know some differences that affect behaviour. Besides, we do not see the error rate as a measure of individual error probabilities but rather as a compound including unexplained differences between subjects. Allowing for different error rates for subjects would blur this measure. Instead, we use known differences, like revealed risk aversion and other personal characteristics to explain the differences in individual choices.

TABLE 7
Logit models: results from maximum likelihood estimates

	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
b_0	-5.98	-2.64	-6.77	-4.14	-5.43	-6.18	-3.54	0.29
b_X	-0.48	-0.58	-0.44	-0.48	-0.39	-0.45	-0.29	-0.42
b_k	-3.55	-4.25	-3.70	-4.03	-2.85	-3.23	-0.51	-0.75
b_N		-0.05		-0.06				
b_r		0.54		0.31		0.02		0.96
b_g		0.53		0.28		0.58		0.21
b_a		0.04		0.06		-0.07		0.09
b_{BS}		-0.08				-0.04		-0.56
b_{DIS}		-0.04				-0.26		0.36
b_{ES}		0.22				0.54		-0.08
b_{TAS}		-0.04				-0.06		-0.23
LL	-1446.6	-1224.6	-1022.5	-944.9	-466.0	-415.0	-341.8	-245.8
avg.LL	-0.399	-0.337	-0.416	-0.384	-0.457	-0.407	-0.518	-0.372
avg.l.	0.671	0.714	0.660	0.681	0.633	0.666	0.596	0.689
Data	$n = 121$ subjects in Frankfurt		$n = 82$ subjects in Cologne		$n = 34$ subjects in Bonn		$n = 22$ subjects in Barcelona	

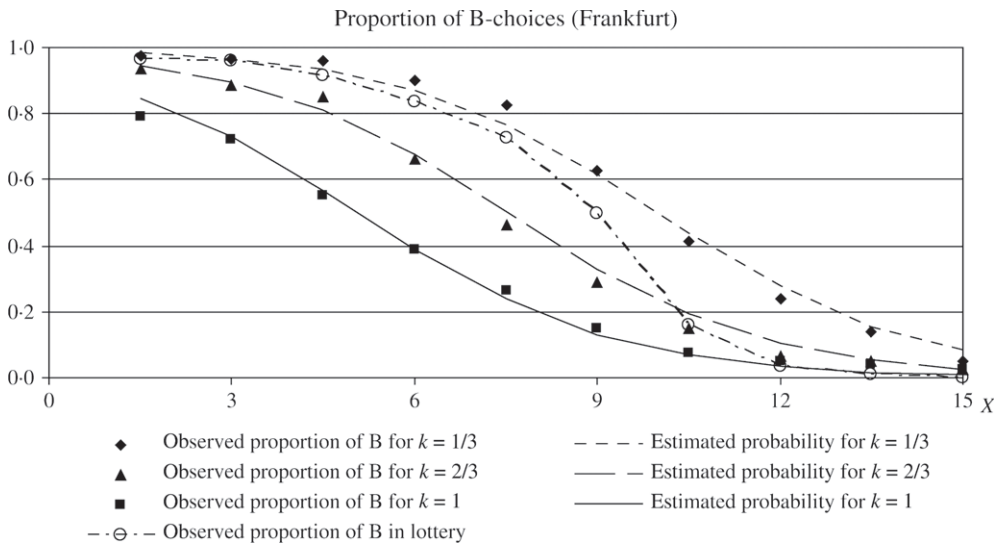


FIGURE 3
 Data from 121 subjects in Frankfurt and estimated Model 1

B with a probability p in the mixed equilibrium, solving $15(1 - \text{Bin}(K - 2, N - 1, p)) = X$. Figure 7 displays the mixed equilibria and the observed proportions of B-players for games with $N = 4$ and $K = 3$ (thus, $k = 2/3$).

While we observe that the proportion of B-choices is decreasing in X , the mixed equilibrium predicts that it rises in X . In coordination games, the mixed equilibrium is contrary to our intuition of human behaviour. It is an equilibrium in beliefs. In order to be indifferent, the expected probability of success must increase if the safe payoff rises. However, it is intuitive that players are more inclined to play B if opportunity costs are low. A similar reversal concerns the response

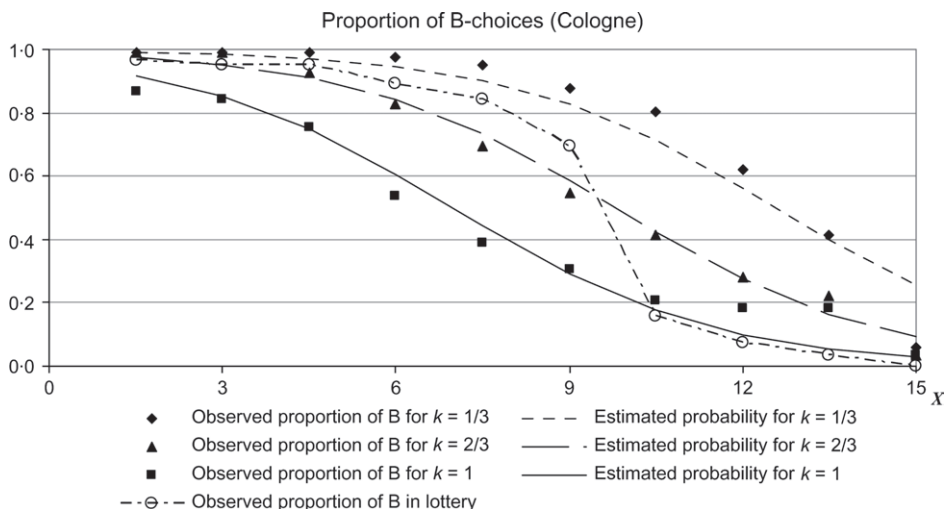


FIGURE 4
Data from 82 subjects in Cologne and estimated Model 1

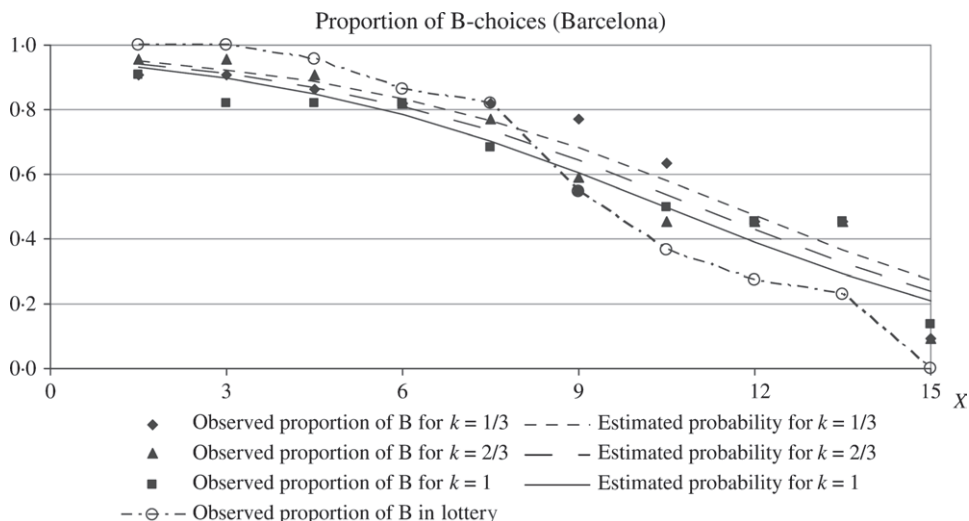


FIGURE 5
Data from 22 subjects in Barcelona and estimated Model 1

to increases in k : in the mixed equilibrium, the probability of a subject choosing B rises with k , while we observe the opposite response.²⁰

In the Bayesian game, we distinguish types by their degree of risk aversion. Using lottery choices as a measure of risk aversion, we distinguish 10 risk types by their choices in the lottery block. An increasing number of *safe* choices in the lottery block is associated with a higher degree of risk aversion α . We exclude subjects who do not follow a threshold strategy in the lottery block

20. Anderson, Goeree and Holt (2001) find that mixed Nash equilibria of a minimum-effort game have counterintuitive comparative static properties. Echeniquea and Edlin (2004) show that "in games with strict strategic complementarities, properly mixed Nash equilibria (...) are unstable for a broad class of learning dynamics".

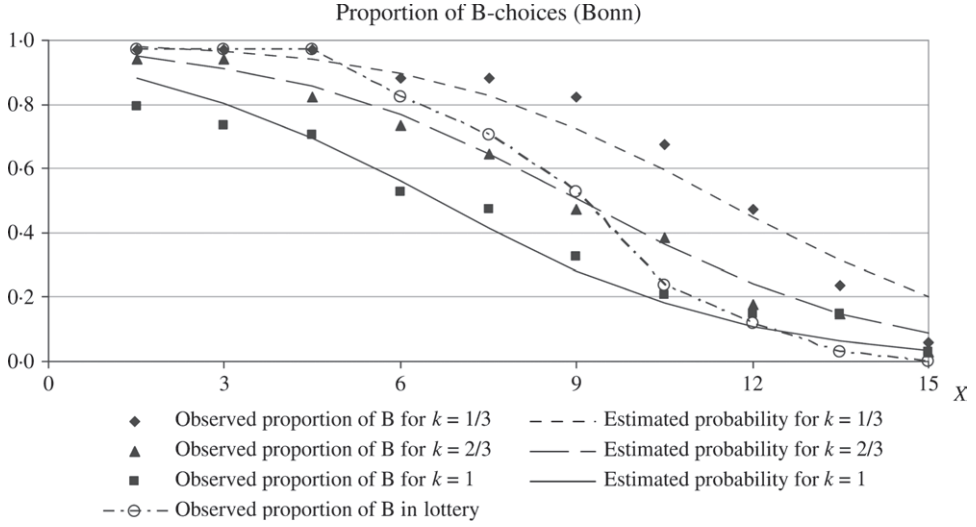


FIGURE 6
Data from 34 subjects in Bonn and estimated Model 1

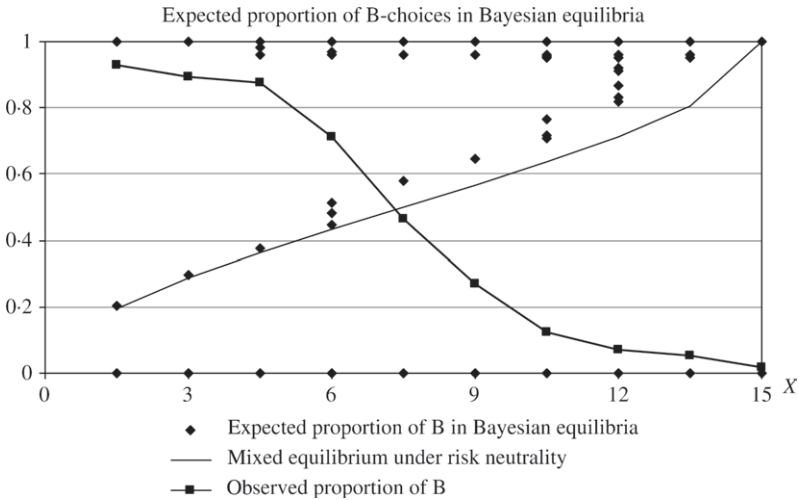


FIGURE 7
Expected proportions of B-choices in Bayesian and mixed Nash equilibria and the observed proportions of B-choices (for $N = 4, k = 2/3$, using Frankfurt data)

or who choose B for $X = 15$. As shown in Section 2.2, the Bayesian equilibrium is characterized by a pair (α^*, π) , satisfying equations (1) and (2).

Utility functions are estimated from observations in the lottery setup. We assume that players have a constant absolute risk aversion (CARA) utility function:

$$U(x) = \frac{1 - \exp(-\alpha_i x)}{\alpha_i},$$

TABLE 8

Estimation results for the global game with private information about monetary payoffs

	Frankfurt	Barcelona	Cologne	Bonn
σ	4.496	8.636	4.825	5.92
α	-0.0086	-0.0649	-0.0678	-0.054
Log likelihood (LL)	-1546.789	-366.831	-1080.934	-499.035
No. observations	3630	660	2460	1020
Average LL	-0.426	-0.556	-0.439	-0.489

where α_i is the subject's degree of absolute risk aversion. For calculating individual degrees of risk aversion α_i , we assume that a subject's true certainty equivalent is given by the median between the highest safe payoff at which the subject chooses B and the lowest safe payoff at which the subject chooses A. Thus, for a subject who chooses the lottery BL_i times, α_i is defined by the solution to:

$$1 - \exp(-\alpha_i 15) = \frac{2}{3} [1 - \exp(-\alpha_i (BL_i 1.5 + 0.75))]. \quad (7)$$

For the distribution of types, we use the observed proportions of risk types in a subject pool. Thus, $f(\alpha)$ is the proportion of players with α safe choices in the lottery block. The expected proportion of B choices is then $p(\alpha^*, \pi)$ as defined by equation (1).

Figure 7 compares $p(\alpha^*, \pi)$ for the Bayesian equilibria for $N = 4$, $k = 2/3$ with the observed proportions of B-choices in Frankfurt. The distribution of risk types is taken from observed lottery choices in Frankfurt. Pure Nash equilibria are also Bayesian. The other Bayesian equilibria are interior in the sense that the expected proportion of players choosing B is between 0 and 1. The large number of equilibria (we find up to nine Bayesian equilibria) is due to the discrete distribution of risk types.

As a descriptive theory, interior Bayesian equilibria suffer from the same reversals of responses to the exogenous parameters as the mixed Nash equilibrium since in equilibrium, probabilities must be such that some type α^* is indifferent. We conjecture that such reversals of comparative statics between theory and observations are a general phenomenon of unstable mixed Nash and interior Bayesian equilibria in games with strategic complementarities.

5.3. Global game with monetary payoff uncertainty

If players have a commonly known utility function but private information regarding monetary payoffs, the equilibrium of a game (N, K, X) is characterized by a threshold signal $Y^*(N, K, X)$ satisfying equation (3). For details, see Section 2.3. The resulting probability for a subject choosing B is given by equation (4).

The model treats the utility function and the variance of private signals as given. Using the CARA utility function $U(x) = [1 - \exp(-\alpha x)]/\alpha$, we estimate risk aversion α and the S.D. of private signals σ . We use separate maximum likelihood estimates for each subject pool. Results are reported in Table 8.

The estimated S.D. of private signals σ is measured in unit of X . Given the range of payoffs in the experiment (0–15), estimated S.D. of 4–9 are surprisingly large. They represent the magnitude of payoff uncertainty that reflects strategic uncertainty in the experiment. Of course, players know the true payoff. Their uncertainty about others' behaviour makes them behave *as if* they are uncertain about payoffs. The estimated ARA (α) is negative. In this model, the coefficient α shifts the mean threshold. The lower the ARA, the more uncertainty is accepted, and the higher is the threshold at which subjects switch from B to A. A negative estimated ARA reflects previous

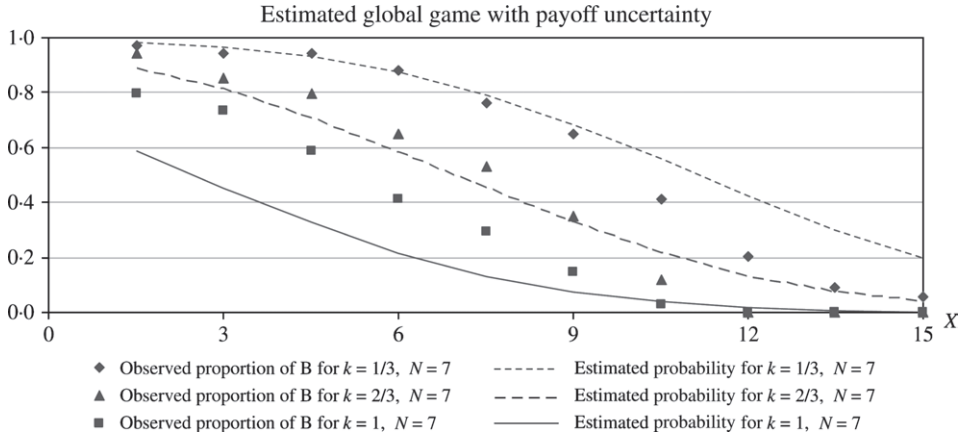


FIGURE 8

Data and estimated global-game equilibrium for $N = 7$ in Frankfurt

findings that subjects tend to deviate from risk-dominant and global-game equilibrium towards more efficient strategies.

Figure 8 shows the relative frequency of B choices given X (as dots) and the estimated probability of a player choosing B in the global-game equilibrium (as curves) for each k (Frankfurt data with $N = 7$). We see some general patterns that hold for most of the other group sizes and locations as well. For $k = 1$, the theory underpredicts the actual frequency of B. The only exception is Cologne with $N = 4$. For $k = 1/3$ and $X > 9$, the theory predicts a higher proportion of B-choices, except for the data from Cologne with $N = 7$. Changes in k and N have smaller effects on observed choices than predicted by the theory.

5.4. Global game with uncertainty about risk aversion

In Section 2.4, we presented a global game, introduced by Hellwig (2002), in which players are assumed to have common knowledge regarding monetary payoffs and a common CARA utility function, but private information concerning the degree of risk aversion. In addition, it was assumed that players make mistakes with some probability ε . As shown in Section 2.4, for $X < 15$, the equilibrium is characterized by a critical signal $\alpha^*(N, K, X, \sigma, \varepsilon)$ for which equation (6) holds. The resulting probability for a subject choosing B is given by equation (5). For $X = 15$, this probability is ε .

Fitting these probabilities $p(N, K, X, \alpha, \sigma, \varepsilon)$ to observed choices yields a joint estimate of parameters $(\alpha, \sigma, \varepsilon)$. Results are reported in Table 9. Again, we use separate maximum likelihood estimates for each subject pool. Figure 9 shows an example for the obtained data fit in Frankfurt with $N = 7$. Figures for the other locations and group sizes share the feature that the predicted probability of B-choices is smaller than the observed relative frequency if $k = 1$ and X is small ($1.5 < X < 6$), while the predicted probability is higher than the observed proportion if $k = 1/3$ and X is large ($10.5 < X < 15$). Changes in k and N have smaller effects on observed choices than predicted by the theory.

A special feature of Hellwig's model is that decisions for B at $X = 15$ are entirely attributed to mistakes, while the theory of global games predicts larger proportions here because subjects are not aware (given their private signals) that it is a weakly dominating strategy to choose A in

TABLE 9

Estimation results for the global game with private information about risk aversion

	Frankfurt	Barcelona	Cologne	Bonn
α	0.208	0.315	0.161	0.224
σ	0.124	0.220	0.143	0.161
ε	0.015	0.041	0.013	0.011
Log likelihood (LL)	-1550.5	-360.9	-1058.0	-489.2
No. observations	3630	660	2460	1020
Average LL	-0.427	-0.547	-0.430	-0.480

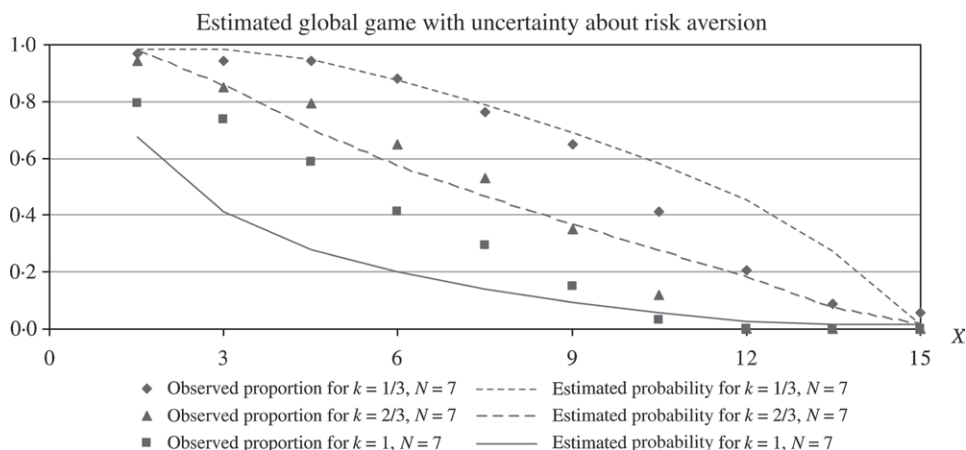


FIGURE 9

Data and estimated Hellwig model for $N = 7$ in Frankfurt

these situations. Despite the differences in modelling, both models produce very similar predictions as seen in Figures 8 and 9.²¹

5.5. Comparing the descriptive theories

In this section, we compare the models we have discussed so far. Different theories require different amounts of information for application and out-of-sample predictions. Logit models can be constructed with any available information. Model 1 uses the games' parameters. Model 2 uses personal characteristics such as risk aversion, age, and the Zuckerman scales (if available). Bayesian equilibria require knowledge of the distribution of types (risk aversion). Nash equilibria and global games do not require this information but just the parameters of the game. The equilibrium models are more general than logit models because they can be applied to games with another payoff structure than in our experiment.

To compare the quality of descriptive theories, we consider (i) the fit of data in sample, (ii) the log likelihood of out-of-sample predictions, (iii) required information, and (iv) external validity. Table 10 summarizes the log likelihoods of the estimated models in sample for all locations.

21. In the limit (when errors or variance go to 0), both models predict the same result.

TABLE 10
Comparing log likelihoods in sample

	Frankfurt	Cologne	Barcelona	Bonn
Model 1	-1447 (-0.40)	-1023 (-0.42)	-342 (-0.52)	-466 (-0.46)
Model 2	-1225 (-0.34)	-945 (-0.38)	-246 (-0.37)	-415 (-0.41)
Global game with uncertainty about monetary payoffs	-1547 (-0.43)	-1081 (-0.44)	-367 (-0.56)	-499 (-0.49)
Global game with uncertainty about risk aversion	-1550 (-0.44)	-1058 (-0.43)	-361 (-0.55)	-489 (-0.48)

Note: Numbers in parentheses indicate average log likelihood.

TABLE 11
Log likelihood of observations, given the estimated coefficients from Frankfurt or Cologne

Observations from:	Model estimated with Frankfurt data			Model estimated with Cologne data		
	Cologne [†]	Barcelona	Bonn	Frankfurt	Barcelona	Bonn
Model 1	-1219	-481	-503	-1717	-394	-478
Model 2	n.a.	-416	-561	-1461	-360	-481
Global game with uncertainty about monetary payoffs	-1317	-543	-581	-1871	-411	-507
Global game with uncertainty about risk aversion	-1278	-520	-571	-1863	-392	-495

[†]Since we had no Zuckerman test in Cologne, we cannot use Model 2 as estimated in Frankfurt (including coefficients for Zuckerman scales) for an out-of-sample test with Cologne data. n.a., not available.

Logit Model 2 yields the highest log likelihood in Frankfurt, Cologne, and Bonn. In Barcelona, the estimate of Model 2 is unreliable because the number of independent observations is small compared to the number of estimated coefficients. Model 2 requires knowledge about the distribution of personal characteristics, while the other models depend just on the games' parameters. Among the latter, Model 1 yields the highest log likelihood. However, the global games are more general. They can be applied to games with another payoff structure that are not covered by the range of games in this experiment. Comparing these two, Hellwig's version does slightly better in three of four locations, but it has three free parameters instead of two for the version with uncertainty about monetary payoffs, and since the models are not nested, likelihood ratio tests do not apply.

Comparing observations between different locations allows us to check the quality of out-of-sample predictions. Out-of-sample predictions provide a real challenge for estimated models because parameters estimated in one location differ from the optimal parameters in another location. If a model responds very sensitively to parameter changes, the estimated function leads to a low log likelihood out of sample. In comparing the quality of out-of-sample predictions, we use model estimates from the two large samples, Frankfurt and Cologne, and calculate, for each, the associated log likelihood of observations in the other locations. Table 11 summarizes log likelihoods out of sample.

Including personal characteristics (Model 2) does not generally improve the quality of out-of-sample predictions over Model 1. Comparing models in which predictions depend just on the games' parameters, Model 1 yields a higher log likelihood than a global game with uncertainty about risk aversion in five cases and just a small difference for the remaining case (-394 compared to -392 in Barcelona with estimates from Cologne). The global game with monetary

payoff uncertainty delivers a lower log likelihood than Model 1 and Hellwig's global game in all cases. Thus, it seems inferior to these models for out-of-sample predictions. However, both global games work fairly well in and out of sample. Promisingly, these general models do not perform much worse than "experienced-based" logit models.

The major difference between the two global games is that the original version (with uncertainty about monetary payoffs) predicts that each subject plays a threshold strategy, while Hellwig's model predicts that subjects deviate from a threshold strategy in every single decision with some probability ε . The estimated error probability is small (Table 9), but applied to 10 decisions, the probability that a subject uses a non-threshold strategy cannot be neglected. In Hellwig's model, the expected proportion of non-threshold players is about 11.4% in Frankfurt and 8.5% in Cologne. Indeed, we observed 10% non-threshold players in Frankfurt, but only 1 of 85 in Cologne. However, Hellwig's model predicts, additionally, that the proportion of B-choices is decreasing in both X and k even for non-threshold players, while we observe that the proportion of B-choices among non-threshold players does not respond to either X or k . Thus, it may be more appropriate to distinguish between two types of players: (i) rational agents, who use threshold strategies and whose behaviour is best explained by a global game with monetary payoff uncertainty and (ii) random players. The proportion of random players is a parameter that can be estimated by a maximum likelihood function that accounts for the threshold restriction for rational agents.

Out-of-sample tests reject the parameters estimated in one location by data from the others. But even though parameters can be rejected across subject pools, the predicted success probabilities are useful information. With an estimated probability \hat{p} of a randomly selected subject choosing B, the estimated probability for coordination on B being successful is $\hat{q} = 1 - \text{Bin}(K - 1, N, \hat{p})$. In 60% of all games, success probabilities \hat{q} that are estimated with data from Frankfurt by Model 1 deviate by less than 5% from objective probabilities of success in Cologne. To be more precise, in 21 of 30 situations in which success probabilities in Cologne are above 0.95, estimates from Frankfurt data predict a success probability above 0.95. Also, in 33 of 39 cases where Frankfurt data predict a success probability below 0.05, the objective probability in Cologne is below 0.05. However, whenever estimated success probabilities are between 0.05 and 0.95, they deviate from objective probabilities in Cologne by more than 5%. Similar results are obtained when estimates from Cologne are compared to data in Frankfurt. We conclude that observations of behaviour in coordination games are useful to detect the extreme cases in which successful coordination is very likely or very unlikely, but they do not give reliable estimates of success probabilities for intermediate cases.

Out-of-sample tests tell us which models can be applied to predict behaviour if the experiment is repeated in another location. They do not tell us how well a theory can be adopted to predict behaviour in another game. This external validity requires that a theory should be sufficiently general and coefficients can be easily adapted to another payoff structure.

Our logit models are explicitly constructed for the range of games covered in the experiment and can be applied only to environments with binary choices between a safe alternative and an option where the payoff can take on two different values, with the higher value requiring that at least a certain fraction of players should decide for this option. The estimated coefficient on X must be adjusted according to the magnitudes of the potential payoffs for B. For example, in a game paying \$10 in case of success with B, \$2 in case of failure, and \$ X for the safe option, the coefficient b_X would have to be adjusted for the relation 8:15 in differences between the two potential payoffs from option B.

As mentioned before, global games can be applied to a large class of games with strategic complementarities. However, the estimated S.D. of private signals in the global game with monetary payoff uncertainty (σ) is in units of the uncertain payoff variable. Applying this

estimate for predicting behaviour in another environment requires an adjustment to the structure of the game, in particular to the size of the interval for which the uncertain payoff variable yields multiple equilibria. The estimated parameter for risk aversion α may be used independent of the payoff structure. In Hellwig's model, the range of degrees of risk aversion, for which (given some error probability $\varepsilon > 0$) multiple equilibria exist, depends on the game's payoff structure. Hence, one cannot apply our estimates to a game with a different structure without adjusting σ and ε in a way that is far from obvious. It is an open question as to how the parameters of a global game depend on the payoff structure. Here, we have provided the first estimation of a global game. Further experiments with other game forms are needed to analyse the external validity of estimated models of strategic uncertainty.

6. BEST RESPONSE STRATEGIES

Coordination games are real-world problems, and it would behoove practitioners to know under which circumstances they should invest in a network or a new standard. When participating in a coordination game, players usually do not know success probabilities. They are in a situation of Knightian uncertainty and must rely either on experience or on abstract recommendations that provide some guidance. In this section, we determine which strategy can be recommended to a player of a one-shot coordination game and thereby contribute to a reduction of strategic uncertainty.

We evaluate and compare subjects' expected utility from their actual choices (given their revealed risk aversion and the observed distribution of others' choices) with that from strategies of different refinement concepts. We check whether subjects would have been better off responding to experience or following an abstract refinement. Recommendations to single players must account for the likely behaviour of others. For a participant, opting for B pays off if at least $K - 1$ other group members decide for B. Thus, his probability of success with B is $1 - \text{Bin}(K - 2, N - 1, p)$, where p is the probability that a randomly selected subject chooses B in this situation.²² The best response of a player is to choose B if and only if $(1 - \text{Bin}(K - 2, N - 1, p))U(15) > U(X)$, where U is the player's utility function.

In Table 12, we compare the best response to the observed distribution of choices with two refinement concepts, the GGS and the risk-dominant equilibrium. We state results for a risk-neutral player and for an agent with a moderate degree of constant absolute risk aversion (CARA). Here, we choose an ARA of $\alpha = 0.092$, which implies that the agent is indifferent between the lottery and a safe payoff of €7.50. First, we define the two strategies.

GGS(α): The GGS for diminishing variance of private signals is identical to the best response of a player who believes that the proportion of other players who choose B has a uniform distribution in $[0, 1]$.²³ As shown in Section 2.3, such a player should choose B if:

$$\left(1 - \frac{K - 1}{N}\right) U_{\alpha}(15) > U_{\alpha}(X),$$

and A if the reverse inequality holds, and is indifferent at the threshold X^* for which both sides are equal. For the latter case, we assume that a player chooses B with probability $1/2$.

RDE(α): The refinement of the risk-dominant equilibrium is the best response to a bi-centric prior or to the assumption that the probability of another player choosing B equals $U_{\alpha}(X)/U_{\alpha}(15)$.

22. For these probabilities, we use the observed proportion of B-choices among all subjects within a subject pool, including non-threshold players.

23. Hellwig (2002) has shown that the equilibrium in his model also converges to the GGS if the variance of private signals and the error probability converge to zero.

TABLE 12
Optimal and theoretical number of B-choices

	$k = 1/3$			$k = 2/3$			$k = 1$		
	$N = 4$	$N = 7$	$N = 10$	$N = 4$	$N = 7$	$N = 10$	$N = 4$	$N = 7$	$N = 10$
(i) Best response of a risk-neutral player									
In Frankfurt	7	7	7	4	4	3	2	1	0
In Barcelona		9			6			2	
In Cologne	8	9	8	5	6	5	3	2	1
In Bonn			8			5			1
Equilibrium refinements assuming risk neutrality									
GGS	7	7	6–7	4–5	4	3–4	2	1	0–1
Risk-dominant equilibrium	6	6	6	4–5	4	4	3	2	1
(ii) Best response of a player with CARA, $\alpha = 0.092$									
In Frankfurt	6	6	7	4	3	3	2	0	0
In Barcelona		8			5			1	
In Cologne	8	8	8	5	5	5	2	2	1
In Bonn			7			4			0
Equilibrium refinements assuming CARA, $\alpha = 0.092$									
GGS	5	5	5	3	2	2	1	0	0
Risk-dominant equilibrium	5	4	4	3	2	2	1	1	1

B is chosen if:

$$(1 - \text{Bin}(K - 2, N - 1, 1 - U_\alpha(X)/U_\alpha(15)))U_\alpha(15) > U_\alpha(X),$$

and A for the reverse inequality. When indifferent at threshold X^* , B is chosen with probability 1/2. The risk-dominant equilibrium is always close to the GGS. Parameters of the experiment have been chosen to yield a notable difference between the two equilibrium refinements.

Table 12 reports the best response to actual behaviour, the GGS, and the risk-dominant equilibrium. The GGS would have constituted a good recommendation for a risk-neutral player in Frankfurt. In the other locations, a risk-neutral player should have chosen higher thresholds, but the GGS would still have served as a good guide. Risk aversion lowers the optimal number of B-choices. However, with increasing risk aversion, the optimal number of B-choices falls by less than prescribed by GGS or risk dominance. In Frankfurt, a risk-averse player could have achieved a higher expected payoff by choosing thresholds that are in between the GGS calculated for risk neutrality and the GGS based on the player's own utility function. In Barcelona and Cologne, even a risk-averse player should have chosen higher thresholds than prescribed by the GGS based on risk neutrality.²⁴

Are theoretic recommendations better than out-of-sample experience? If a risk-averse subject in Cologne had known the data from Frankfurt, he or she would have been better off playing the best response to Frankfurt data than following the GGS. The reverse is not true: in Frankfurt, the GGS is closer to the optimal strategy than the best response to Cologne data. This is a surprising result because in most experiments, experience provides better guidance to optimal behaviour than any theory.

Whether risk aversion is described by CARA or by constant relative risk aversion has no big effects, either on best responses or on equilibrium refinements. This still holds, when results are calculated for the degrees of risk aversion associated with other thresholds in the lottery setup.

24. This is in line with previous observations by Heinemann *et al.* (2004) and Cabrales, Nagel and Armenter (2003), who detect systematic deviations of behaviour from the GGS towards the payoff-dominant equilibrium.

Are theoretic recommendations better than actual behaviour, so that players may expect to profit from these suggestions, and which refinement concept provides the best recommendation? To answer these questions, we calculate the expected utility that arises from these strategies for each subject. Besides GGS and risk-dominant equilibrium, we consider the following two further refinements.

P2/3(α) is the best response to other players choosing B with probability $2/3$. This strategy prescribes to choose B if:

$$(1 - \text{Bin}(K - 2, N - 1, 2/3))U_\alpha(15) > U_\alpha(X),$$

and A for the reverse inequality. When both sides are equal, B is chosen with probability $1/2$. We include this strategy because it gives the best prediction in Heinemann *et al.* (2004).

LLE(α) is the limiting logit equilibrium, introduced by McKelvey and Palfrey (1995). For any non-negative λ , a quantal response equilibrium describes the probability that a player chooses B by the solution to:

$$p(\lambda) = \frac{1}{1 + \exp(\lambda[U_\alpha(X) - (1 - \text{Bin}(K - 2, N - 1, p(\lambda))U_\alpha(15)])}.$$

The limit of the continuous path of the solution correspondence $p(\lambda)$ for $\lambda \rightarrow \infty$ defines the limiting logit equilibrium. The associated threshold X^* is given by:

$$(1 - \text{Bin}(K - 2, N - 1, 1/2))U_\alpha(15) = U_\alpha(X^*).$$

It amounts to the best response of a player who believes that others choose B with probability $1/2$.

We distinguish strategies based on a subject's own risk aversion, where ARA $\alpha = \alpha_i$ defined by equation (7), and strategies based on $\alpha = 0$, for which $U(x) = x$. Strategies based on risk neutrality are easier to calculate and do not require the knowledge of one's own risk aversion. However, neglecting risk aversion may lead to losses in expected utility.

It is informative to examine the magnitude of expected payoffs arising from these strategies. Figure 10 displays (for each location) expected payoffs associated with strategies GGS(0), RDE(0), P2/3(0), and LLE(0). They are compared with the expected payoff from a best response to actual behaviour (which is *ex ante* unknown), with average expected payoffs from subjects' actual choices, and with those from random choices with $\text{prob}(B) = 50\%$. All four theoretical strategies deliver a higher expected payoff than actual choices, which are in turn much better than random behaviour. GGS(0) and RDE(0) perform very well in all four locations by yielding at least 95% of the expected payoff from a best response.

Accounting for subjects' risk aversion, the left columns in Table 13 compare how many subjects could have improved their expected utility by choosing any of the theoretical strategies instead of their actual choices. For these comparisons, we consider all subjects who played threshold strategies and did not choose the lottery when the alternative safe payoff was 15. Success probabilities are calculated from actual choices of the respective subject pool. A vast majority of subjects could have improved their expected utility with any of the considered refinement strategies, except for P2/3(0) which would have led to an expected utility lower than or equal to that from actual behaviour for 71% of the subjects in Frankfurt.

The right three columns in Table 13 compare expected utilities from GGS(0) with expected utilities from other strategies. For most subjects in Frankfurt and Cologne, GGS(0) would have led to a higher expected payoff than any other strategy in pair-wise comparisons. In Barcelona, however, a majority of subjects would have achieved a higher expected utility using P2/3(0) or P2/3(α_i). In Bonn, a majority would have been better off with strategies RDE(0) or RDE(α_i).

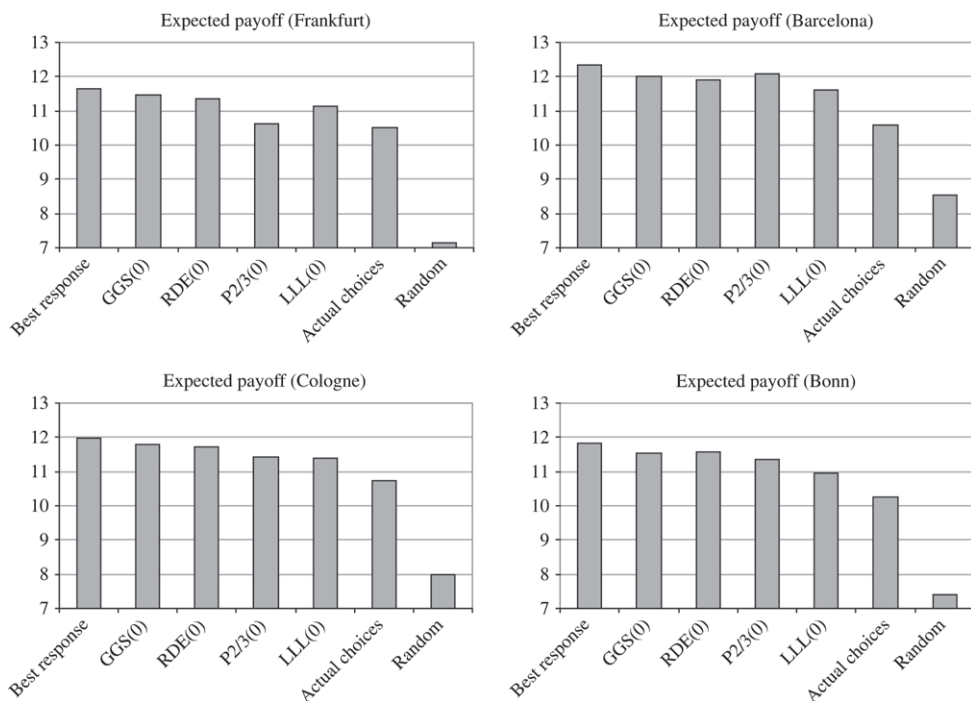


FIGURE 10
Expected payoffs from various strategies

TABLE 13
Comparison of expected utilities arising from different refinements and actual choices

Refinement strategy	Percentage of subjects who would have achieved higher expected utility using the refinement strategy instead of their actual choices				Percentage of subjects for whom GGS(0) would have led to a higher expected utility than the other refinement strategy			
	Frankfurt (%)	Barcelona (%)	Cologne (%)	Bonn (%)	Frankfurt (%)	Barcelona (%)	Cologne (%)	Bonn (%)
GGs(0)	91	74	81	94				
RDE(0)	81	65	76	91	100	100	74	29
P2/3(0)	29	74	60	76/	99	43	74	88
LL(0)	67	52	63	65	94	85	95	100
GGs(α_i)	78	74	75	85	65	61	81	74
RDE(α_i)	73	78	73	85	94	61	71	44
P2/3(α_i)	43	87	70	85	91	43	65	94
LL(α_i)	64	70	60	71	94	63	89	85
No. subjects	124	23	84	34	124	23	84	34

Note: These comparisons include subjects who did not complete the questionnaire.

A recommendation can account for a subject’s degree of risk aversion. Table 14 shows, for each risk type, which strategy would have yielded the highest expected utility in the various coordination game setups. GGS(0) is the main competitor of RDE(0) and P2/3(0) for being the optimal strategy of approximately risk-neutral agents ($\alpha_i = 0.01$). A moderate degree of risk

TABLE 14
Best strategies depending on an agent's degree of ARA in the various environments

No. B-choices in lottery block	ARA α_i	Frankfurt			Barcelona	Cologne			Bonn
		$N = 4$	$N = 7$	$N = 10$	$N = 7$	$N = 4$	$N = 7$	$N = 10$	$N = 10$
0	1.465	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$
1	0.488	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$	$\underline{GGS(0)}$	$P2/3(\alpha_i)$	$\underline{GGS(0)}$	$P2/3(\alpha_i)$	$\underline{GGS(0)}$
2	0.286	$P2/3(\alpha_i)$	$P2/3(\alpha_i)$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$RDE(0)$	$\underline{GGS(0)}$
3	0.187	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$
4	0.121	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$
5	0.065	$\underline{GGS(0)}$	$\underline{GGS(\alpha_i)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$RDE(0)$	$\underline{RDE(0)}$
6	0.010	$\underline{GGS(\alpha_i)}$	$\underline{GGS(0)}$	$\underline{GGS(\alpha_i)}$	$P2/3(0)$	$\underline{GGS(0)}$	$P2/3(0)$	$RDE(0)$	$\underline{RDE(0)}$
7	-0.065	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$= P2/3(\alpha_i)$	$\underline{GGS(\alpha_i)}$	$= P2/3(\alpha_i)$	$= RDE(\alpha_i)$	$= RDE(\alpha_i)$
8	-0.161	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$\underline{GGS(0)}$	$RDE(\alpha_i)$	$\underline{GGS(\alpha_i)}$	$\underline{GGS(\alpha_i)}$	$\underline{GGS(\alpha_i)}$	$\underline{GGS(\alpha_i)}$
9	-0.540	$P2/3(0)$	$P2/3(0)$	$P2/3(0)$	$\underline{GGS(\alpha_i)}$	$P2/3(0)$	$\underline{GGS(\alpha_i)}$	$\underline{GGS(\alpha_i)}$	$P2/3(0)$
								$= RDE(\alpha_i)$	

Notes: The underlined entries indicate GGS(0). In some cases, two strategies give identical prescriptions.

aversion favours GGS(0). For agents with an extremely high risk aversion, $P2/3(\alpha_i)$ may have been a better recommendation. Risk-lovers might have benefited more from $\underline{GGS(\alpha_i)}$, $RDE(\alpha_i)$, or $P2/3(0)$. However, note that differences between the best two or three strategies usually would have been very small in magnitude because, for most games, they would have given the same recommendation.

It is a striking and surprising result that $\underline{GGS(0)}$ does so well in particular in the two large subject pools, although aggregate behaviour differs significantly between these locations. The threshold associated with $\underline{GGS(0)}$ is given simply by:

$$X^* = 15 \left(1 - \frac{K-1}{N} \right).$$

This strategy is easy to calculate and does not even require the knowledge of one's own risk aversion. For other binary-choice games with strategic complementarities, $\underline{GGS(0)}$ is given by the best response to a uniform distribution of the proportion of other players choosing either of the two alternatives.

Concerning external validity: When payoffs are scaled up, one should expect thresholds to fall because Holt and Laury (2002) show that risk aversion rises with higher payoffs. Given our results on the close relationship between risk aversion and thresholds in coordination games, we expect high-scale payoffs to also reduce certainty equivalents for strategic games. This effect shows up in the data from two high-stake coordination games in Heinemann *et al.* (2004). As a consequence, we expect that strategies that account for risk aversion may be better than $\underline{GGS(0)}$ in experiments with higher payoffs. Firms should be less risk-averse than subjects in an experiment. When advising a player in an environment with low risk aversion, we recommend strategies that deviate from the \underline{GGS} towards more efficient strategies.

7. SUBJECTIVE BELIEFS

In the questionnaire, 83.6% of all subjects answered yes to the question of whether they expected other subjects to play threshold strategies. In this section, we discuss what we can learn from the experiment about subjective expectations in coordination games. There are two ways of

measuring subjective beliefs: asking subjects directly and rewarding them if their beliefs came close to observations (direct mechanism) or deducing beliefs from their actions (indirect mechanism). In standard sessions, we did not ask for beliefs because we did not want to impose probabilistic beliefs and because staking a price on predictions changes the game and might affect behaviour. In standard session, we can deduce beliefs from actions, using the hypothesis that subjects' actions are best responses to their beliefs.

Two control sessions in Bonn asked for beliefs directly. With these data, we analyse whether actions are best responses to stated beliefs and we compare subjective probabilities with observed frequencies. We can also check another dimension of strategic uncertainty: in which situations do subjects disagree most about the likely behaviour of others.

7.1. *Deduced beliefs*

A measure of subjective beliefs can be deduced from behaviour by assuming a particular utility function. Again, we use the CARA utility function, but we confirmed that the CRRA utility function yields approximately the same results.

Let X_L be a subject's certainty equivalent of the lottery and X_c his or her threshold in a coordination game. Then, $U(X_L) = 2/3 U(15) + 1/3 U(0)$ and $U(X_c) = qU(15) + (1 - q)U(0)$, where q is the subjective probability for successful coordination on B when the alternative safe payoff from A is X_c . Solving the second equation, we get:

$$q = \frac{U(15) - U(0)}{U(X_c) - U(0)}.$$

Replacing $U(15)$ by the first equation yields a measure for subjective beliefs:

$$q(X_L, X_c) = \frac{2}{3} \cdot \frac{U(X_c) - U(0)}{U(X_L) - U(0)}.$$

Note that $q(X_L, X_c)$ is increasing in X_c and decreasing in X_L . In our experiment, we measure certainty equivalents only in intervals of 1.50. Consider a subject who chooses the lottery B when the payoff for A is smaller or equal to X_L euros, but chooses B in a coordination setup when the safe alternative is smaller or equal to X_c euros. We can deduce that $q(X_L + 1.5, X_c) < q(X_L, X_c) < q(X_L, X_c + 1.5)$.

We say that a subject overestimates the probability for successful coordination if $q(X_L + 1.5, X_c)$ exceeds the objective probability of success in the coordination game with safe payoff X_c . We say that a subject underestimates the probability for successful coordination if $q(X_L, X_c + 1.50)$ is lower than the objective probability of success in the coordination game.

Subjects for whom neither of the two conditions mentioned above holds are said to have subjective probabilities that are approximately equal to the objective ones. Table 15 presents these comparisons and shows that most subjects overestimate the probability of success in games with a high coordination requirement but underestimate success in games with a low hurdle.

Most subjects underestimate the probability of successful coordination when they need only one-third of the other players to be successful, while most subjects overestimate the probability of successful coordination in games with $k = 2/3$ or $k = 1$. The proportion of subjects who overestimate probabilities of success in the coordination game tends to rise in k and N , while the proportion of subjects who underestimate success probabilities tends to fall in k and N .

It remains an open question as to whether subjective beliefs are formed over the outcome of the order statistic (here, success or failure of coordination on B) or over the individual choices of other players. Our experiment provides some answers to this question. If player i attributes subjective probability p_i to another randomly selected subject choosing B, then his or her subjective

TABLE 15

Estimated number of subjects, who overestimate or underestimate the probability of successful coordination

Game	$k = 1/3$			$k = 2/3$			$k = 1$		
	$N = 4$	$N = 7$	$N = 10$	$N = 4$	$N = 7$	$N = 10$	$N = 4$	$N = 7$	$N = 10$
Subjects who overestimate success probability	12	6	10	23	19	18	36	26	25
	23%			48%			70%		
Subjective probability approximately equal to objective	18	13	6	17	4	11	7	6	7
	31%			26%			16%		
Subjects who underestimate success probability	26	14	16	16	12	5	13	3	2
	46%			26%			16%		

Note: Data are from sessions in Frankfurt.

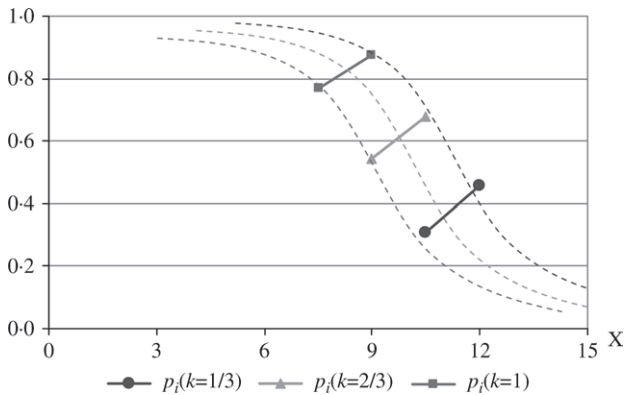


FIGURE 11

Subjective probabilities p_i for a particular subject of another subject choosing B

probability for success with action B is $q_i = 1 - \text{Bin}(K - 2, N - 1, p_i)$. This function is invertible, so that we can deduce the subjective probability for another player choosing B by the value p_i that solves the equation for the estimated q_i .

For each subject, we obtain intervals for subjective probabilities at three different combinations of k and X_c . For a single subject, p_i and q_i should decrease with rising k and X_c . When we increase the hurdle k , subjects respond with a lower threshold X_c . If the direct effect of rising k dominates the indirect effect from the associated change in the threshold, beliefs are increasing in k . Otherwise, they are decreasing.

For more than 70% of subjects, estimated values of q_i at the respective threshold are decreasing in k ; in Frankfurt, there were only three cases with a reverse order. However, estimated values for p_i are increasing in k for two-thirds of all subjects. A typical example is a risk-neutral subject who chose B seven, six, and five times in the setups with $k = 1/3, 2/3$, and 1, respectively. Estimated subjective beliefs for another player choosing B are inside the intervals depicted by thick lines in Figure 11. Beliefs as functions of X for a particular k should be decreasing in

X and pass through the respective interval. In addition, the belief function for $k = 1/3$ should be above that for $k = 2/3$, which itself should be above that for $k = 1$, as is shown by the broken curves in Figure 11. This requires that beliefs should be rather steep functions and minimally affected by changes in k . This is in stark contrast to the observed distribution of choices (displayed in Figures 3–6) that is much flatter in X and more sensitive to k . In Frankfurt, 47 of 121 subjects changed their number of B-choices by at most two between $k = 1/3$ and $k = 1$, analogously to the subject displayed in Figure 11.

There are two possible explanations for this result: (i) behaviour might be inconsistent with probabilistic beliefs about individual behaviour and (ii) subjects might be overly confident in their assessment of others' behaviour. Two control sessions in Bonn, in which we asked subjects to state their beliefs, were designed to shed more light on this question.

7.2. Stated beliefs

In one session [Individual Beliefs (IB)], we asked subjects to state a probability p of a randomly selected player other than one's self choosing B. In the other session [Group belief (GB)], we asked subjects to state a probability q for success of B in each coordination game. Theoretically, $q = 1 - \text{Bin}(K - 2, N - 1, p)$. Since both sessions used different elicitation procedures, we first check whether behaviour is comparable between the two sessions. Comparing the number of B-choices in each block between these two sessions with a Mann-Whitney U test, we find no significant differences in behaviour. Comparing the thresholds in coordination games conducted in Bonn with the other sessions where $N = 10$, we find significant differences between Bonn and Frankfurt but not between Bonn and Cologne (see Result 2). The sessions in Bonn are the only ones without correlation between B-choices in lottery setups and coordination games (Result 5).

In lottery situations, subjects were asked in both sessions to estimate the probability with which another player would choose B. There is a significant correlation (at 1%) between subjects' choices and beliefs about others' choices: risk-averse subjects expect a higher average risk aversion than risk-neutral or risk-loving subjects. There is no correlation between stated beliefs in lottery choices and in coordination games with $k = 1$ or $k = 2/3$. There is a significant correlation (at 1%) between beliefs in lotteries and in games with $k = 1/3$. Note that in these sessions, the lottery block immediately followed the decision block with $k = 1/3$. The correlation of beliefs between these blocks may be a framing effect. There is no significant correlation ($p > 25\%$) between lottery choices and choices in coordination games.

In coordination games, stated beliefs and choices are highly correlated ($p < 0.00001$). Subjects who state a higher success probability for B are more likely to choose B. However, subjects do not always choose a best response to their beliefs. Using ARA utility functions with an interval of compatible risk aversion defined by the revealed certainty equivalents in lottery choices, we calculate the best response to stated beliefs. As risk aversion is potentially defined over an interval instead of a point value, both options can be consistent with stated beliefs in some situations. In Session IB, 40% of chosen thresholds are consistent with beliefs, 35% of the chosen thresholds are higher than they should be in a best response to stated beliefs, and 25% are lower. In Session GB, 37% of chosen thresholds are consistent with stated beliefs, 50% are too high, and 13% too low. Figure 12 displays these deviations. These results do not differ significantly under the assumption that all subjects are risk-neutral.

There is no significant difference (at the 10% level using a one-sided Mann-Whitney U test) in the distribution of these deviations between the two sessions. Thus, we cannot support the hypothesis that subjects make more mistakes when responding to beliefs about individual behaviour than to success probabilities. This is insofar surprising as the expected payoff from

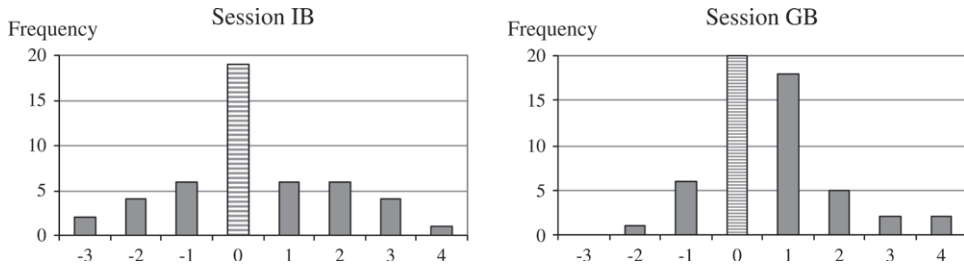


FIGURE 12

Frequency of differences between actual number of B-choices and best response to stated beliefs in decision blocks with coordination games. For example, the column above “4” indicates the frequency with which subjects chose B four times too often compared to their stated beliefs

option B is rather easy to calculate for a given success probability. It is much more complicated to calculate the binomial function, given a probability for individual behaviour.

The strong correlation between beliefs and actions shows that subjects respond to their beliefs. Rey-Biel (2006) and Costa-Gomes and Weizsäcker (2007) test best response rates to stated beliefs in one-shot 3×3 games. They find best response rates ranging from 55% to 73% that are significantly higher than for random behaviour. In our game, random behaviour is a poor reference point, and we know already that subjects do not choose their thresholds randomly because these thresholds respond to k . Comparing chosen thresholds with best responses to other subjects’ stated beliefs leads to much larger differences than in comparison with best responses to their own stated beliefs. This also demonstrates the close relation between stated beliefs and actions.

In the GB session, subjects often chose higher thresholds than in a best response to their beliefs, even though the expected payoff was easy to calculate. An explanation might be that subjects want to be “kind” and opt for the efficient outcome out of an altruistic motive even when they attribute a low probability to success. If this is generally true, then subjective probabilities that are deduced from observed behaviour overestimate true beliefs. In the IB session, however, chosen thresholds are almost symmetrically distributed around the best response to stated beliefs. The “kindness” effect might be cancelled out by mistakes if subjects systematically underestimate the probability of success for a given probability of an individual B-choice (effect of the binomial function). However, comparing the stated beliefs between the two sessions hints at the opposite mistake.²⁵

If subjects perceive the mathematical relation between the two probabilities correctly, the aggregate distribution of stated beliefs should differ between the two sessions. The probability that all the nine (other) subjects choose B is smaller than the probability that a single, randomly selected (other) subject chooses B. Hence, we should expect that, for $k = 1$, stated probabilities for success in the GB session should be lower than stated probabilities in the IB session. In games with a low hurdle ($k = 1/3$), we should observe the opposite relation because it is more likely that at least three of nine choose B than a single player chooses B.

25. Ambiguity aversion can also be a reason for deviations between actual choices and the best response to stated beliefs because it leads to behaviour consistent with subadditive probabilities. If stated probabilities are not subadditive, subjects would tend to choose a lower threshold than in a best response to beliefs. Camerer and Karjalainen (1994) conduct a series of experiments to measure ambiguity aversion. They find only small degrees of subadditivity. However, Hsu *et al.* (2005) show that different degrees of ambiguity change brain activity in the amygdalae and the orbitofrontal cortex.

In none of the 10 situations with $k = 1$ did we find a difference between the stated beliefs of the two sessions that is significant at a 10% level. In situations with $k = 2/3$, we found a significant difference (at 5%) only for $X = 15$. At a 10% level, there were three other situations (of the 10) with significant differences pointing in the right direction. In situations with $k = 1/3$, we found two situations with differences significant at 9%. Taken together, these results indicate that subjects state almost the same probabilities whether asked for individual behaviour or for successful coordination. They seem to ignore the effect of the binomial function that relates these probabilities.

If we calculate success probabilities from stated individual beliefs (using the binomial function), these probabilities are significantly smaller than those of the directly elicited success probabilities when all members of a group are needed ($k = 1$) for success and vice versa when $k = 1/3$; there are no significant differences when $k = 2/3$. This is a direct consequence from the observation that the stated beliefs are the same in both sessions.

Comparing beliefs with objective probabilities as derived from observed behaviour, average stated probabilities (in both sessions) are close to the proportion of B-choices but not to the probability of success. Figure 13 compares the average stated beliefs (dots) from both sessions with the proportion of B-choices (thin curves) and with the objective success probability (thick curves). For this figure, we used the data from all 40 subjects in Bonn.

Stated beliefs respond to k and X in the right direction, and deviations of behaviour from best responses to stated beliefs are the same for both treatments. In this respect, we cannot support the hypothesis that behaviour is less consistent with beliefs about individual behaviour than with beliefs about success. However, subjects do not differentiate between the two concepts. In this respect, they make mistakes. When asked about individual behaviour, they give (on average) a very good estimate. But individuals are overconfident in the sense that, for each subject, stated probabilities decrease rapidly within a small range of X . Average stated beliefs (the dots in Figure 13) are flatter because subjects differ in their opinions about the values of X for which the probability should decrease the most. This holds for both sessions.

In Table 16, we compare how many subjects overestimated or underestimated probabilities, compared with objective probabilities. In the IB session, a majority overestimates the probability

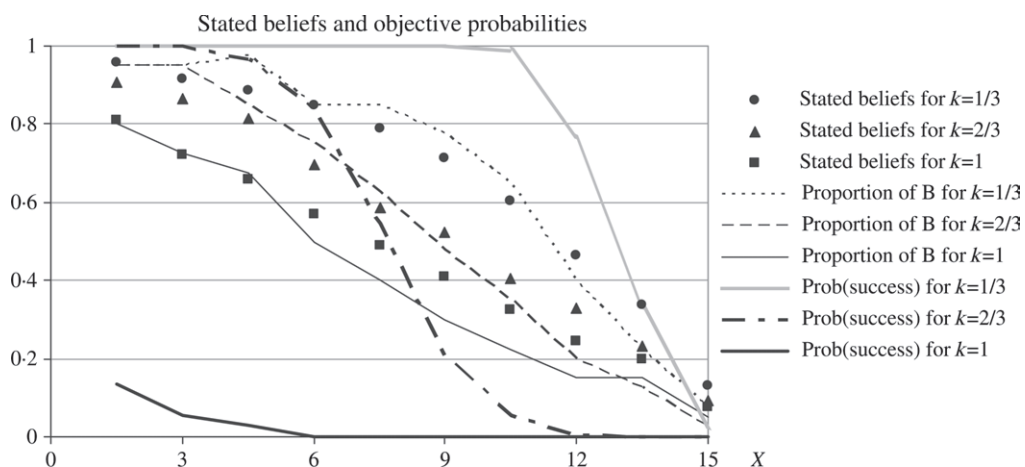


FIGURE 13

Stated beliefs, observed proportions of B-choices (thin curves), and derived success probabilities (thick curves). The highest curves refer to $k = 1/3$, the lowest to $k = 1$

TABLE 16

Percentage of stated beliefs that are above, equal, or below objective probabilities for success, given actual behaviour

	IB (%)			GB (%)		
	$k = 1/3$	$k = 2/3$	$k = 1$	$k = 1/3$	$k = 2/3$	$k = 1$
Stated belief higher than objective probability	55	59	67	23	38	78
Stated belief equal to objective probability	8	9	1	22	13	0
Stated belief lower than objective probability	37	32	32	54	48	22

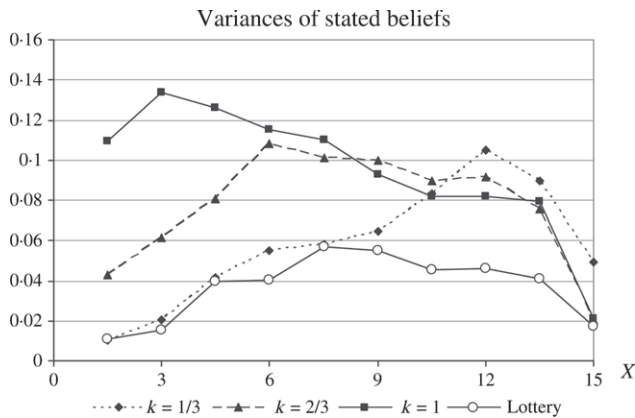


FIGURE 14

Variances of stated beliefs conditional on the situation

that another subject chooses B, while in the GB session, results are similar to those from deduced beliefs about success reported in Table 15: For $k = 1$, most subjects overestimate success probability, and for $k = 1/3$, most subjects underestimate success probability.

A last aspect of strategic uncertainty that we want to mention concerns the differences between individual beliefs. Figure 14 displays the variances of stated beliefs for each situation. For coordination games, we display the median of the two variances associated with IB and GB sessions. The overall picture is about the same if we restrict data to either session.

For each X , the dispersion of beliefs is lowest in the lottery setup. The dispersion within a block is an inverse U-shape, with the highest point at $X = 3$ for $k = 1$, at $X = 6$ for $k = 2/3$, and at $X = 12$ for $k = 1/3$. These are situations in which we had the largest differences in success probabilities across subject pools (Table 6). Subjects disagree on the expected behaviour of others, in particular, in those situations in which experimental results leave us with the highest uncertainty about which outcome to expect.

This hints at a possible instrument to predict regime changes in models of financial crises. If the variances of traders' expectations are rising, the economy is endangered by a switch from one equilibrium to another.

8. CONCLUSIONS

We designed an experiment that allows for the measurement of strategic uncertainty and the estimation of subjective probabilities in coordination games with multiple equilibria. Strategic uncertainty associated with a certain number of group members required to coordinate is measured by a certainty equivalent, that is the certain payoff that a subject is willing to forgo for the uncertain payoff from coordination. The lower the certainty equivalent of a coordination requirement is, the more risk seems to be associated with it.

The outcome of a coordination game with multiple equilibria can be highly predictable, especially when the attitudes of a population towards risk and strategic uncertainty are known. The same knowledge allows recommendations for behaviour to be formulated and will, thereby, enhance efficiency in the process of achieving coordination. Without precise knowledge of the environment, the GGS can be recommended to agents who are engaged in one-shot coordination games. Note that this is advice for a single agent. Advice given to the whole group should always try to achieve a more efficient outcome.

Strategic uncertainty can be modelled by global games with private information about monetary payoffs or private information about risk aversion. Two models with this feature deliver good descriptions in sample and good predictions out of sample. In contrast to equilibrium refinements, these theories allow for behaviour to not be fully coordinated. We have shown how to estimate the parameters of these models, a procedure that can be applied to other games with strategic complementarities as well. Thereby, we estimate the distribution of private signals, which, in the literature thus far, has been taken as given exogenously. The quality of a descriptive model can be improved if individual characteristics are taken into account. However, they are less useful for out-of-sample predictions.

The design of our experiment points the way for measuring strategic uncertainty in other games as well. A generic approach would ask subjects to decide between safe payoffs of various amounts or lotteries with different success probabilities on the one side and participation in a strategic game on the other. Should a subject decide to participate in the game, he or she must also state his or her chosen action in the game. His or her beliefs about the payoff from the strategic game can then be measured by the marginal payoff or lottery, at which he or she switches actions. This procedure can actually be applied to a wide variety of games. Analysing strategic uncertainty assists in the forecasting of behaviour and in the formulation of advice to players.

Acknowledgements. For valuable comments, we are grateful to Fabrizio Germano; Itay Goldstein; Werner Güth; Alexander Meyer-Gohde; Stephen Morris; Christian Schade; Burkhard Schipper; Andreas Schröder; Joachim Weimann; various seminar participants in Barcelona, Darmstadt, Exeter, Frankfurt, and at the game theory conference in Stony Brook; the editor of this issue, Juuso Välimäki; and two anonymous referees. We thank Irina Cojuharenco and Aniol Llorente-Saguer for technical assistance. All errors remain our responsibility. Frank Heinemann acknowledges financial support by the German Research Foundation through the SFB 649 “Economic Risk”. Rosemarie Nagel acknowledges financial support of Spain’s Ministry of Education under Grant SEC2002-03403 and thanks the Barcelona Economics Program of CREA and Human Frontier Science Program for support. Nagel also thanks CES-München and Caltech for their hospitality.

REFERENCES

- ANDERSON, S. P., GOEREE, J. K. and HOLT, C. A. (2001), “Minimum-Effort Coordination Games: Stochastic Potential and Logit Equilibrium”, *Games and Economic Behavior*, **34**, 177–199.
- AUMANN, R. J. and DREZE, J. H. (2004), “Assessing Strategic Risk” (<http://ideas.repec.org/p/huj/dispa/p/dp361.html>).
- BERNINGHAUS, S. and ERHART, K.-M. (1998), “Time Horizon and Equilibrium Selection in Tacit Coordination Games: Experimental Results”, *Journal of Economic Behavior and Organization*, **37**, 231–248.
- BERNINGHAUS, S. and ERHART, K.-M. (2001), “Coordination and Information: Recent Experimental Evidence”, *Economic Letters*, **73**, 345–351.
- BOHNET, I. and ZECKHAUSER, R. (2004), “Trust, Risk and Betrayal”, *Journal of Economic Behavior and Organization*, **55**, 467–484.

- BOSCH-DOMÈNECH, A., MONTALVO, J. G., NAGEL, R. and SATORRA, A. (2002), "One, Two, Three, Infinity, ...: Newspaper and Lab Beauty-Contest Experiments", *American Economic Review*, **92**, 1687–1701.
- BOSCH-DOMÈNECH, A., NAGEL, R., SÁNCHEZ-ANDRÉS, J. V. and MORERA, J. (2007), "Preserved Social Capabilities in Alzheimer's Patients" (Working Paper, Universitat Pompeu Fabra 1020).
- BRANDENBURGER, A. (1996), "Strategic and Structural Uncertainty in Games", in R. Zeckhauser, R. Keeney and J. Sibenius (eds.) *Wise Choices: Games, Decisions, and Negotiations* (Boston: Harvard Business School Press) 221–232.
- CABRALES, A., NAGEL, R. and ARMENTER, R. (2007), "Equilibrium Selection through Incomplete Information in Coordination Games: An Experimental Study", *Experimental Economics*, **10**, 221–234.
- CAMERER, C. and KARJALAINEN, R. (1994), "Ambiguity-Aversion and Non-Additive Beliefs in Non-Cooperative Games: Experimental Evidence", in B. Munier and M. J. Machina (eds.) *Models and Experiments in Risk and Rationality* (Dordrecht: Kluwer Academic Publishers) 325–358.
- CARLSSON, H. and VAN DAMME, E. (1993), "Global Games and Equilibrium Selection", *Econometrica*, **61**, 989–1018.
- COOPER, R. W. (1999), *Coordination Games: Complementarities and Macroeconomics* (Cambridge, UK: Cambridge University Press).
- COSTA-GOMEZ, M. A. and WEIZSÄCKER, G. (2007), "Stated Beliefs and Play in Normal-Form Games", *Review of Economic Studies*, **75**, 729–762.
- CROSON, R. (2000), "Thinking Like a Game Theorist: Factors Affecting the Frequency of Equilibrium Play", *Journal of Economic Behavior and Organization*, **41**, 299–314.
- ECHENIQUE, F. and EDLIN, A. (2004), "Mixed Equilibria are Unstable in Games of Strategic Complements", *Journal of Economic Theory*, **118**, 61–79.
- FISCHBACHER, U. (2007), "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments", *Experimental Economics*, **10** (2), 171–178.
- FRANKEL, D., MORRIS, S. and PAUZNER, A. (2003), "Equilibrium Selection in Global Games with Strategic Complementarities", *Journal of Economic Theory*, **108**, 1–44.
- GOEREE, J. K. and HOLT, C. A. (2005), "An Explanation of Anomalous Behavior in Models of Political Participation", *American Political Science Review*, **99**, 201–213.
- GOEREE, J. K., HOLT, C. A. and PALFREY, T. R. (2003), "Risk Averse Behavior in Generalized Matching Pennies Games", *Games and Economic Behavior*, **46**, 97–113.
- GROSSKOPF, B. and NAGEL, R. (2008), "The Two-Person Beauty Contest", *Games and Economic Behavior*, **62**, 93–99.
- GÜTH, W., SCHMIDT, C. and SUTTER, M. (2003), "Fairness in the Mail and Opportunism in the Internet—A Newspaper Experiment on Ultimatum Bargaining", *German Economic Review*, **4**, 243–265.
- HARSANYI, J. C. and SELTEN, R. (1988), *A General Theory of Equilibrium Selection in Games* (Cambridge, MA: MIT Press).
- HEINEMANN, F. (2008), "Measuring Risk Aversion and the Wealth Effect", in J. C. Cox and G. W. Harrison (eds.) *Risk Aversion in Experiments* (Bingley, UK: Emerald), *Research in Experimental Economics*, **12** (forthcoming).
- HEINEMANN, F., NAGEL, R. and OCKENFELS, P. (2004), "Global Games on Test: Experimental Analysis of Coordination Games with Public and Private Information", *Econometrica*, **72**, 1583–1599.
- HELLWIG, C. (2002), "Imperfect Common Knowledge of Preferences in Coordination Games" (mimeo, University of California, Los Angeles).
- HO, T., CAMERER, C. and WEIGELT, K. (1998), "Iterated Dominance and Iterated Best-Response in Experimental 'P-Beauty-contests'", *American Economic Review*, **88**, 947–969.
- HOLT, C. A. and LAURY, S. K. (2002), "Risk Aversion and Incentive Effects", *American Economic Review*, **92**, 1644–1655.
- HSU, M., BHATT, M., ADOLPHS, R., TRANEL, D. and CAMERER, C. F. (2005), "Neural Systems Responding to Degrees of Uncertainty in Human Decision-Making", *Science*, **310** (5754), 1680–1683.
- KNIGHT, F. H. (1921), *Risk, Uncertainty and Profit* (Boston: Houghton Mifflin).
- LANGE, A., LIST, J. A. and PRICE, M. K. (2004), "Auctions with Resale When Private Values are Uncertain: Theory and Empirical Evidence" (NBER Paper No. 10639).
- LUCE, D. and RAIFFA, H. (1957), *Games and Decisions* (New York: John Wiley & Sons).
- MANSKI, C. F. (2004), "Measuring Expectations", *Econometrica*, **72**, 1329–1376.
- McKELVEY, R. D. and PALFREY, T. R. (1995), "Quantal Response Equilibria for Normal Form Games", *Games and Economic Behavior*, **10**, 6–38.
- McKELVEY, R. D., PALFREY, T. and WEBER, R. A. (2000), "The Effects of Payoff Magnitude and Heterogeneity on Behavior in 2 x 2 Games with Unique Mixed Strategy Equilibria", *Journal of Economic Behavior & Organization*, **42**, 523–548.
- MORRIS, S. and SHIN, H. S. (2002), "Measuring Strategic Uncertainty" (<http://hyunsongshin.org/www/barcelona.pdf>).
- MORRIS, S. and SHIN, H. S. (2003), "Global Games: Theory and Applications", in M. Dewatripont, L. Hansen and S. Turnovsky (eds.) *Advances in Economics and Econometrics, the Eighth World Congress* (New York: Cambridge University Press).
- NYARKO, Y. and SCHOTTER, A. (2002), "An Experimental Study of Belief Learning Using Elicited Beliefs", *Econometrica*, **70**, 971–1005.

- R DEVELOPMENT CORE TEAM (2005), *R: A Language and Environment for Statistical Computing*, (Vienna, Austria: R Foundation for Statistical Computing; <http://www.R-project.org>).
- REY-BIEL, P. (2006), "Equilibrium Play and Best Response to (Stated) Beliefs in Constant-Sum Games" *Games and Economic Behavior* (forthcoming).
- SCHECHTER, L. (2007), "Traditional Trust Measurement and the Risk Confound: An Experiment in Rural Paraguay", *Journal of Economic Behavior and Organization*, **62**, 272–292.
- SCHMIDT, D., SHUPP, R., WALKER, J. M. and OSTROM, E. (2003), "Playing Safe in Coordination Games: The Roles of Risk Dominance, Payoff Dominance, and History of Play", *Games and Economic Behavior*, **42**, 281–299.
- STAHL, D. O. and WILSON, P. (1994), "Experimental Evidence of Players' Models of Other Players", *Journal of Economic Behavior and Organization*, **25**, 309–327.
- VAN HUYCK, J. B., BATTALIO, R. C. and BEIL, R. O. (1990), "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure", *American Economic Review*, **80**, 234–248.
- VAN HUYCK, J. B., BATTALIO, R. C. and BEIL, R. O. (1991), "Strategic Uncertainty, Equilibrium Selection, and Coordination Failure in Average Opinion Games", *Quarterly Journal of Economics*, **106**, 885–910.
- ZUCKERMAN, M. (1994), *Behavioural Expressions and Biosocial Bases of Sensation Seeking* (Cambridge: Cambridge University Press).