Signaling without a common prior: Results on experimental equilibrium selection ☆

Michalis Drouvelis [a], Wieland Müller [b,c], Alex Possajennikov [d,*]

[a] Department of Economics, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom
[b] Department of Economics, University of Vienna, Brünnerstrasse 72, 1210 Vienna, Austria
[c] CentER, TILEC, Tilburg University, Netherlands
[d] School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

A R T I C L E   I N F O

A B S T R A C T

The common prior assumption is pervasive in game-theoretic models with incomplete information. This paper investigates experimentally the importance of inducing a correct common prior in a two-person signaling game. Equilibrium selection arguments predict that different equilibria may be selected depending on whether the common prior is induced or not. Indeed, for a specific probability distribution of the sender's type, the long-run behavior without an induced common prior is shown to be different from the long-run behavior when a common prior is induced, while for other distributions long-run behavior is similar under both regimes. We also present a learning model that allows players to learn about the other players' strategies and the prior distribution of the sender's type. We show that this learning model accurately accounts for all main features of the data.

1. Introduction

This paper studies the effect of not inducing a correct common prior about the probability distribution of players' types on long-run outcomes of play. This is done using a specific example of a two-person signaling game that has two strict equilibria. We find that for some prior, average play indeed converges to a different outcome when the correct common prior belief is induced than when a common prior is not induced.

Game theory usually assumes that all components of a game such as the number of players, strategy sets, and payoffs are commonly known. Arguably, the assumption of commonly known game components is an exception rather than the rule in many real-world applications of game theory. If players are uncertain about any specific component of a game, Harsanyi (1967) suggested that players might agree on the set of possible realizations of this component and that a chance move at

the beginning of the game assigns an element out of this set of possible components to each player. But it is still usually assumed that the probability distribution of this chance move is common knowledge among players. However, what happens if players either do not agree on, or have no information about the probability distribution of the chance move? In this paper we concentrate on this question, that is, on the case where the common prior is initially not induced among players.

The common prior assumption clearly plays an important role in game theory in general and in information economics in particular. Although it is possible to build models without this assumption, it would be difficult, if not impossible, to solve them.[1] Almost all applications we are aware of make this assumption. However, it is not clear whether the assumption is actually satisfied in the situations that these applications are trying to describe. For instance, in labor market relationships it is hard to imagine that managers are aware of the (exact) probability distribution of skilled and unskilled job candidates, and it seems likely that their employment decisions are often based on previous learning experience.[2] This previous experience is one possible argument for the common prior assumption. We show that this argument needs to be qualified as this experience may differ depending on how play starts, which in turn depends on what is known about the game at the beginning of a series of interactions.

In this paper we experimentally investigate what happens if a common prior belief about the underlying probability distribution of player types is not induced at the beginning but can be learned over time. For this purpose we use a two-person signaling game that has two separating equilibria.[3] We choose a signaling setup because it allows for a simple situation that reflects the potential importance of having (or learning) correct beliefs. The sequential nature of the signaling game reduces strategic uncertainty about actions and thus allows us to focus on the consequences of what players know (or learn) about types of other players, rather than about their strategies. Signaling applications are common in economic modeling, making it potentially important to extend the analysis to situations where a common prior belief is not established at the start of repeated and random encounters. Various signaling games have been subjected to earlier experimental tests (e.g. Brandts and Holt, 1992, 1993; Banks et al., 1994; Cooper et al., 1997). Thus, there is already a certain amount of knowledge about players' behavior in such situations that can be built upon.

Since we are interested in the long-run outcomes that arise after the players have gained experience with playing the game and possibly learned its probabilistic structure, we use equilibrium selection arguments to predict whether initially knowing or not knowing the distribution of the chance move makes a difference for these long-run outcomes. In our game, selection is between two strict equilibria that both satisfy signaling game refinements (such as the intuitive criterion), in contrast to the previous experimental studies which analyzed an important but different question of selection of more refined equilibria in such games.

Equilibrium selection among strict equilibria is an old and persistent problem in game theory. The literature distinguishes between deductive and inductive equilibrium selection principles (for a recent comparison, see e.g. Haruvy and Stahl, 2004). Deductive equilibrium selection principles build on the idea that players coordinate on a specific equilibrium through a process of reasoning. Examples of such principles are payoff dominance and risk dominance (Harsanyi and Selten, 1988). These principles have sometimes been shown to be rather poor predictors of long-run outcomes in experimental games (Cooper et al., 1990; Van Huyck et al., 1990, 1991; see Straub, 1995, for a comparison of the principles). Inductive equilibrium selection principles are based on adaptive dynamics where players coordinate on a specific equilibrium through repeated interaction and learning.[4] Inductive selection dynamics have been reported to outperform deductive selection principles (e.g. Haruvy and Stahl, 2004).

As inductive equilibrium selection models are dynamical systems, their long-run outcomes not only depend on the way behavior adapts from one period to the next, but also strongly on initial conditions. Brandts and Holt (1992, 1993), for example, demonstrate that payoff changes in signaling games that leave equilibria unchanged may trigger different adjustment patterns, leading play to converge to a less refined equilibrium. In the context of learning models, initial conditions usually refer to initial beliefs held by players and the initial population shares playing the available strategies at time zero.

However, initial conditions may also be influenced by what is *initially known* about components of a game. In the context of inductive equilibrium selection, there are various results available regarding outcomes in games with "little" information.[5] We contribute to this literature by showing that suppressing information about a move by nature may sufficiently change

---

[1] See Dekel et al. (2004) for a thorough exposition of the issues that arise when players not only learn about opponents' strategies but also about nature's moves in Bayesian games.

[2] The same holds for the case of auctions: information about how the valuations of other bidders are distributed will often be unknown or is likely to be imprecise. Also, financial investors' information about the distribution of returns, states of nature or the distribution of future prices is often likely to be very limited.

[3] Note that although prior beliefs do not matter in a separating equilibrium, they matter along the adjustment path in a repeated play of the game. It is the latter that is analyzed in this paper.

[4] For applications of adaptive dynamics to equilibrium selection see, for example, Van Huyck et al. (1994, 1997). For an overview of learning models see e.g. Fudenberg and Levine (1998).

[5] Haruvy and Ünver (2007) find that a lower degree of knowledge about other market participants' preferences slows down convergence and often precludes coordination on more efficient equilibria in matching markets. Van Huyck et al. (2007) analyze the evolution of play in a coordination game when players know only their feasible choices and experienced payoffs. They show that "subjects converge to an absorbing state [...] faster than reinforcement learning algorithms, but slower than under complete information" (p. 53). Oechssler and Schipper (2003) analyze whether subjects learn the payoff structure in simple matrix games. The authors find that the games subjects appear to be playing (called "subjective" games) often differ from the real games, although play is close to an equilibrium of the subjective game.

players' initial beliefs to lead play to different outcomes, even after many opportunities for learning. More precisely, the main question the paper tries to answer is whether play converges to a different equilibrium when the common prior is induced than when it is not induced. To answer this question, we report the results of six experimental treatments based on a signaling game for which we vary both the prior distribution of the sender's type and whether this prior distribution is commonly known or not.

In our game the sender has two possible types, and the probability p of the sender being of type 1 is either 1/4, 1/2, or 3/4. We conjectured that knowing or not knowing the prior distribution (characterized by p) may change the long-run outcome of this game for one value of p, whereas the outcome would not be changed for the other values. This conjecture is based on a "shortcut" version of the "naive" learning model suggested by Brandts and Holt (1993, 1996) where players learn the opponent's strategy over time. Indeed, when the prior is known, this learning model predicts that play converges to one equilibrium of our signaling game if p is small (1/4) but converges to the other equilibrium if p takes on higher values (1/2 or 3/4). If the prior is not known, it is reasonable to assume that subjects start with a uniform prior about the sender's type.[6] In this case we demonstrate that, in an adjusted learning model (that also allows learning about the probability of the sender's type), under some circumstances play converges to the equilibrium predicted if the prior p is equal to 1/2 and known, independently of the true underlying prior. Hence, knowing or not knowing the prior is predicted to lead to different long-run outcomes only if $p = 1/4$.

Our experimental results support this hypothesis. Indeed, we find that in treatments with the prior $p = 1/2$ or $p = 3/4$, behavior appears to be the same independently of whether or not the prior is known, with play clearly converging to the equilibrium predicted by the learning process. But if the prior of $p = 1/4$ is not known, we find that although play does not converge to a pure equilibrium (as predicted by the naive learning model), there is a clear difference in comparison to the case of a known prior of $p = 1/4$. We thus clearly demonstrate the effect of varying initial information about the prior on long-run outcomes in an experimental game.

We further estimate the learning model parameters, which include noise in best response and inertia in belief revision. The estimated model tracks the data across all treatments reasonably well, and thus is quite parsimonious. Using one set of estimated model parameters, we show that the learning model predicts no differences in behavior when the prior is $p = 1/2$ or $p = 3/4$, independently of whether the prior is commonly known or not, but predicts clear differences when the prior is $p = 1/4$. The model thus provides further evidence that such a belief adjustment is a plausible way to describe how subjects play the game (see e.g. Brandts and Holt 1993, 1996, for signaling games and Haruvy and Stahl, 2004, for coordination games, for similar results). An interesting feature of the model is that beliefs about the underlying type distribution are revised more slowly than beliefs about the opponents' strategies. This is similar to the psychological evidence of forming more confident beliefs about a person's performance than about an "objective" uncertainty based on a small sample (e.g. Nickerson, 2004, Ch. 8).

Earlier experimental work by Güth and Ivanova-Stenzel (2003) has shown that knowing or not knowing the distribution of types in specific asymmetric auctions does not lead to significant differences in behavior. While this result is remarkable, in their study these authors confine themselves to only reporting the aforementioned result. In our study we go beyond the work presented by these authors. First, we use a simple shortcut version of a naive learning model to derive hypotheses about when knowing or not knowing the prior makes a difference in terms of which equilibrium play converges to in a signaling game. Second, we present experimental evidence confirming the hypotheses about differences in adjustment patterns for different values of the prior, thereby providing an example in which not knowing the prior distribution leads to differences in long-run behavior. Third, we propose an adaptation of the naive learning model of Brandts and Holt (1993, 1996) which allows players to also learn about the prior distribution in case it is not initially induced. More importantly, we show that this learning model tracks the evolution of outcomes in the various treatments quite accurately. Fourth, we find that subjects' behavior is consistent with an initially uniform prior about the sender's type if the prior is not induced, and that learning about strategies is faster than learning about the prior. We also provide further support for the model by checking its robustness to varying initial conditions regarding subjects' beliefs.

2. The signaling game

Our experiment is based on the following signaling game:

		Receiver					Receiver	
	Type t_1	a_1	a_2			Type t_2	a_1	a_2
Sender	m_1	15, 10	80, 80		Sender	m_1	80, 80	15, 30
	m_2	25, 10	50, 50			m_2	50, 50	25, 30

Nature (N) first selects the type of the sender and the sender observes his type: with probability p the type is t_1 and with probability $1 - p$ the type is t_2. Each type of the sender can then send either message m_1 or message m_2. The receiver
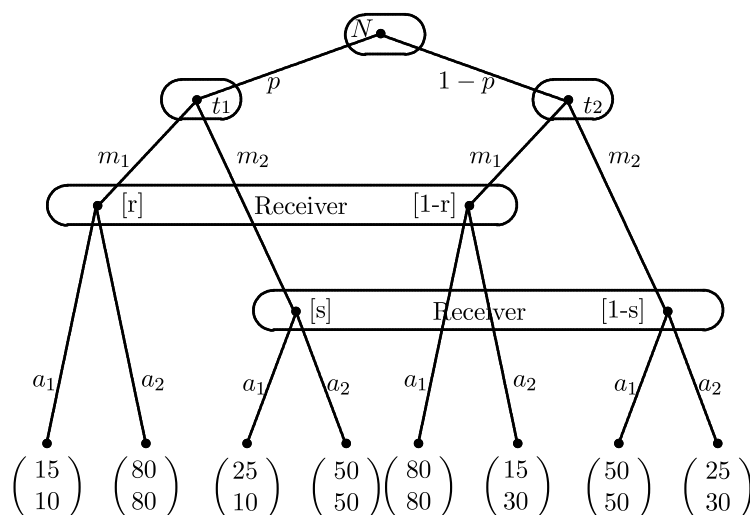

Fig. 1. The signaling game used in the experiment.

observes the message but not the type of the sender, and decides between actions a_1 and a_2. The extensive form of the game is shown in Fig. 1. The notation for the beliefs of the receiver is indicated in brackets [.] in her information sets: the receiver's posterior belief that the sender is of type t_1 is r after observing message m_1 and s after observing m_2. The payoffs of the sender (receiver) at terminal nodes are the upper (lower) entries in the payoff vectors.

In the experiment we implemented the priors $p \in \{1/4, 1/2, 3/4\}$. If p is commonly known, weak perfect Bayesian equilibria have the following properties. First, for any value of p the game has two pure separating equilibria: $[(m_1, m_2), (a_2, a_1), r = 1, s = 0]$ and $[(m_2, m_1), (a_1, a_2), r = 0, s = 1]$.[7] Second, the game has no pooling equilibria. Third, for each of the values of p mentioned above, the game has a hybrid equilibrium.[8] Because of the multiple pure strategy equilibria that are more efficient than the hybrid one, the game has a coordination structure.

If p is not commonly known, Bayesian equilibrium analysis is not possible.[9] We look at a dynamic learning model instead, which is also useful if p is known, since static equilibrium concepts do not say where beliefs come from. Brandts and Holt (1993, 1996) suggest that learning in a repeated signaling game like the one in Fig. 1 could be based on a naive reasoning process. We describe a version of such a learning process that combines "naive" initial beliefs with best reply behavior. The signaling game in Fig. 1 was constructed in such a way that for a small probability p of type 1, the learning process *could* lead to a different outcome depending on whether the prior is commonly known or not; and that the outcome of the learning process *could* be the same independently of the knowledge of the prior for larger values of p.

Suppose the sender initially believes that the receiver will choose any of her two actions with equal probability, independently of the message sent. Then both types of the sender will initially send message m_1 since, given the initial belief about the receiver's reaction, the expected payoff of sending m_1 (47.5) is higher than the expected payoff of sending m_2 (37.5) for both types of the sender.

If the prior is $p = 1/4$ and common knowledge, the best response of the receiver after observing either of the two messages is to choose action a_1. (The receiver's expected payoff after observing m_1 is 62.5 for a_1 and 42.5 for a_2, and the receiver's expected payoff after observing m_2 is 40 for a_1 and 35 for a_2.) So, initial play would be $[(m_1, m_1), (a_1, a_1)]$. Given the receiver's initial reaction, the best response for type t_1 of the sender would be m_2 while the best response for type t_2 remains unchanged. Given this strategy (m_2, m_1) of the sender, the receiver's best response will still be a_1 after observing message m_1, but a_2 after observing message m_2. These two strategies together—(m_2, m_1) for the sender and (a_1, a_2) for the receiver—constitute equilibrium behavior.

If $p = 1/2$, the best response of the receiver against the initial strategy of the sender mentioned above is action a_2 after m_1 (expected payoffs are 45 for a_1 and 55 for a_2) and a_2 after m_2 (expected payoffs are 30 for a_1 and 40 for a_2). So, in this case initial play would be $[(m_1, m_1), (a_2, a_2)]$. Given the receiver's initial reaction, the best response for type t_1 of the sender remains unchanged while the best response of type t_2 of the sender would be m_2 (change in payoff from 15 to 25). Given this strategy (m_1, m_2) of the sender, the receiver's best response will be a_2 after observing message m_1, and a_1 after observing message m_2. The two strategies together—(m_1, m_2) for the sender and (a_2, a_1) for the receiver—constitute, again,

---

[7] Here, the first entry in the sender's strategy is the message chosen as type t_1, and the second entry is the message chosen as t_2. Analogously, the first entry in the receiver's strategy is the action chosen after observing message m_1, and the second entry is the action chosen after message m_2.

[8] The game with $p = 1/4$ has the hybrid equilibrium $[(m_1, \frac{7}{15}m_1 + \frac{8}{15}m_2), (\frac{7}{13}a_1 + \frac{6}{13}a_2, a_1), r = \frac{5}{12}, s = 0]$, the game with $p = 1/2$ has the hybrid equilibrium $[(\frac{5}{7}m_1 + \frac{2}{7}m_2, m_1), (\frac{6}{13}a_1 + \frac{7}{13}a_2, a_2), r = \frac{5}{12}, s = 1]$, and the game with $p = 3/4$ has the hybrid equilibrium $[(\frac{5}{21}m_1 + \frac{16}{21}m_2, m_1), (\frac{6}{13}a_1 + \frac{7}{13}a_2, a_2), r = \frac{5}{9}, s = 1]$.

[9] However, in our experiment players receive feedback on realizations of types such that, over time, the probability distribution of Nature's move could be learned. The Bayesian equilibrium predictions for the appropriate value of p could then become relevant.

M. Drouvelis et al. / Games and Economic Behavior 74 (2012) 102–119

Table 1
Predictions regarding long-run behavior based on the naive learning process.

	Prior commonly known		Prior not commonly known	
	Sender	Receiver	Sender	Receiver
$p = 1/4$	(m_2, m_1)	(a_1, a_2)	(m_1, m_2)	(a_2, a_1)
$p = 1/2$	(m_1, m_2)	(a_2, a_1)	(m_1, m_2)	(a_2, a_1)
$p = 3/4$	(m_1, m_2)	(a_2, a_1)	(m_1, m_2)	(a_2, a_1)

Note: The first entry in the sender's strategy is the message chosen as type t_1 and the second entry is the message chosen as t_2. The first entry in the receiver's strategy is the action chosen after observing message m_1 and the second entry is the action chosen after m_2.

equilibrium behavior. If the commonly known prior is $p = 3/4$, the dynamic is the same as in the case of $p = 1/2$. We summarize these predictions in the second and third columns of Table 1.

If the prior probability p is not commonly known, it is perhaps reasonable to assume that subjective beliefs may be (close to) $p = 1/2$.[10] In that case, initial play should be as described above when $p = 1/2$. Further adjustment depends on the speed of belief revision about the prior[11] and about the opponent's strategy.[12] If players revise beliefs about p faster than beliefs about the opponent's strategy, the receiver may switch to (a_1, a_1) and play may be attracted to the equilibrium $[(m_2, m_1), (a_1, a_2)]$. On the other hand, if beliefs about the opponent's strategy are revised faster than beliefs about the probability p, players may continue playing as if $p = 1/2$ for long enough to get attracted to the equilibrium $[(m_1, m_2), (a_2, a_1)]$. Hence, it is possible that play in the above game will converge to the equilibrium $[(m_1, m_2), (a_2, a_1)]$ when the probability p is not commonly known. This is indicated in the last two columns in Table 1. Together with the predictions when p is commonly known, this implies that we expect long-run behavior to be quite similar when $p = 1/2$ or $p = 3/4$, independently of whether the probability is commonly known or not. On the other hand, we expect that long-run behavior will be different depending on the prior probability being commonly known or not, when $p = 1/4$.

Note that the basic mechanism of the short-cut version above is that some initial beliefs are fixed and that agents then alternate in choosing best responses. The naive learning model just described starts with the assumption that the sender initially believes that the receiver will choose any of her two actions randomly and independently of the message sent, and that the sender best-responds to this belief. That is, the sender thinks the receiver is a level-0 player, while the sender is a level-1 player in the depth-of-reasoning model of Stahl and Wilson (1994). If we started with the assumption that the sender is a level-0 player while the receiver is a level-1 player, then we would reach the same conclusions as shown in Table 1. This is also true for other assumptions about initial beliefs. For instance, if the receiver believes that the sender plays his maximin strategy (followed by alternating best responses), then the adjustment dynamics would again lead to the conclusions spelled out in Table 1.[13] In Section 5.2 we discuss the implications of various other alternative initializations of the learning model.

We want to emphasize that we view the above analysis as a "shortcut" version of a more elaborate learning model that we introduce in Section 5.[14] Of course, we do not expect subjects to converge to one of the equilibria within the first few periods. After all, subjects do not have to indicate an entire strategy, but only react at one information set at a time. Also, so far we assumed that subjects always choose a best reply given their beliefs, without making any decision errors. This assumption is relaxed in Section 5. Here, we only wanted to show that a reasonable adjustment theory gives rise to the hypothesis that for some values of p it might make a difference whether or not players know the prior.

Although it is possible to consider different adjustment processes, we concentrate on belief revision and myopic best response dynamic (with noise, in Section 5). We view this kind of adjustment as natural in an experimental setting with binary decisions, and variants and extensions of it have successfully been used previously in experimental signaling games (see, e.g., Brandts and Holt 1992, 1993, 1996; Cooper et al., 1997; Anderson and Camerer, 2000) and coordination games (e.g. Haruvy and Stahl, 2004). The process is rich enough to generate various predictions depending on initial conditions, which can then be tested.

---

[10] Indeed, it may be even more reasonable to assume a flat prior belief about Nature's move than about the opponent's move as the "principle of insufficient reason" (see footnote 6) seems to be more applicable to exogenous events.

[11] Similarly to learning about strategies, with respect to learning about the unknown prior we assume that, starting with the flat prior of $p = 1/2$, players update this prior in each round given the feedback about the selected type.

[12] In Section 5 we estimate parameters determining these speeds. There, we also consider variants in which one type of learning is much faster than the other, to see the importance of relative speeds. Nickerson (2004, Ch. 8) cites psychology research showing that beliefs about "objective" uncertainty takes a long time to form (slow revision), while a small sample may be sufficient to form beliefs about a person's performance (fast revision).

[13] Suppose the receiver initially believes that each type of the sender chooses his maximin message, that is, each type of the sender chooses message m_2. If $p = 1/4$ is known, the best response of the receiver is a_1 after each message. Then the sender sends m_2 as type t_1 and m_1 as type t_2. The receiver then plays a_1 after m_1 and a_2 after m_2, and, hence, the equilibrium $[(m_2, m_1), (a_1, a_2)]$ is reached. If $p = 1/2$ or $p = 3/4$ is known or if p is unknown, the best response of the receiver is a_2 after each message. Then the dynamic converges to the equilibrium $[(m_1, m_2), (a_2, a_1)]$. Taken together, we have the same predictions as in Table 1. However, should the sender believe that the receiver plays her maximin strategy (choosing action a_2 after each message), then one round of alternating best responses would lead to the equilibrium $[(m_1, m_2), (a_2, a_1)]$, independently of the prior.

[14] In this formal learning model we will spell out explicitly how the updating of beliefs about strategies and the prior works and what we mean when we refer, as above, to the speed of belief revision.

The game with $p = 1/2$ is included in our experiment because this value is a natural first guess when p is not known. We do not expect that the behavior will depend on whether $p = 1/2$ is commonly known or not; treatments with $p = 1/2$ rather serve as a benchmark. We use other values of p on both sides of $1/2$ for completeness. If play in the game in Fig. 1 follows the "naive" dynamic described above, the game with value $p = 1/4$ can be an example where knowing the prior distribution matters in the long run, while the game with value $p = 3/4$ is another example where it does not matter.

The payoffs in the game were primarily chosen in such a way that for some values of p the learning process could lead to a different outcome depending on whether or not p is commonly known, and that for other values of p the outcome is the same independently of the knowledge of the prior. We chose a game with two separating equilibria because such equilibria would remain for any value of p. A strategy profile that is a pooling equilibrium for one value of p may cease to be an equilibrium for other values of p. The difference in observed behavior across games with different values of p can then be influenced by the different structure of the equilibrium set, in addition to our main treatment variable (knowledge of the value of p). Clearly, we wanted to avoid this.

In an effort to weaken their possible confounding influence on adjustment patterns, we also kept an eye on various behavioral effects when designing the game. These were the similarity of payoffs for the two players and for the two sender types (to avoid relative payoff effects), the similarity of payoffs in the two long-run outcomes (to lessen a possible focal effect of the efficient outcome), a small conflict between expected payoff maximization and maximin (to diminish risk aversion and ambiguity aversion effects), and the relative steepness of the best response (to give enough incentive to find it). Further desirable features we considered were that all messages are sent during the adjustment period (to give subjects experience in all information sets of the game), and that the dynamic is relatively fast (so that the convergence to a long-run outcome can occur in a reasonable period of time).

3. Experimental design and hypotheses

3.1. Experimental design

Our experiments are based on a 3×2 factorial design. It includes treatments with three different values of the prior, namely $p \in \{1/4, 1/2, 3/4\}$ and it includes treatments with and without a commonly known prior. A treatment is denoted K-p if the prior probability $p \in \{1/4, 1/2, 3/4\}$ of type 1 of the sender is commonly known and N-p if the prior is not known. Hence, the six treatments are referred to as K-1/4, K-1/2, K-3/4, N-1/4, N-1/2, and N-3/4.

In the instructions, senders were referred to as "A-participants" and receivers as "B-participants". Subjects were informed that A-participants could be of two types ("Type 1" and "Type 2") and that at the beginning of each round, a random draw would determine the type of an A-participant. For the treatments with unknown prior, the instructions contained the following sentences: "The random draw is such that with an $X\%$ chance the A-participant will be of Type 1, and with a $(100 - X)\%$ chance of Type 2. You receive no information about the value of X, except that X is constant over all rounds of the experiment." In the treatments with a commonly known prior, the latter sentence was omitted and X and $100 - X$ were explicitly given by either 75 and 25, 50 and 50, or 25 and 75. In all treatments subjects were informed that, after the random draw, the A-participant would be informed about his type while the B-participant would not be informed about the type of the A-participant.

The sender's messages and the receiver's actions were generally referred to as "decisions" and labeled "C" and "D" (for messages) and "E" and "F" (for actions). Each session consisted of 40 rounds to give subjects room for learning. Payoffs were given to senders in the form of two tables corresponding to the two types. Receivers saw payoff tables corresponding to different messages as this makes it clearer what they know and what they do not know.[15]

Given that our research topic concerns beliefs, we could have asked questions with regard to which beliefs players actually entertain during the game and whether the true distribution of types is eventually learned. However, in this study we chose to limit ourselves to the analysis of behavior. The game appeared to us to be relatively complex, and proper elicitation of beliefs may have created excessive demands on the cognitive effort of subjects, perhaps making behavior unnecessarily noisy.

The experiment was computerized using the *z-Tree* software (Fischbacher, 2007) and was conducted in March 2007 at the Centre for Decision Research and Experimental Economics (CeDEx) laboratory at the School of Economics of the University of Nottingham, United Kingdom. The recruitment of the subjects from the pool of University of Nottingham students registered with CeDEx for experiments was done using the *ORSEE* software (Greiner, 2004). Subjects were students from various fields of studies including economics.

There were three sessions each for treatments K-1/4 and N-1/4 (since this is the comparison we were mostly interested in), and two sessions for each of the other four treatments. Each session consisted of 16 subjects who were divided into two matching groups of eight subjects each (four senders and four receivers). The design is summarized in Table 2.

We randomly rematched subjects in every period within each matching group in order to create an environment that most resembles a single-period interaction between subjects. The same matching protocol was used in all individual matching groups. Player roles were randomly assigned at the beginning of the experiment and were then kept fixed for the entire

---

[15] See Appendix A for further details of the instructions.

Table 2
Overview of the experimental design.

Treatment	Value of the prior	Prior	No. of sessions	No. of independent matching groups	No. of subjects
K-1/4	1/4	known	3	6	$6 \times 8 = 48$
N-1/4	1/4	not known	3	6	$6 \times 8 = 48$
K-1/2	1/2	known	2	4	$4 \times 8 = 32$
N-1/2	1/2	not known	2	4	$4 \times 8 = 32$
K-3/4	3/4	known	2	4	$4 \times 8 = 32$
N-3/4	3/4	not known	2	4	$4 \times 8 = 32$

experiment. The assignment of types to senders in a matching group was random in each of the 40 rounds of the experiment but was done prior to the experiments for each value of p. This assignment of types to senders was then used for all matching groups in treatments with this value of p.

After each period, each subject observed the type of the sender, the message sent by the sender, the action taken by the receiver, and the realized payoffs in his/her own match. This is needed to update beliefs. Some parts of this information can be inferred from other parts of the feedback, but we decided to give all of this information explicitly to facilitate learning. We decided not to provide the whole history of play up to the current period because our predictions are based on a short-memory best-reply adjustment process. Subjects were allowed to take notes (though almost nobody did).

Sessions lasted approximately 90 minutes. Subjects were paid according to accumulated earnings over all rounds. With the points converted to money at the rate of £0.05 for 10 points, the average payments were £11.73 ($22.56) per subject.

3.2. Hypotheses

We want to compare behavior when the prior probability is unknown with behavior when this probability is commonly known. Based on the simple learning process described above, we expect long-run behavior to be quite similar when the prior probability is $p = 1/2$ or $p = 3/4$, independently of whether or not the value of p is commonly known. On the other hand, when the probability is $p = 1/4$ we expect the long-run behavior to be different depending on whether the value of p is commonly known or not (see Table 1). Based on the "naive" reasoning assumed in the learning process, we should also see short-run behavior in treatments with an unknown value of p to be similar to behavior when $p = 1/2$ is known.

4. Results

4.1. A first look

We first look at the time series of senders' and receivers' strategies, averaged across sessions and subjects of the same treatment. Figs. 2, 3, and 4 show data for the treatments with prior $p = 1/4$, $p = 1/2$, and $p = 3/4$. The data are grouped in blocks of 5 periods. The figures show the relative frequency of decisions made. More precisely, the two top panels in each figure show the relative frequency of m_1 choices for each of the two types of the sender, whereas the two bottom panels in each figure show the relative frequency of a_1 choices after each of the two possible messages.

Consider first Fig. 2, which shows behavior in the treatments with prior $p = 1/4$. Concentrate first on behavior in treatment K-1/4 where the prior is known. In this case we see clear convergence towards the equilibrium $[(m_2, m_1), (a_1, a_2)]$ as predicted in Table 1. That is, senders of type 2 learn to send message m_1 while receivers learn to respond to it with action a_1. Senders of type 1 learn to send m_2, after which receivers react mostly with a_2. Next, we observe that there is a clear gap between the graphs representing behavior in treatments K-1/4 and N-1/4. Furthermore, the graphs representing behavior in the two treatments by and large run in parallel (with the exception of the early behavior of receivers after observing m_2). Finally, while behavior in treatment K-1/4 converges to a pure equilibrium, this does not seem to be the case in treatment N-1/4.[16]

The average frequency in treatment N-1/4 masks some heterogeneity across matching groups. One group seems to converge to the $[(m_2, m_1), (a_1, a_2)]$ equilibrium, another one is close to the $[(m_1, m_2), (a_2, a_1)]$ equilibrium, while the other matching groups are in between, not close to any of the two pure equilibria or the hybrid one.[17] By contrast, in the K-1/4 treatment all matching groups exhibit the same pattern of choices, close to $[(m_2, m_1), (a_1, a_2)]$. Thus, even though the behavior in treatment N-1/4 does not converge to an equilibrium, it is clear that there is a difference in strategy choices between the two treatments.

---

[16] Recall from footnote 8 that the game with $p = 1/4$ has the hybrid equilibrium $[(m_1, \frac{7}{15}m_1 + \frac{8}{15}m_2), (\frac{7}{13}a_1 + \frac{6}{13}a_2, a_1), r = \frac{5}{12}, s = 0]$ and the game with $p = 1/2$ has equilibrium $[(\frac{5}{7}m_1 + \frac{2}{7}m_2, m_1), (\frac{6}{13}a_1 + \frac{7}{13}a_2, a_2), r = \frac{5}{12}, s = 1]$. In these equilibria, either type 1 or type 2 of the sender chooses one of the messages with certainty and the receiver chooses one of the actions with certainty after observing m_2. Fig. 2 clearly suggests that this is not what play converges to over time in treatment N-1/4.

[17] The data for these observations are reported in Section 2.2 of the Supplementary material.
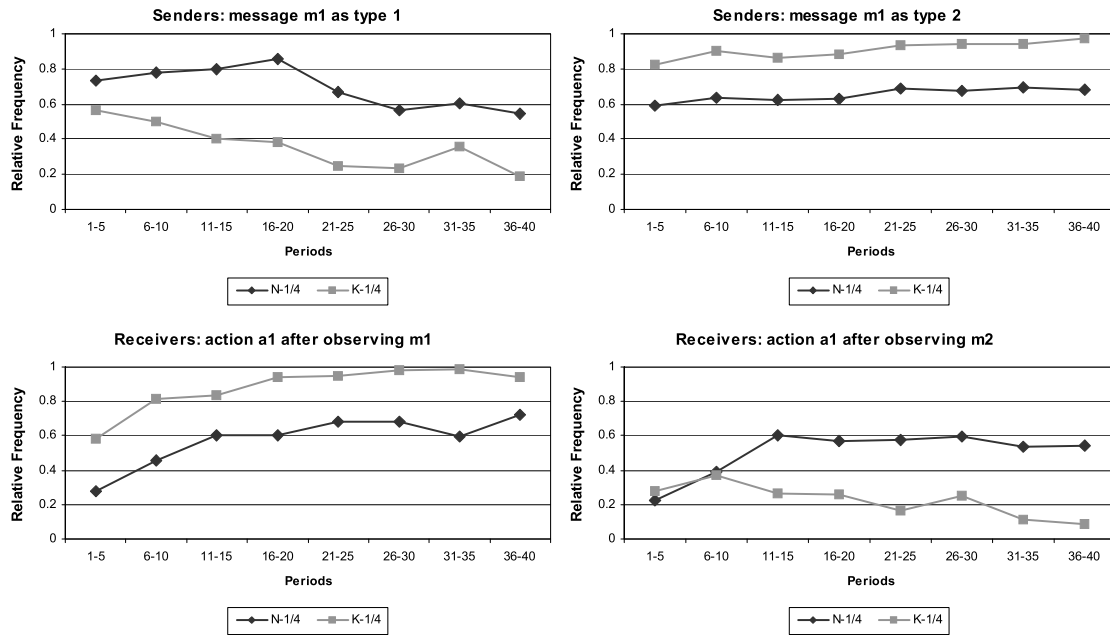
Fig. 2. Time series of average strategies in treatments *N*-1/4 and *K*-1/4.
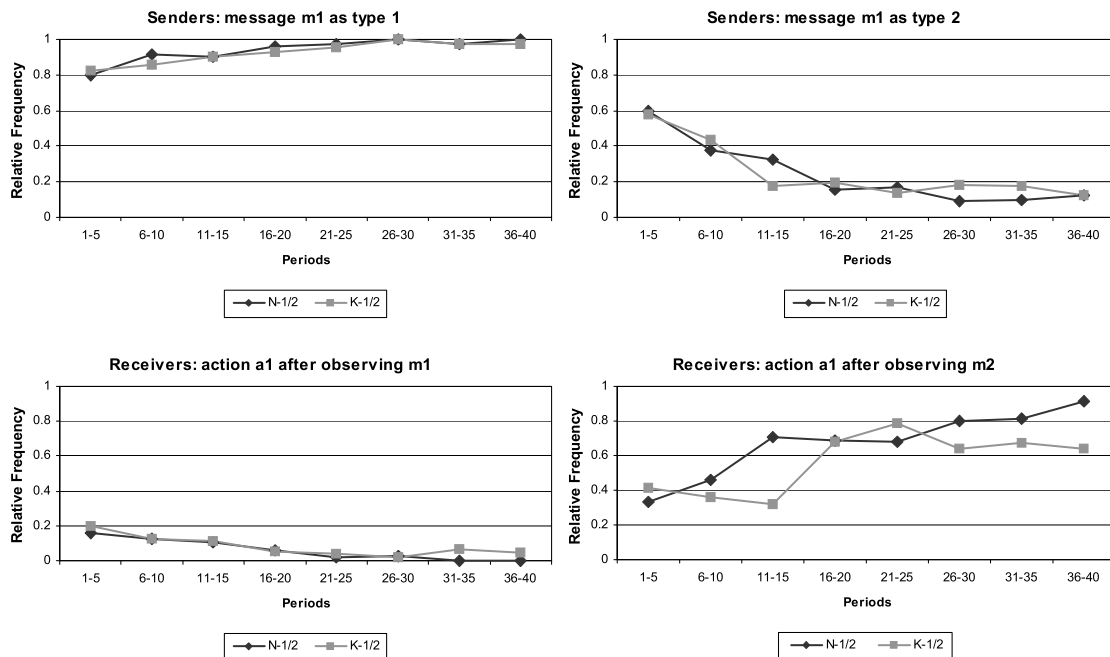

Fig. 3. Time series of strategies in treatments *N*-1/2 and *K*-1/2.

Let us now consider behavior in the treatments with prior $p = 1/2$ and $p = 3/4$ as shown in Figs. 3 and 4. The results in these two cases are very similar and straightforward. Inspecting the graphs in each of the four panels of these two figures reveals that there is not much difference between behavior in treatments *N-x* and *K-x* for $x \in \{1/2, 3/4\}$. Also comparing the panels between the two figures, there does not seem to be much difference across all four treatments. In particular, when the sender is of type 1, almost exclusively m_1 is played and this message is answered by action a_2. When the sender is of type 2, play converges to the senders choosing message m_2 which the receivers answer by mostly choosing a_1. This pattern is replicated across matching groups. Thus, we conclude that in all treatments with $p = 1/2$ or $p = 3/4$ play converges to the equilibrium $[(m_1, m_2), (a_2, a_1)]$ which is what we predicted in Section 2 (see Table 1).

Summarizing, Figs. 2–4 suggest the following. First, in treatments with prior $p = 1/2$ or $p = 3/4$ behavior appears to be the same, independently of whether the prior is known or not, with play clearly converging to the equilibrium predicted by the learning process illustrated in Section 2. Second, in the treatment with known prior $p = 1/4$, play clearly converges to the equilibrium predicted by the naive learning process. However, when the prior $p = 1/4$ is not known, play does not
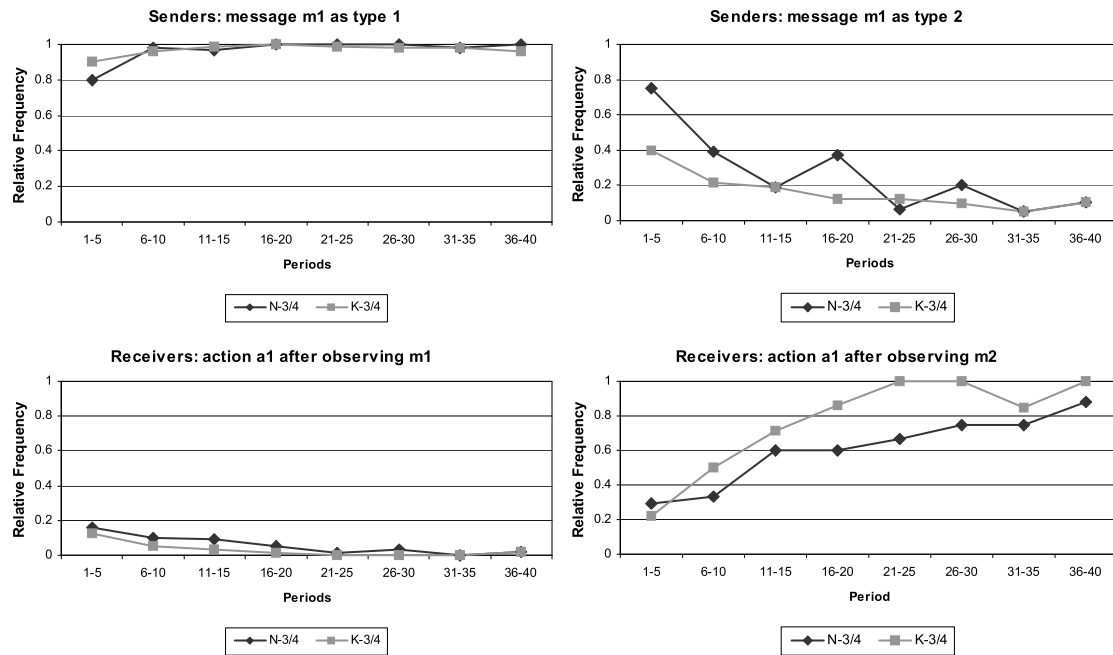
Fig. 4. Time series of strategies in treatments N-3/4 and K-3/4.

always converge to the same pure equilibrium and there is a clear difference in comparison to the case of a known prior $p = 1/4$. Hence, this first graphical inspection of the data supports our hypotheses stated at the end of Section 3.

Another observation, common to all treatments, is that the time series are fairly constant in the late periods, while there are discernible trends in the early ones. We will consider the relatively stable behavior in the second half of the sessions as the basis for the statistical tests for long-run behavior.

4.2. A formal analysis

Let us now check the observations from the graphs more formally. We first ask for each of the priors $p \in \{1/4, 1/2, 3/4\}$ separately, whether there is a difference in behavior if the prior is known or not known.

We use the Wilcoxon–Mann–Whitney non-parametric test (Siegel and Castellan, 1988, Ch. 6), using matching groups as independent observations. This gives six observations for the treatments N-1/4 and K-1/4 and four observations for each of the other treatments. We use Periods 21–40 for the main tests, checking the robustness of the results for other choices of periods.

To test whether, conditional on type realization, senders behave differently if the prior is known or not, we calculate the proportions of times senders choose message m_1 within each matching group. Similarly, to test whether, conditional on the message observed, receivers behave differently if the prior is known or not, we calculate the proportion of times receivers choose action a_1.[18]

The results are presented in Table 3. The tests presented in this table reject the one-sided null hypotheses of differences between the strategy proportions in treatments N-1/4 and K-1/4 at the 5% significance level. Therefore, there is a difference between the treatments in the direction predicted by the model in Section 2. On the other hand, the two-sided null hypotheses of no differences between treatments N-1/2 and K-1/2, and between treatments N-3/4 and K-3/4 are not rejected. The results are similar when tests are done on the data from all periods, or on the data from the last block (Periods 36–40) only.[19]

Since the power of the test is low if there are only four observations per treatment, we also ran subject-specific random-effect probit regressions to compare strategies across treatments. These regressions confirm the difference between treatments N-1/4 and K-1/4, and fail to detect significant differences between treatments with known and unknown prior when $p \neq 1/4$ (except in N-3/4 vs K-3/4 in the response of receivers to the less frequent message m_2). That the behavior in the second half of the experiment is stable is confirmed by the trend regressions, as time trends are not significant for almost all strategy choices in any of the treatments.[20]

Furthermore, we employed a regression technique introduced by Noussair et al. (1995) to analyze the asymptotic behavior of the strategy time series of senders and receivers. Using the actually observed data, this technique allows one

---

[18] The data these tests are based on are reported in Section 1 of the Supplementary material.

[19] The details of these tests are reported in Section 2 of the Supplementary material.

[20] The details of all these regressions are reported in Section 3 of the Supplementary material.

Table 3
Results of tests regarding differences in behavior across treatments in Periods 21–40.

Proportions of strategies for Periods 21–40 and comparison tests				
	Senders		Receivers	
	$m_1\|t_1$	$m_1\|t_2$	$a_1\|m_1$	$a_1\|m_2$
N-1/4 vs K-1/4	0.59 vs 0.26 (0.032)[**]	0.68 vs 0.95 (0.023)[**]	0.67 vs 0.96 (0.018)[**]	0.56 vs 0.15 (0.019)[**]
N-1/2 vs K-1/2	0.99 vs 0.98 (0.495)	0.12 vs 0.16 (0.237)	0.01 vs 0.04 (0.281)	0.80 vs 0.68 (0.149)
N-3/4 vs K-3/4	1.00 vs 0.98 (0.850)	0.11 vs 0.10 (0.653)	0.02 vs 0.00 (0.166)	0.78 vs 0.96 (0.237)

Note: p-Values in parentheses. N-1/4 vs K-1/4: 6 observations for each treatment, H_0: $Prop_{N-1/4} \leqslant Prop_{K-1/4}$ for $m_1|t_1$ and $a_1|m_2$; H_0: $Prop_{N-1/4} \geqslant Prop_{K-1/4}$ for $m_1|t_2$ and $a_1|m_1$. N-1/2 vs K-1/2 and N-3/4 vs K-3/4: 4 observations for each treatment, for all tests H_0: $Prop_{N-x} = Prop_{K-x}$, where $x = 1/2, 3/4$.
[**] $p < 0.05$.

Table 4
Results of tests regarding differences in behavior in Periods 1–5.

Proportions of strategies in Periods 1–5 and comparison tests				
	Senders		Receivers	
	$m_1\|t_1$	$m_1\|t_2$	$a_1\|m_1$	$a_1\|m_2$
N-1/4 vs K-1/2	0.73 vs 0.83 (0.590)	0.59 vs 0.58 (0.745)	0.28 vs 0.20 (0.398)	0.22 vs 0.42 (0.392)
N-1/2 vs K-1/2	0.80 vs 0.83 (0.505)	0.60 vs 0.58 (0.769)	0.16 vs 0.20 (0.885)	0.33 vs 0.42 (1.000)
N-3/4 vs K-1/2	0.80 vs 0.83 (0.772)	0.75 vs 0.58 (0.186)	0.16 vs 0.20 (0.773)	0.29 vs 0.42 (1.000)
N vs K-1/2	0.78 vs 0.83 (0.704)	0.61 vs 0.58 (0.451)	0.21 vs 0.20 (0.790)	0.27 vs 0.42 (0.668)

Note: p-Values in parentheses. N-1/4: 6 observations, other treatments: 4 observations; N: 14 observations pooled across N treatments. H_0: $Prop_{N-x} = Prop_{K-1/2}$, where $x = 1/4, 1/2, 3/4$ and H_0: $Prop_N = Prop_{K-1/2}$.

to conjecture which strategies play would have converged to, had the experiment been repeated for many more rounds (possibly for infinitely many rounds).[21] We find that, whereas the asymptotes for treatments K-1/4 and N-1/4 are significantly different from each other, they are not different from each other at any conventional significance level in the relevant pairwise comparisons for the treatments with $p \geqslant 1/2$. This confirms the findings from Section 4.1 (especially the visual impressions from Figs. 2 to 4).

Result 1. There is a significant difference in behavior between treatments with known and unknown prior when $p = 1/4$ in the directions predicted by the model in Section 2 (see Table 1). There are no significant differences in behavior between treatments with known and unknown prior when $p = 1/2$ or $p = 3/4$.

The predictions of our simple learning model in Section 2 (as well as the more elaborate version presented in the next section) are based on the assumption that subjects in treatments with an unknown prior entertain the initial belief that each of the sender's type is equally likely. A comparison of behavior between treatments N-p and K-1/2 in the early rounds (rounds 1–5) can show whether or not this assumption is justified. Note, though, that testing for possible differences in *behavior* only provides an indirect test of the hypothesis that the subjective prior belief maintained in the initial periods of the treatments with an unknown p is the same as the prior in treatment K-1/2. However, since we did not explicitly elicit beliefs about the prior, this indirect method is the only way to shed light on the issue of what subjects' initial beliefs are when p is not known.

The tests on the behavior in the first five periods of the experiments compare, separately for senders and receivers, proportions of actions used in the treatments with an unknown prior with those used in the treatment with known prior $p = 1/2$. The results of the Wilcoxon–Mann–Whitney tests are presented in Table 4. The test results show that we cannot reject the hypothesis that behavior in the early periods of the treatments with unknown prior is the same as behavior

---

[21] The details of these regressions are reported in Section 4 of the Supplementary material. Note that the technique employed here assumes that the strategies to which play converges are common to all matching groups of the same treatment. Because of some heterogeneity across matching groups, this assumption is violated in treatment N-1/4, and therefore the results for the comparison between N-1/4 and K-1/4 are not fully reliable.

in treatment K-1/2. This in turn implies that we cannot reject the hypothesis that subjects in the treatments with an unknown prior entertain beliefs about the prior which are similar to the ones in treatment K-1/2. The results are also corroborated by subject-specific random-effects probit regressions. These statistical tests also fail to detect differences when the N treatments are pooled (see the bottom row of Table 4).[22]

Result 2. There are no significant differences in the behavior in early periods between treatments with an unknown prior and the treatment with a known prior $p = 1/2$.

The simple version of the learning model described in Section 2 makes specific predictions about the behavior in Period 1 and its subsequent adjustment in Period 2. While the actual behavior in these periods is often consistent with the predictions,[23] the analysis of only the first two periods cannot provide more than an indication that the basic learning model captures qualitative trends. In the next section, we describe and estimate an extended learning model that is based on the naive model of Section 2. Importantly, it also takes into account further plausible features such as learning about the prior distribution, imperfect best response, and inertia in belief revision.

5. Simulations of the learning model

5.1. Baseline simulations

In this section we present the results of simulations which are based on an extension of a learning model first proposed in Brandts and Holt (1993, 1996). In this model, players learn about the strategies of the other players; the extension allows the players to also learn about the distribution of types in the case where it is not initially induced.

The main idea of the learning model is as follows. Players initially have prior beliefs about the other player's strategies and choose a noisy best response given these beliefs. In later rounds they update their beliefs in view of experience and continue choosing noisy myopic best responses. In treatments in which players do not know the probability of the chance move, they initially have a prior belief about it and update this belief in each round given the feedback about the selected type.

Let us first concentrate on receivers' beliefs about senders. Let $A_\tau^S(t_i, m_j)$ be the "belief propensity" that type t_i sends message m_j in period τ, and let $A_\tau^S(t_i) = A_\tau^S(t_i, m_1) + A_\tau^S(t_i, m_2)$. The belief at time τ that type t_i sends message m_j is $B_\tau^S(m_j|t_i) = A_\tau^S(t_i, m_j)/A_\tau^S(t_i)$. Let $A_1^S(t_i, m_j)$ be given and assume $A_1^S(t_1) = A_1^S(t_2)$.[24] Updating proceeds as follows: suppose that a receiver observes that type t_i of the sender chose message m_j at time τ. Then

$$A_{\tau+1}^S(t_i, m_j) = A_\tau^S(t_i, m_j) + 1,$$
$$A_{\tau+1}^S(t_i, m_{-j}) = A_\tau^S(t_i, m_{-j}),$$
$$A_{\tau+1}^S(t_{-i}, m_j) = A_\tau^S(t_{-i}, m_j) \quad \text{for } j = 1, 2.$$

That is, only the belief propensity of type t_i who sent message m_j is increased by 1, whereas all other propensities remain unchanged.

Next we turn to senders' beliefs about receivers. Let $A_\tau^R(m_j, a_k)$ be the belief propensity that message m_j is answered by action a_k in period τ, and let $A_\tau^R(m_j) = A_\tau^R(m_j, a_1) + A_\tau^R(m_j, a_2)$. The belief that message m_j is answered by action a_k at time τ is $B_\tau^R(a_k|m_j) = A_\tau^R(m_j, a_k)/A_\tau^R(m_j)$. Let $A_1^R(m_j, a_k)$ be given. We assume that $A_1^R(m_1) = A_1^R(m_2) = A_1^S(t_i)$.[25] Suppose a sender observes that message m_j is answered by action a_k at time τ. Then the propensities are updated as

$$A_{\tau+1}^R(m_j, a_k) = A_\tau^R(m_j, a_k) + 1,$$
$$A_{\tau+1}^R(m_j, a_{-k}) = A_\tau^R(m_j, a_{-k}),$$
$$A_{\tau+1}^R(m_{-j}, a_k) = A_\tau^R(m_{-j}, a_k) \quad \text{for } k = 1, 2.$$

Only the belief propensity of the action that was chosen in response to the actual message sent is increased by 1, whereas all other propensities remain unchanged.

To extend the model to learning about the chance move, let $A_\tau^T(t_i)$ be the belief propensity that type t_i is the outcome of the chance move at time τ, and let $A_\tau^T = A_\tau^T(t_1) + A_\tau^T(t_2)$. The belief that type t_i is selected in period τ is $B_\tau^T(t_i) = A_\tau^T(t_i)/A_\tau^T$. Let $A_1^T(t_i)$ be given. If the probability p of type 1 is known, then $A_\tau^T(t_1) = p \cdot A_\tau^T$ and $A_\tau^T(t_2) = (1-p) \cdot A_\tau^T$. This implies $B_\tau^T(t_1) = p$, and $A_\tau^T(t_i)$ are not updated further. If p is not known, then beliefs are updated via propensities. When

---

[22] The data these tests are based on and the test results are reported in Section 5 of the Supplementary material.

[23] We provide evidence of this in Section 6 of the Supplementary material.

[24] $A_1^S(t_i)$ corresponds to the parameter α in Brandts and Holt (1996).

[25] This means that updating will proceed with the same speed for senders and receivers.

a player observes that type t_i was realized at time τ, then

$$A^T_{\tau+1}(t_i) = A^T_\tau(t_i) + 1,$$
$$A^T_{\tau+1}(t_{-i}) = A^T_\tau(t_{-i}).$$

For the naive learning model of Section 2, we assume $A^S_1(t_i, m_j) = A^S_1(t_i)/2$, $A^R_1(m_j, a_k) = A^R_1(m_j)/2$, and, if the value of p is not known, $A^T_1(t_1) = A^T_1/2$. These assumptions imply the initial beliefs that players use each strategy with equal probability, and that each type is equally likely if p is not known. We begin with these assumptions and check the robustness of the results against alternatives.[26] More generally, different initial beliefs can be specified by setting $A^S_1(t_i, m_1) = p^S_i A^S_1(t_i)$, $A^R_1(m_j, a_1) = p^R_j A^R_1(m_j)$, $A^T_1(t_1) = p^T A^T_1$, where p^S_i, p^R_j, p^T represent initial beliefs about the sender's strategy, the receiver's strategy, and the type respectively (and can be chosen by the modeler).

Given the beliefs, players calculate expected payoffs $E^S[m_j|t_i] = S(t_i, m_j, a_1) \times B^R(a_1|m_j) + S(t_i, m_j, a_2) \times B^R(a_2|m_j)$ for senders and $E^R[a_k|m_j] = R(t_1, m_j, a_k) \times B^S(t_1|m_j) + R(t_2, m_j, a_k) \times B^S(t_2|m_j)$ for receivers, where $S(t_i, m_j, a_k)$ and $R(t_i, m_j, a_k)$ are the payoffs of the sender and of the receiver if the sender's type is t_i, the message sent is m_j, and the action chosen by the receiver is a_k. Receivers' posterior beliefs that the sender is of type t_i given that message m_j was sent are $B^S(t_i|m_j) = B^S(m_j|t_i)B^T(t_i)/(B^S(m_j|t_i)B^T(t_i) + B^S(m_j|t_{-i})B^T(t_{-i}))$.

We assume that players make (logistic) decision errors when choosing their actions. More precisely,

$$\Pr(m_j|t_i) = \frac{\exp(\lambda \cdot E^S[m_j|t_i])}{\exp(\lambda \cdot E^S[m_j|t_i]) + \exp(\lambda \cdot E^S[m_{-j}|t_i])},$$

$$\Pr(a_k|m_j) = \frac{\exp(\lambda \cdot E^R[a_k|m_j])}{\exp(\lambda \cdot E^R[a_k|m_j]) + \exp(\lambda \cdot E^R[a_{-k}|m_j])},$$

where $\Pr(m_j|t_i)$ is a sender's probability of playing message m_j when he is of type t_i and $\Pr(a_k|m_j)$ is a receiver's probability of playing action a_k after observing message m_j. The parameter λ represents the strength with which agents choose a best response.

Given the choice of the initial beliefs $A^S_1(t_i, m_j) = A^S_1(t_i)/2$, $A^R_1(m_j, a_k) = A^R_1(m_j)/2$, and $A^T_1(t_1) = A^T_1/2$ if p is not known, the parameters of our learning model are (i) the strength of the initial propensities on strategies $A^S_1(t_i) = A^R_1(m_j) = A_1$, (ii) the strength of the initial propensities on types A^T_1 (in treatments with an unknown prior), and (iii) the strength of the best response λ.

Mimicking the procedures used in our experiments, the simulation model was run over 40 periods with eight simulated agents (four agents in each of the two player roles), using the same matching protocol as in the actual experimental sessions. In the simulations, we used the same outcomes of the chance moves that selected senders' types in the experimental sessions. To smooth out the effects of randomness, we use the average outcomes of 100 runs of the simulation model for each treatment and compare the simulated outcomes with the actual data. In each period, there are eight proportions of observed actual outcomes $p_{obs}(m_j, a_k|t_i)$ and eight proportions of simulated outcomes $p_{sim}(m_j, a_k|t_i)$. As a goodness-of-fit measure, for each treatment we use the metric

$$g = \sum_{\tau=1}^{40} \big(p_{obs}(m_j, a_k|t_i) - p_{sim}(m_j, a_k|t_i)\big)^2.$$

We add up the metrics for all treatments and look for the set of parameters A_1, A^T_1, and λ that give the best fit to the observed data.

To find the best-fitting parameters, we did a grid search. After a few initial trials to locate the likely values of parameters, the grid search concentrated on the interval $[0.5, 5]$ with step size 0.5 for A_1, the interval $[11, 20]$ with step size 1 for A^T_1, and the interval $[0.05, 0.25]$ with step size 0.05 for λ. In Figs. 5–7 we present the results of simulations that use the set of parameters

$$A_1 = 2, \qquad A^T_1 = 15, \qquad \lambda = 0.15$$

that give the best fit over all six treatments. These figures also present the observed behavior and are organized in the same way as Figs. 2–4.

Let us first consider the results of the simulations for the treatments with $p = 1/4$, shown in Fig. 5. From our analysis in Section 4, the main feature of the observed data is that there is a clear gap between choices depending on whether or not the prior $p = 1/4$ is known. Fig. 5 shows that the learning model accounts for this observation as it predicts the same differences in choice behavior in all subpanels of Fig. 5. The learning model, though, underestimates the choice of message m_1 for type t_1 of the sender and overestimates it for type t_2 in both treatments N-1/4 and K-1/4 (the simulated senders of type t_1 learn much quicker not to choose message m_1 if the prior $p = 1/4$ is known, see upper left panel in Fig. 5). The

---

[26] These assumptions are also largely consistent with the data. See Result 2 and Section 6 of the Supplementary material.
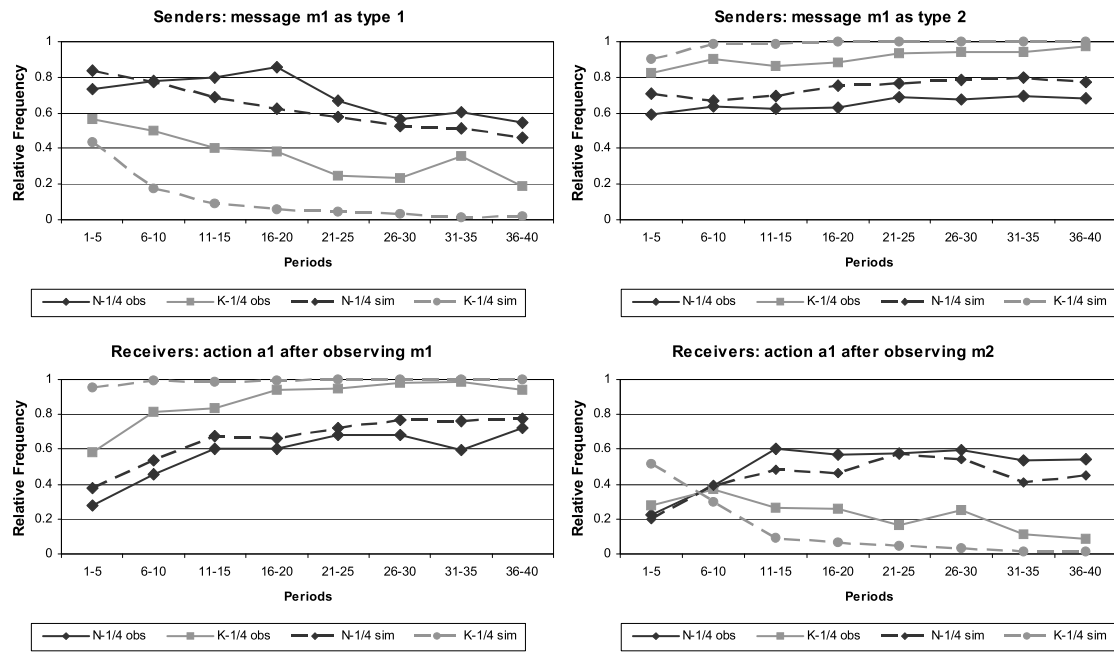
Fig. 5. Observed and simulated time series of average strategies in treatments N-1/4 and K-1/4.
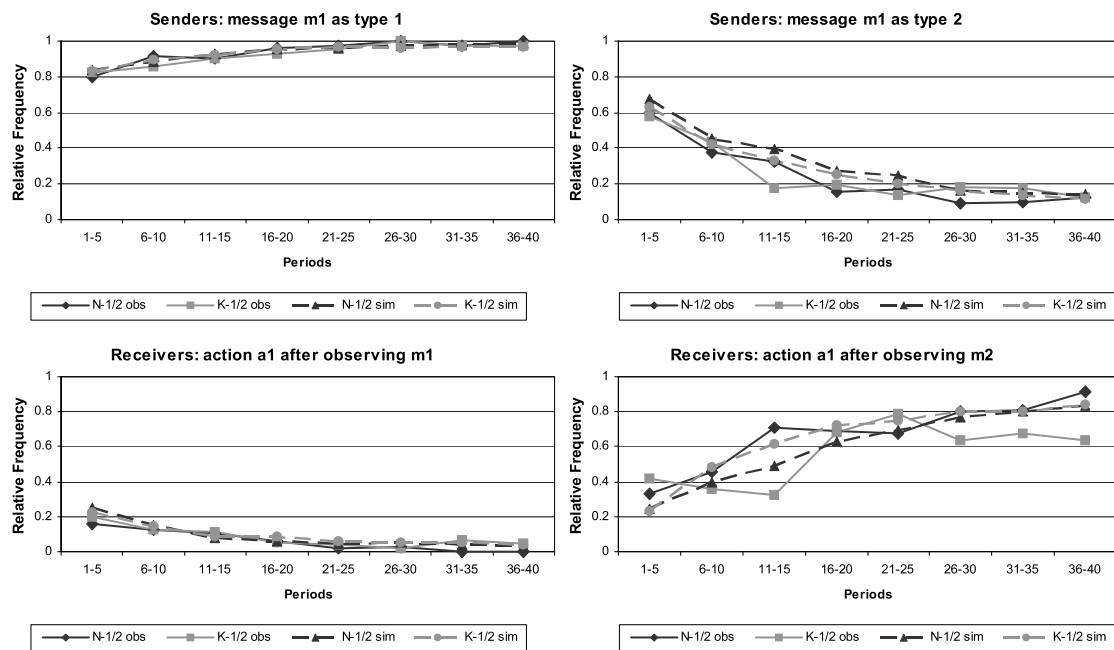

Fig. 6. Observed and simulated time series of average strategies in treatments N-1/2 and K-1/2.

model also slightly overestimates the choice of action a_1 after observing message m_1 and underestimates it after message m_2 in both treatments (simulated receivers after observing message m_1 learn more quickly to choose action a_1 if the prior $p = 1/4$ is known, see lower left panel in Fig. 5).

Let us now turn to Figs. 6 and 7 that show the results of the simulations in the treatments with $p = 1/2$ and $p = 3/4$, respectively. In all these treatments the fit is very good for the sender if the type is t_1 and for the receiver conditional on observing message m_1. In fact, the fit in these cases is so good that the lines representing observed and simulated data can hardly be distinguished. Furthermore, whereas the fit for type t_2 senders is also quite good in these treatments, the learning model only has some problems tracking the behavior of receivers conditional on observing message m_2 in treatment K-3/4, as the predicted share is consistently lower than the observed share.

Summarizing the results of the simulations, it seems fair to say that they account for the primary patterns found in the data. Strategies of simulated players follow a similar evolution as in the actual experimental sessions. Note that our learning model is quite parsimonious as it has only three parameters (A_1, A_1^T, and λ), which we did not allow to depend on the treatment. Nevertheless, the model replicates the patterns of strategy adjustment in all six treatments reasonably well.
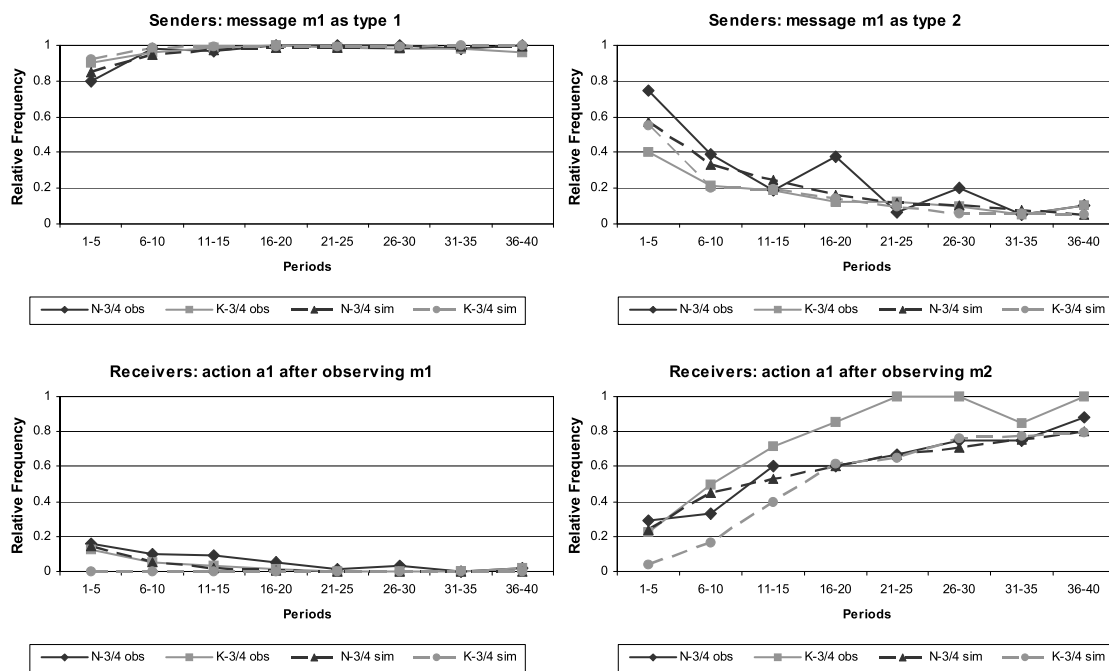
Fig. 7. Observed and simulated time series of average strategies in treatments N-3/4 and K-3/4.

Since in some treatments convergence of behavior is faster than in others, the fit is less than perfect, as discussed above. To obtain more insight into this issue and also to check the predictive power of the model, we estimated parameters for each individual treatment separately, and also conducted out-of-sample estimations of parameters for each treatment.[27] For the latter, we used data of 5 of the 6 treatments to estimate the best-fitting parameters for these 5 treatments and then used these estimated parameters to predict the evolution of strategies in the 6th treatment. For treatment K-1/4, the fit is noticeably better if the parameters are estimated from K-1/4 data only. In this treatment, the convergence of behavior appears to be much slower in the actual data than in the other treatments. For treatment N-1/2 the fit is worse and convergence is slower than in the data if the parameters are estimated out-of-sample. However, for the other four treatments these alternative estimation methods either did not change the estimated parameters or led to only a small deterioration of the goodness-of-fit metric. Therefore we are confident that our parsimonious model with the same parameters for all treatments describes behavior quite well.

Note that the best-fitting values of the initial propensities on strategies (A_1) and types (A_1^T), respectively, are estimated to be different ($A_1 = 2$, $A_1^T = 15$). Since both of these belief propensities are updated by 1 with each observation, the lower value for A_1 means that the players adjust their beliefs about the opponent's strategy faster than their beliefs about the probabilities of the chance move. As discussed in the preliminary description of the learning model in Section 2, this is precisely what is needed for the model to produce the differences we observe. We note that there is psychological evidence that, indeed, people are more ready to form beliefs about a person's performance than about "objective" uncertainty based on a small sample (Nickerson, 2004, Ch. 8). This is similar to a faster revision of beliefs about strategies used by a rival than about types.

Below we check the robustness of our model by considering alternative initial beliefs and analyzing different speeds of learning further.

5.2. Robustness checks

As already pointed out earlier, we view belief revision and best response adjustment as natural in our binary decision setting with sufficient feedback about the environment. Since the basic learning model is capable of generating different patterns of adjustment depending on parameters and on initial conditions, for the robustness checks we keep its structure while varying initial beliefs and parameters.[28] We perform two sets of robustness checks. In the first set we vary initial beliefs and in a second set we check what happens when learning about initial beliefs or learning about strategies is (essentially) suppressed.[29]

---

[27] The details of these estimations are reported in Section 7 of the Supplementary material.

[28] Brandts and Holt (1992, 1993, 1996), Cooper et al. (1997), and Anderson and Camerer (2000) discuss different adjustment theories ("naive" dynamic, fictitious play, EWA) for experimental signaling games. All of these theories are based on belief revision in view of experience and on best response. Thus their long-run predictions are likely to be similar, although the speed of adjustment may differ.

[29] The technical details of these checks are presented in Section 8 of the Supplementary material.

5.2.1. Robustness checks with alternative initial beliefs

One of the key assumptions in the model of Section 2, also adopted in the baseline simulations, is that initial beliefs are uniform. We run simulations and estimate best-fitting parameters for the following alternative specifications of initial beliefs:

Maximin beliefs: Beliefs that the opponents employ the maximin criterion (the sender chooses m_2 for both of the types and the receiver chooses a_2 in response to both messages).

Level-2 beliefs: Beliefs that the opponents choose a strategy that is a best response to uniform beliefs.

Period-1 beliefs: Beliefs equal to the actual distribution of behavior in Period 1, assuming that players—unrealistically—are able to predict this distribution.

Alternative initial priors: In all the alternative specifications above we assume that $p^T = 0.5$, i.e., in the treatments with an unknown prior distribution over the sender's types, players are assumed to entertain a flat prior. We also run simulations with alternative values, namely $p^T = 0.25$ and $p^T = 0.75$ (with beliefs on strategies assumed to be uniform).

The estimation results for all of the alternative specifications show that the model with uniform beliefs does no worse and most of the time actually does better than models with alternative initial beliefs.[30]

The above alternative models focus on different initial beliefs while keeping the basic adjustment process itself constant. Below we check what happens when the main parameters of the learning model (speed of belief revision, represented by the strength of the initial propensities A_1 and A_1^T) are taken to the extreme.

5.2.2. Robustness checks with learning about types or strategies suppressed

To check the importance of learning about the strategies of the opponent and about the distribution of types. We run simulations and estimate best-fitting parameters for the following two alternative specifications (assuming initial beliefs on types and the opponent's strategies to be uniform)[31]:

Strategy learning only: We fix the parameter A_1^T at a very high value such that there is almost no updating of initial beliefs about sender-type probabilities and learning is essentially about strategies used by the opponents. The main result of this exercise is that the fit for treatment *N*-1/4 becomes much worse than for the baseline model with uniform beliefs and a "freely" estimated parameter A_1^T. The fit for the other treatments does not change much, since in the *K* treatments there is no learning about the type distribution anyway, while for treatments *N*-1/2 and *N*-3/4, learning the type distribution does not change the best response structure with respect to a known (or initially believed) $p = 1/2$. We interpret this result as evidence in favor of a type component of the learning process of subjects, which manifests itself as important in treatment *N*-1/4, where it indeed matters.

Type learning only: Here we fix the parameter A_1 at a very high value such that there is almost no updating about strategies and learning is essentially about the types of the sender. The fit becomes much worse for all treatments except for *N*-1/4, where it is still worse than for the uniform belief model with a lower (estimated) value for A_1, but not by much. This shows that there is learning about strategies in all treatments while it is perhaps less important than learning about types in the *N*-1/4 treatment.

The results of the models with suppressed learning about one aspect of the environment (type or strategy) indicate that both aspects are important but to a different extent. In treatments with a known p, learning can be about strategies only. However, the results confirm the intuition from the model in Section 2 that in treatments *N*-1/2 and *N*-3/4, learning the type distribution is not that important because it does not change the best response structure compared with uniform initial beliefs. By contrast, in treatment *N*-1/4 subjects do seem to learn about the distribution of types and adjust their behavior accordingly.

Overall, alternative models of initial beliefs do not lead to a better fit than the baseline model assuming uniform beliefs about strategies and a flat prior about types. Furthermore, and as might be expected, models with extreme assumptions regarding the strength of the initial propensities on strategies and types produce a worse fit, but they also give an idea on how important the speeds of learning are for different components of the model. We are thus quite confident that our baseline simulations provide a reasonable account for the aggregate adjustment behavior observed in the experiment. An important result of the model estimation is that learning about strategies is faster than learning about types—a condition that was needed for the intuition of the simple model in Section 2 to produce the different outcomes for known and unknown distributions of the sender's type.

6. Summary and conclusions

In this paper we report the results of an experiment designed to test the effect of varying the initial information about the prior belief in incomplete-information situations using a two-person signaling game. Specifically, we study the effect

---

[30] These estimation results are reported in the first six rows of the table in Section 8 of the Supplementary material.

[31] The results of these estimations are reported in the last two rows of the table in Section 8 of the Supplementary material.

of this manipulation on the outcome that the signaling game converges to. We implemented six versions of this signaling game that differ in the underlying prior distribution of the sender's type ($p \in \{1/4, 1/2, 3/4\}$) and whether or not the prior distribution is commonly known. Feedback in the experiments was such that players had the opportunity to learn over time about the strategy of the other player and the prior distribution of the sender's type. Furthermore, the game was designed so that, according to a plausible adjustment theory, early play in the games with a low value of p may direct behavior towards different long-run outcomes depending on p being commonly known or not, while initial information about p should be inconsequential with regard to long-run behavior in the games with higher values of p.

The main result of our paper is that these predictions are indeed borne out by the data. More precisely, whereas we find significant differences in long-run outcomes in the two games with $p = 1/4$, we cannot reject the hypothesis that there is no difference in long-run outcomes between corresponding pairs of games for the other values of p. The second main result is that the learning model we present (in which players learn about both the other player's strategy and the prior distribution of the sender's type in case it is not known) quite accurately tracks the main features of the adjustment process in our observed data. Furthermore, we show that assuming a flat prior regarding initial beliefs about the sender's type in case it is not known fits the data quite well and that learning about the other player's strategy is faster than learning about the sender's type.

We view our results as an existence proof that there are games (or, more speculatively, classes of games) with incomplete information for which inducing or not inducing a common prior may make a difference with respect to agents' long-run behavior and, hence, the outcomes in (experimental) games. Our results should be of some interest in the light of the results by Güth and Ivanova-Stenzel (2003) who showed that commonly known beliefs are inessential for bidding behavior in specific asymmetric first-price and second-price independent-value auctions.

Our results also contribute to the discussion of equilibrium selection principles by providing further evidence that a dynamic process of adjustment can predict behavior and organize the data well. This approach was previously used in e.g. Brandts and Holt (1993, 1996) for signaling games, and in Haruvy and Stahl (2004) for coordination type games with multiple strict equilibria. We also show that initial play (and thus subsequent adjustment) can be manipulated towards the basin of attraction of a different equilibrium, in our case by withholding the information about the prior distribution of types.

So far the results presented here and elsewhere on the effect of (not) inducing a common prior on long-run outcomes in games are not more than isolated findings and more research effort is necessary and desirable to get a fuller picture of this effect. We see various avenues for fruitful future research. First, it would be interesting to consider different sender–receiver games and to test the ability of the learning model to predict adjustment patterns in these games, depending on whether or not a common prior is induced. Also, it may be useful to analyze financial markets in which investors learn about the distribution of returns or states of nature (for related issues, see Bossaerts's, 2002, "efficient learning markets" in which agents learn about possibly incorrect priors). Second, a more systematic inquiry into the relationship between the feedback given to players and the adjustment path in games with and without a common prior is called for. In this paper we concentrated on the arguably most simple case where subjects were informed about nature's move in their individual encounters at the end of each period. What happens if this information is not given to subjects or if information about the realization of the players' type can only be derived indirectly or ambiguously from payoffs? Third, in this paper we chose to abstain from properly eliciting beliefs due to the assumed complexity of the situation from a subject's perspective. But it could be interesting to extend the analysis to directly learning about what beliefs subjects hold. Finally, other important classes of games with incomplete information, (not) commonly known priors and multiple equilibria such as market games with incomplete information about costs or demand (e.g. price competition in homogeneous markets where firms have convex costs) could be subjected to thorough experimental tests.

Appendix A. Experiment instructions

A.1. Instructions

Please read these instructions carefully! There are other people taking part in the experiment. Please do not talk to them and remain quiet throughout the experiment. If you have a question, please raise your hand. We will come to you to answer it.

In this experiment you can earn varying amounts of money, depending on which decisions you and other participants make. The experiment will consist of 40 rounds, in each of which you can earn Points. Your payoff at the end of the experiment is equal to the sum of your own payoffs from all rounds. For every 10 Points you will be paid 5p.

A.2. Description of the experiment

In each round of the experiment, two participants are randomly matched and interact with each other: one participant will be called "A-participant" and the other "B-participant". The A-participant can be of two types that we call "Type 1" and "Type 2". At the beginning of each round, a random draw determines the type of the A-participant. The random draw is such that with an $X\%$ chance the A-participant will be of Type 1, and with a $(100 - X)\%$ chance of Type 2. You receive no information about the value of X, except that X is constant over all rounds of the experiment. After the random draw, the

A-participant will be informed about his/her type. However, the B-participant will not be informed about the type of the A-participant. Knowing his/her type, the A-participant has to decide between options "C" and "D". Then, the B-participant will be informed about which option was chosen by the A-participant. Knowing the option chosen by the A-participant, but not knowing his/her type, the B-participant will now have to choose between options "E" and "F". After that, the payoffs of the two participants are determined according to the tables overleaf.

[Instead of the sentences with X, it was "The random draw is such that with a 25% chance the A-participant will be of Type 1, and with a 75% chance of Type 2." in the treatment with known prior distribution $p = 1/4$, and the corresponding probabilities for the other treatments with known prior distribution.]

A.3. Payoffs

Generally, the payoffs of both participants depend on the A-participant's type, the option chosen by the A-participant and the option chosen by the B-participant.

A.4. The A-participant's payoffs

The payoffs of the A-participant (in blue) in each round are given in the following two tables (along with the B-participant's payoffs in red). For the A-participant of Type 1 the table on the left applies and for the A-participant of Type 2 the table on the right applies:

Payoff table for Type 1 of the A-participant:

		Decision of the B-participant	
		E	**F**
Decision of the	**C**	**15**, 10	**80**, 80
A-participant	**D**	**25**, 10	**50**, 50

Payoff table for Type 2 of the A-participant:

		Decision of the B-participant	
		E	**F**
Decision of the	**C**	**80**, 80	**15**, 30
A-participant	**D**	**50**, 50	**25**, 30

A.5. The B-participant's payoffs

The payoffs of the B-participant (in blue) in each round are given in the following two tables (along with the A-participant's payoff in red). If the A-participant chose option "C", the table on the left applies, if the A-participant chose option "D", the table on the right applies:

Payoff table for the B-participant if A-participant chose option "C":

		Decision of the B-participant	
		E	**F**
Type of the	**1**	15, **10**	80, **80**
A-participant	**2**	80, **80**	15, **30**

Payoff table for the B-participant if A-participant chose option "D":

		Decision of the B-participant	
		E	**F**
Type of the	**1**	25, **10**	50, **50**
A-participant	**2**	50, **50**	25, **30**

A.6. Summary

To give you an overall picture of the rules, the timing of events in each round can be summarized as follows:

1. The computer randomly matches participants in pairs.
2. The computer randomly determines the A-participant's type. With an X% chance the A-participant will be of Type 1 and with a $(100 - X)$% chance of Type 2. You receive no information about the value of X, except that X is constant over all rounds of the experiment. [The last two sentences were "With a 25% chance the A-participant will be of Type 1, and with a 75% chance of Type 2." in the treatment with known prior distribution $p = 1/4$, and corresponding probabilities for the other treatments with known distribution.]
3. The A-participant is informed about his/her type. Then the A-participant chooses between options "C" and "D".
4. The B-participant is informed about the choice of the A-participant, but not about his/her type. Then the B-participant chooses between options "E" or "F".
5. Payoffs result as described above.

A.7. Number of rounds, role assignment and matching

The experiment consists of 40 rounds.

The role of either the A-participant or the B-participant will be randomly assigned to each participant in the room at the beginning of the experiment. (The computer program makes sure that half of the participants will be assigned the role of an A-participant and the other half the role of a B-participant.) You will then keep the same role during the entire experiment.

In each round the computer will randomly match participants in pairs of two (one A-participant and one B-participant) from a group of eight subjects. The matching is completely random, meaning that there is no relation between the participant you have been matched with last round (or any other previous round) and the participant to whom you will be matched with in the current round.

Supplementary material

The online version of this article contains additional Supplementary material.
Please visit doi:10.1016/j.geb.2011.05.004.

References

Anderson, C.M., Camerer, C.F., 2000. Experience-weighted attraction learning in sender–receiver signaling games. Econ. Theory 16, 689–718.

Banks, J., Camerer, C., Porter, D., 1994. An experimental analysis of Nash refinements in signaling games. Games Econ. Behav. 6, 1–31.

Bernoulli, J., 1713. Wahrrscheinlichkeitsrechnung, third and fourth parts. Ostwald, Klassiker der exakten Wissenschaften, vol. 108 (Leipzig, 1896), translation of "ars conjectandi".

Bossaerts, P., 2002. The Paradox of Asset Pricing. Princeton University Press, Princeton, NJ.

Brandts, J., Holt, C.A., 1992. An experimental test of equilibrium dominance in signaling games. Amer. Econ. Rev. 82, 1350–1365.

Brandts, J., Holt, C.A., 1993. Adjustment patterns and equilibrium selection in experimental signaling games. Int. J. Game Theory 22, 279–302.

Brandts, J. Holt, C.A., 1996. Naive Bayesian learning and adjustment to equilibrium in signaling games. Working paper, University of Virginia.

Cooper, D.J., Garvin, S., Kagel, J.H., 1997. Signalling and adaptive learning in an entry limit pricing game. RAND J. Econ. 28, 662–683.

Cooper, R.W., DeJong, D.V., Forsythe, R., Ross, T.W., 1990. Selection criteria in coordination games: some experimental results. Amer. Econ. Rev. 80, 218–233.

Dekel, E., Fudenberg, D., Levine, D.K., 2004. Learning to play Bayesian games. Games Econ. Behav. 46, 282–303.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10, 171–178.

Fudenberg, D., Levine, D.K., 1998. The Theory of Learning in Games. MIT Press, Cambridge, MA.

Greiner, B., 2004. An online recruitment system for economic experiments. In: Kremer, K., Macho, V. (Eds.), Forschung und Wissenschaftliches Rechnen, 2003. In: GWDG Bericht, vol. 63. Ges. für Wiss. Datenverarbeitung, Göttingens, pp. 79–93.

Güth, W., Ivanova-Stenzel, R., 2003. Asymmetric auction experiment with(out) commonly known beliefs. Econ. Letters 80, 195–199.

Harsanyi, J.C., 1967. Games with incomplete information played by 'Bayesian' players; Part 1: The basic model; Part II: Bayesian equilibrium points; Part III: The basic probability distribution of the game. Manage. Sci. 14, 159–182, 320–334, 486–502.

Harsanyi, J.C., Selten, R., 1988. A General Theory of Equilibrium Selection in Games. MIT Press, Cambridge, MA.

Haruvy, E., Stahl, D.O., 2004. Deductive versus inductive equilibrium selection: experimental results. J. Econ. Behav. Organ. 53, 319–331.

Haruvy, E., Ünver, M.U., 2007. Equilibrium selection and the role of information in repeated matching markets. Econ. Letters 94, 284–289.

Jaynes, E.T., 2003. Probability Theory: The Logic of Science. Cambridge University Press, Cambridge, UK.

Keynes, J.M., 1921. A Treatise on Probability. MacMillan, New York, NY.

Nickerson, R.S., 2004. Cognition and Chance. Lawrence Erlbaum Associates, Mahwah, NJ.

Noussair, C.N., Plott, C.R., Riezman, R.G., 1995. An experimental investigation of the patterns of international trade. Amer. Econ. Rev. 85, 462–491.

Oechssler, J., Schipper, B., 2003. Can you guess the game you're playing? Games Econ. Behav. 43, 137–152.

Siegel, S., Castellan Jr., N.J. , 1988. Nonparametric Statistics for the Behavioral Sciences. McGraw–Hill, New York.

Sinn, H.-W., 1980. A rehabilitation of the principle of insufficient reason. Quart. J. Econ. 94 (3), 493–506.

Stahl II, D.O., Wilson, P.W., 1994. Experimental evidence on players' models of other players. J. Econ. Behav. Organ. 25, 309–327.

Straub, P.G., 1995. Risk dominance and coordination failures in static games. Quart. Rev. Econ. Finance 35, 339–365.

Van Huyck, J.B., Battalio, R.C., Beil, R.O., 1990. Tacit coordination games, strategic uncertainty, and coordination failure. Amer. Econ. Rev. 80, 234–248.

Van Huyck, J.B., Battalio, R.C., Beil, R.O., 1991. Strategic uncertainty, equilibrium selection, and coordination failures in average opinion games. Quart. J. Econ. 106, 885–910.

Van Huyck, J.B., Cook, J.P., Battalio, R.C., 1994. Selection dynamics, asymptotic stability and adaptive behavior. J. Polit. Economy 102, 975–1005.

Van Huyck, J.B., Cook, J.P., Battalio, R.C., 1997. Adaptive behavior and coordination failure. J. Econ. Behav. Organ. 32, 483–503.

Van Huyck, J.B., Battalio, R.C., Rankin, F.W., 2007. Selection dynamics and adaptive behavior without much information. Econ. Theory 33, 53–65.