On the influence of imitation costs on information acquisition through imitative behavior: an experimental analysis

Christian Feige[1]

Chong-Dae Kim[2]

Nannan Zhou[3]

University of Cologne

[1] no affiliation, Gell'sche Str. 78, 41472 Neuss, Germany, phone: +49 2131 5245094, e-mail: chrisfeige@gmx.de.
[2] no affiliation, e-mail: cdekim@gmx.de.
[3] University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany, phone: +49 221 470-6311, fax: +49 221 470 5078, e-mail: nannan.zhou@uni-koeln.de.

Abstract

We investigate the effects of imitation costs on individual imitative behavior. For this purpose, we introduce a model of a modified prediction game with informed and uninformed individuals and imitation costs. The theoretical implications are tested in a laboratory experiment. We show theoretically that there is a negative relation between imitation costs and imitative behavior. Moreover, we derive threshold values for imitation costs for (non-)imitation to be the strictly dominant strategy. However, the theoretical implications can only be partially confirmed by the experimental results. While a significant number of uninformed subjects do not imitate in a low-cost scenario, imitative behavior can still be observed in a high-cost scenario.

Key Words: Imitation costs, imitation, information transfer, expert advice, belief learning, predictions

JEL Classification: D8, D83,C91, C73, D82,C53

1. Introduction

Information is a good that is only traded with difficulties, as the value of a piece of information is not only subjective, but can moreover decrease abruptly once the trade is accomplished. As soon as information becomes public knowledge, its value is zero, so it is worth the less, the more people know it. In general, people are only willing to pay for information which they *don't* have.[4]

In particular on the financial market, information is of crucial importance in a competition for the best predictions (e.g. of stock values). For instance, an illegal competitive advantage by means of insider knowledge is worth hard cash. A legal, but more work-intensive, and thus costly, method of procuring a comparable informational advantage is, for instance, a market analysis on which financial service providers base their forecasts. But how can this informational advantage be monetarily quantified in terms of money, that is, what is a fair price for an investment fund that gives more weight to stocks predicted to perform better in the future or for a financial advisor's recommendation to buy?

In this paper, we introduce a model of a modified prediction game with imitation costs and test the theoretical implications in a laboratory experiment. Our model aims to contribute to the analysis of pricing information exchange and its effects on the utility of the market participants and, not the least, also on the individuals' imitative behavior and the quality of the predictions made. In doing so, no distinction will initially be made between the factors besides information value that might influence the price of this information exchange. Instead, it is only assumed that this exchange is hindered in some way (as opposed to the notion that the transfer of knowledge is an exchange of non-rival and non-exclusive goods), which only in one interpretation is actually a price for the communicated piece of information, but more generally is a cost factor, namely, the marginal cost of making another copy of a particular piece of information.[5]

What are imitation costs? Generally speaking, imitation costs are costs occurring during or because of an imitation process, in which an "imitation source" (a piece of information, an entire technology or strategy, or a technical appliance or other economic good) is reproduced or copied in a certain way, so that the copy can be employed by the imitator for other purposes. There are several ways in which such a cost factor can be involved in the imitation process:[6]

- a *production factor*, i.e. the cost of production materials, time, and effort used to make or retrieve such a copy
- a *purchasing price*, i.e. money (e.g. a consultant's salary or risk premium) paid to the imitated party to reimburse them for their own development efforts

[4] See e.g. DeLong and Froomkin (2000).
[5] Ibid.
[6] Since imitation processes are usually associated with copying errors, like distortions of the transferred information which then cause misunderstandings, this might be perceived as another kind of imitation costs. All imitation is certainly probabilistic to some extent or this notion wouldn't be this important to theories of cultural evolution as a cause of mutation. However, this kind of costs doesn't lend itself to be easily modeled as an additive cost term and therefore isn't covered by our model.

- a *transaction fee*, i.e. money (e.g. an advisory tax) paid to a third party (usually the government) to regulate transfers of this kind.

The model we use tends more towards the third option of a transaction fee, because our *experts* don't actually receive the money deducted from their imitators as imitation costs. An extension of our model to involve the *experts* more strongly in the decision by awarding them some or all of the money paid to imitation them is certainly possible, however. Furthermore, since we focus mainly on the imitation of information here, the notion of actual reproduction costs involving time or other production factors plays only a negligible role for our purposes.

Finally, a usually preliminary stage of imitation, which is included in our model, too, involves the search for the best imitation source. Such a search is a learning process, which in itself is again associated with costs. These learning costs include again time and effort, but also "mistakes" in the form of lost payoff through choosing the wrong (that is, not the best) imitation source. In our model, we don't specifically differentiate between learning costs and imitation costs, but instead regard learning costs as just one further component included in the imitation cost factor.

There is a broad number of studies concerning imitation and learning from the behavior of others. In the following, we present several important perspectives and relate them to our own research. For an overview paper on social learning in the area of economics see e.g. Bikhchandani et al. (1998).

1.1. Prediction algorithms with expert advice

The theoretical roots of our model ultimately trace back to Vovk (1996) and DeSantis et al. (1988) from the field of computational (or resp. formal) learning theory, modeling predictions drawing on "expert advice". In these models, one's own prediction for a given period is an aggregate of all other available predictions for this period. Additional variants of these models are described comprehensively e.g. by Cesa-Bianchi and Lugosi (2006).

Schurz (2008) describes this algorithmic way of generating predictions as a kind of imitative behavior, calling it "meta-induction". The strategy of using "meta-induction" to make a prediction can be defined as follows:

> *"Meta-induction"*: After all actors have already made their predictions, the imitator gains access to one (or several) of his competitors' predictions and is then allowed to change his own forecast.

This corresponds to the procedure of an inductive inference, in which future predictions are generated from data collected in the past. The results presented in our paper should by and large also apply to "meta-induction". However, the major aspect that distinguishes our model from models used by computational learning theory is that we assume imitation to happen "blindly", i.e. the imitating player neither observes the imitated prediction nor receives any (direct) information on the decision choice taken by the person he imitates, but only obtains payoff as if he has made the prediction himself.

Regarding the probability distribution of the individual outcomes that have to be predicted by the players, authors investigating comparable models make quite general assumptions. The sequence of outcomes considered by Cesa-Bianchi and Lugosi (2006) can for example be a stochastic or a deterministic process or can even adapt to, or resp. counteract, the players' predictive behavior. By contrast, in our model we assume outcomes to be distributed identically and independently from each other to make the mathematical treatment easier. Furthermore, the imitation costs introduced in our model again also don't appear in any of the theoretical approaches discussed by Cesa-Bianchi and Lugosi.

1.2. Imitation costs in the area of product imitation

The idea of costly imitation has been pursued before in economics, namely in connection with product imitation. Several authors (e.g. Mansfield et al. (1981), Harabi (1991), as well as Laurent (2008)) have contrasted the costs of developing a new product or new product attributes (innovation costs) with the costs of reverse-engineering an already existing product and then selling a copy under a new name (imitation costs)

Mansfield et al. (1981) present an empirical study comparing imitation costs in relation to innovation costs for 48 products developed by American firms in the industrial sectors of chemicals, drugs, as well as electronics and machinery. Their results indicate that imitation is on average cheaper than innovation (on average 65% of innovation costs for the same product) and usually also take less time to carry out, i.e. imitation time is on average 70% of innovation time. However, they also observe that, overall, imitation is either comparably cheap (less than 40% of imitation costs) or rather expensive (more than 90% of imitation costs), splitting the sample of products about half and half into a group of very profitable targets for imitation and one not so profitable. Harabi (1991) arrives at similar results in a study among 358 Swiss firms.

Laurent (2008) describes a theoretical model dealing with the imitation or resp. innovation of new product attributes (e.g. materials used, product color, etc.), providing that imitating a certain product is advantageous in terms of cost to newly developing a comparable product. He shows that, depending on the level of imitation costs, firms may change the attributes of their products to prevent imitation or at least to keep it at an acceptable level.

Our imitation model differs from these approaches mostly in that it simulates a repeated interaction between the competitors and that the outcome of the decisions made is in part determined by chance. Product imitation, by contrast, is usually more of a one-shot situation, because a product need only be reverse-engineered once, in order to be produced again in the future. In our model, this reverse-engineering would compare to the "expert" not only providing the "correct answer" for a given decision problem, but also a "recipe" for more correct answers in the future, or more precisely, the technology of producing predictions of expert quality by oneself.

1.3. Other relevant economic research

Schlag (1998) describes a dynamic model of imitative behavior when facing a multi-armed bandit, which put simply is a more complex decision under risk. He models imitation as a recurring choice between a boundedly rational individual's own strategy and that of a single other player, giving several decision rules on when to switch to a different strategy. While showing that this imitative behavior can on an aggregate level be approximated by a replicator dynamic, Schlag also proves a result (his "Lemma 1") that is quite similar in structure to our Proposition 1 (see Section 3). More precisely, similar to our own approach, he compares a difference of probabilities of certain actions occurring with a difference of payoffs, either of which can be obtained through these actions.

Kirchkamp and Nagel (2003), who investigate imitation of cooperative strategies in the prisoner's dilemma, show that imitation is indeed a possible reason for why players choose a particular action, but so are other factors like reciprocity and experience gained through learning processes. The framework used by Kirchkamp and Nagel (2003) differs from our model in that they assume imitation to happen only in a local area, i.e. between "neighboring" players. Although this might be a straightforward extension of our model, we limit ourselves to the more simple case in which all members of an environment are available for imitation.

Along other lines, imitation is often mentioned in conjunction with informational cascades and herd behavior, both of which have been extensively studied theoretically (see e.g. Bikhchandani et al., 1992; Welch, 1992), and observed both in the field (see e.g. Graham, 1999; Ashiya and Doi, 2001) and in the laboratory (see e.g. Anderson and Holt, 1997; Drehmann et al., 2004). Moreover, the presence of private information as a source for imitative behavior has, for instance, been demonstrated by Conlisk (1980).

In addition, there is a contextual relation to the concept of "prediction markets" (e.g. Wolfers and Zitzewitz, 2004) and more application-related topics like e.g. informed trading, with the idea that there is a group of participants ("experts" or "insiders") who are informed about a state of nature in contrast to other market participants in the stock market (e.g Grossman and Stiglitz, 1976). On these markets, cascade-like behavior is sometimes observed, i.e. the uninformed participants tend to imitate the trading decisions of the informed "insiders", which leads to convergent earnings of all market participants over time (see e.g. Plott and Sunder, 1982).

On behalf of experimental economic research, one should also mention e.g. Heemeijer et al. (2006), Offerman et al. (1997). As in all of these models a large number of small decisions are made, which taken by themselves each only have little impact, many aspects of behavioral learning theory also become important, which are comprehensivelydescribed e.g. by Erev and Haruvy (forthcoming).

While Heemeijer et al. (2006),who show that price expectations can influencefuture prices and even drive them away from the market equilibrium, and Offerman et al. (1997),who show that the more information firms have about their competitors, the more readily they imitate other firms, concentrate on the idea of arriving at an equilibrium outcome through the use of imitative behavior, we in contrast put the spotlight on the actual imitation process and leave such questions of efficient behavior alone (at least for the moment). Furthermore, we establish that, independently from the existing information level, imitation behavior decreases with the costs that this behavior entails.

All of these different theoretical approaches studying imitative behavior share the fact that the costs of performing such an imitation are ignored.[7] Conlisk (1980) even assumes that the imitation of private information is the less expensive alternative because of "optimization costs", referring to the effort of looking for the best solution to a problem on one's own. But why should an agent who possesses private information just give this advantage away without at least trying to protect it from imitation or even to exploit it for monetary gain? In contrast, we therefore include such a cost factor in our model to shed light on some new aspects in this regard.

The remainder of this paper is structured as follows: Section 2 presents a model of a prediction game with imitation costs, followed by the equilibrium analysis in Section 3. Section 4 describes our experiment with respect to experimental procedure,experimental design, and behavioral predictions. Section 5 presents the results regarding the effects of imitation costs on imitative behavior in different treatments. Section 6 concludes and gives an outlook on additional possible research based on these findings.

2. The Model

Similar to Cesa-Bianchi and Lugosi (2006) a **prediction game** basically consists of a **sequence of outcomes** $y_1, y_2, \ldots$ whose future development is predicted by the participating **players** (finite ber $k \in \mathbb{N}$). This sequence of outcomes is spread evenly across a finite number of $t \in \mathbb{N}$**periods** (one outcome obtaining in every period).

As mentioned in Section 1.2, in order to facilitate the mathematical treatment of the model, it will be assumed here that the **individual outcomes are distributed randomly**, to wit, independently of each other with an identical discrete probability distribution. The set of all possible outcomes Y is called the outcome space. For the sake of simplification, it will also be assumed in the following that the distribution parameters are common knowledge, so that a rational decision-maker doesn't rely on past outcome data.

The independence assumption makes sure that, with regard to their forecast, the players have to treat every period separately, without being able to base their predictions on past outcomes (at least not as a

[7] An exception may be Welch (1989) who uses a cost factor to model the imitation of "high-quality" firms, although what he calls "imitation costs" are actually signaling costs to profess a higher level of quality.

part of a “rational” strategy). Without this assumption, you might e.g. be able to infer from very high results in the past that future results will likewise be very high (or very low) with a larger than chance probability. Of course, the players’ actual behavior (e.g. the decision between imitation and a self-generated forecast) may still depend on past outcomes (and the experimental results can in fact establish a correlation here; see Section 5). The assumption of an identical distribution is less crucial, but makes a comparison of the different periods easier for the players.

The prediction $p_{i,t}$ made by player i in period t stems from a decision space D which is not necessarily identical to the outcome space Y. For example, a financial advisor might only predict a stock value to rise, fall, or stay about the same (that is, advise to either buy, sell, or hold it), while the stock value itself can attain any positive number. By player i’s **predictive accuracy** we mean the probability $P[p_{i,t} = y_t]$ of this player making a true prediction in any given period t, that is, the higher a player’s probability of making a true prediction, the more accurate his actual predictions will be.

Constituents of the prediction game

$y_1, \ldots, y_t, t \in \mathbb{N}$	sequence of randomly determined outcomes until period t
$y_t \in Y$	outcome in period t
$k \in \mathbb{N}$	number of players
$p_{i,t} \in D$	prediction of player i in period t
$s = s(p_t, y_t)$	score function as an estimator of predictive accuracy
w_H, w_L	possible scoring ($w_H > w_L \geq 0$)
c	imitation costs
$\pi_{i,t}$	payoff of player i in period t

Every single period proceeds as follows: The players first simultaneously and covertly make their predictions p about the realization of this period’s outcome – **either directly by means of theirown prediction or indirectly by way of imitation**, which means that player i selects another player j and adopts his forecast for this period. This imitation happens “blindly”, meaning that player i learns player j’s precise prediction not until the outcome has been realized. Our experimental design assumes, more specifically, that the players only learn whether the other players have predicted correctly or falsely. After the outcomes have been realized, the players are told neither the other players’ outcomes nor their actual prediction. This aims to make sure that the players can't differentiate between player types by looking for clues not related to player performance. [8]

As an example for how blind imitation works, imagine that you endow your financial consultant with a sum of money to let him work with it, instead of investing the money yourself. Accordingly, imitation is understood here as a conditional action depending on another player's choice: "I predict that which player j predicts." A player's imitation can thus happen simultaneously to making own predictions. In addition,it is assumed that always only a single player can be imitated at one time. Afterwards, the outcome obtains and the players receive points for their (or resp. the adopted) prediction according to the following **score function** s:[9]

$$s(p_{i,t}, y_t) = \begin{cases} w_H & p_{i,t} = y_t \\ w_L & p_{i,t} \neq y_t \end{cases}, with\, w_H > w_L \geq 0 \qquad (1)$$

A player's score obviously correlates with his predictive accuracy, i.e., the higher his chances of making a true prediction, the higher his (average or resp. total cumulative) score will be. For this reason, any score-based ranking of the players can also be used to estimate a ranking of these players by their predictive accuracy (compare Corollary 1 in Section 3).

Furthermore, in the investigation of imitative behavior it's plausible to assume that there are certain differences between the players, justifying imitation in the first place. In the here assumed simplest case there are **two types of players –** ***experts*** **and** ***non-experts***. The former differ from the latter by a higher probability of a correct prediction, e.g. due to insider knowledge, i.e. private information about the outcome of a given period. Apart from this, both player types act in a rational and risk-neutral manner.

At the beginning of the game, player types are also private information, that is to say, known only the respective players themselves.[10] In order to make possible a decision for or against imitation, the actors therefore need to be provided with information, based on which they are able to at least indirectly infer their competitors' player type. One way of doing this is a score ranking, made available to all players during their decision and updated every period, which at the same time acts as a performance measure.[11]

Another important component of the model is a **cost factor*c***, incurred in the imitation of predictions. The exact type of these costs (transmission errors, purchase price, etc.) isn't specified in this model. A

[8] In our experiment, for example, the *expert* can't choose certain alternatives as possible predictions, because his private information tells him that these are false. Accordingly, if a player is revealed to have made such a choice, he can't be of type *expert*.
[9] This score function can also be chosen more generally without lessening the validity of the following derivations. At the least it should have a single global maximum at the value of the correct prediction, though the further behavior of the function (continuity, slope, and curvature) is irrelevant.
[10] The exact ratio of player types (e.g. x "*experts*", y "*non-experts*") can nevertheless be generally known. This is also assumed in the experimental investigation conducted here.
[11] See Section 4 for details.

financial advisor will usually demand a small fee for his advice, or he might be required to pass on a small fraction of any invested money to the state as a tax on financial trading.

With consideration of all of the parameters mentioned here, the following **payoff function** results for player i for a given period t:

$$\pi_{i,t}(p_{i,t}, p_{j,t}, y_t) = \begin{cases} s(p_{j,t}, y_t) - c & i \text{ imitates } j \\ s(p_{i,t}, y_t) & i \text{ makes own forecast} \end{cases} \quad (2)$$

Accordingly, a player receives points either based on his own forecast or based on the prediction of an imitated player j, additionally subtracting imitation costs in this case, however.

3. Theoretical Derivations

The model described in the previous section is a sequential game with decisions under uncertainty. It's easy to see that it's a (weakly) dominating strategy for player type *expert* not to imitate in any period, because on average no other player can make better predictions, and to maximize the score function with his own forecast (that is, to make a prediction which is as truthful as possible). The optimal strategy forplayer type *non-expert* corresponds to a weighing of his own predictive accuracy and that of the other players (considering imitation costs). For a given round t this weighing decision can be described theoretically as follows:

Proposition 1 (imitation in period t)

Let $\hat{P}$ denote a subjective or respectively estimated probability. In the prediction game described above, a player i imitates a given player j in a given period t only if

$$\hat{P}[p_{j,t} = y_t] \geq P[p_{i,t} = y_t] + \frac{c}{w_H - w_L}. \quad (3)$$

Proof: See Appendix 2.

First, note that this proposition only states a necessary condition for imitation; it doesn't give any advice on which player to imitate, if (3) is true for more than one other competitor. One way to deal with such a case is given in Corollary 1 below.

The right-hand side of (3) contains the parameters whose values a player knows for certain (at least in this model), because they don't depend on another player's action and remain constant over time. This means that a *non-expert* can use this term as a benchmark to compare against his competitors' performance. For imitation to be worthwhile, it's not enough for the designated *expert* to have a higher predictive accuracy than that of a *non-expert*. He must outperform the *non-experts* by a significant margin to make up for the imitation costs.

In particular, imitation becomes the more advantageous,

- the smaller the imitation costs c,
- the smaller the probability $P[p_{i,t} = y_t]$ of making a true prediction by oneself,
- the larger the difference of payoffs for true and false predictions $w_H - w_L$, and of course
- the higher the probability $P[p_{j,t} = y_t]$ of the imitated player j making a true prediction

This last point, the left-hand side of (3), requires some more explanation. As written, it represents player i's belief that imitating a particular player j ultimately leads to a true prediction for player i, nothing more and nothing less. However, this belief is in turn potentially affected by a number of factors, first and foremost the results from other rounds (both outcomes and actions of other players), in particular if a player might (correctly or mistakenly) believe that his competitors don't act rationally (or at least not all the time). A more detailed version of (3), which includes these factors, also has to take into account the possibility that the imitated player j in turn imitates another player, who may or may not be of type *expert*.

As imitation costs are of a particular interest to us, we provide more detailed results for given ranges of cost factors in the following proposition:

Proposition 2 (imitation costs):

Let c_L and c_H denote lower and upper bounds for the imitation costs c, with $c_L < c_H$. There is always a unique payoff-maximizing strategy for players of type *non-expert* in the prediction game described above. In particular, a *non-expert* player should

(1) always choose imitation if the imitation costsare sufficiently low (i.e. $c < c_L$);

(2) choose imitation in a certain period t if the imitation costs are on a moderate level (i.e. $c_L \leq c \leq c_H$) and if(3) is satisfied at the same time, don't imitate if (3) doesn't hold;

(3) never choose imitation if the imitation costs are too high (i.e. $c_H < c$).

Proof: See Appendix 2.

These three cases of the size of imitation costs characterize three quite different situations:

In the first case, imitation costs are so low that a player doesn't even need to know his competitors' types to imitate. That's because in this case, even the average predictive accuracy of the other players is at least weakly better than his own. Of course, discovering the actual *expert* will still improve a player's expected payoff. Yet the decision in this case isn't about imitating or not, just about <u>who</u> to imitate. Another way to motivate this case is to assume that indirect imitation is highly effective, i.e.

even if one doesn't catch an *expert* player by oneself, there is still a high probability that the *non-expert* that was imitated instead will pick the *expert* (whether through luck or experience).

The second case is probably the most interesting one, as a player's optimal action can change in every round depending on his perception of his competitors. In one round, a given competitor might seem like the best bet for being (or benefiting from) an *expert*, only to make several false predictions in a row in the following few rounds. And with this fluctuation of outcomes, a *non-expert's* confidence in his "designated" *expert* can fluctuate, too, inducing him to prefer imitating in one round and not imitating (or imitating a different player) in the next.

Finally, the third case describes a situation, in which the gain from imitating another player's predictions becomes only theoretical at best. Since imitation costs are too high to be on average outweighed by a better chance of earning a high payoff for a true prediction, a player needs to be lucky to still benefit from imitation. Even luck doesn't help anymore, if costs exceed another bound, to wit, if $w_H < c$. In such a high-cost environment, every player will ignore the choices of his competitors and focus only on his own predictions.

As mentioned above, it is very well possible that inequality(3)is true with regard to <u>several</u> other players, while however only <u>one</u> player can be imitated at a given time. In addition, we must therefore specify which of these players is imitated optimally. In general, this is the player whose imitation yields the maximum payoff compared to the alternative "no imitation". Yet this still leaves the problem of determining, which of the other players will generate this optimal payoff, especially when taking into account the subjective beliefs influencing not only one's own choice, but also that of one's competitors.

In his Theorem 1, Schurz (2008) proves that, under very general conditions, imitating the so far most successfully predicting player is at least approximately optimal. Our experimental design uses an objective ranking based on the players' predictive accuracy to provide the players with some way of weighing their alternatives against each other. This allows players to make use of Schurz's result, employing a strategy that has been shown to perform well in comparable situations, namely, "take the best" (see Gigerenzer and Goldstein, 1996), i.e. pick that alternative which has performed the best up until now according to a given ranking of all available alternatives. While the players might just as well fall back on their own subjective ranking instead of the provided objective one, this design makes it easier to determine whether a given player has actually "taken the best". By the way, "take the best" is by no means always the optimal choice in the game presented here: Since *non-experts* know that there is at least one *expert* in the game, a *non-expert* who currently ranks best should not "take the best" (meaning predict on his own), but imitate another player whom he has more reason to believe to be an *expert*.

Corollary 1 (equilibrium strategies by player type):

The prediction game with imitation costs described in Section 2 of this paper has the following equilibrium strategies:

expert **(independently of (3)):** does not imitate; makes true (i.e. score-maximizing) prediction.

non-expert **((3) is true):** imitates the player that maximizes the left-hand side of (3); makes true (i.e. score-maximizing) prediction.[12]

non-expert **((3) is false)**: does not imitate; makes true (i.e. score-maximizing) prediction.

Proof: A single round of this game has a Bayesian Nash Equilibrium as follows: Inequality (3) states the players' beliefs p about the competitors' predictive success in the game. We will abbreviate this inequality by $p \geq p^* + c^*$ to point out that the right-hand side contains only known information and consists of a prediction term and a cost term. Type *expert* has a dominant strategy not to imitate for all beliefs $0 \leq p \leq 1$. This is because he knows that he has the highest predictive success and (rationally) can't benefit from imitating another player (even if that one was an *expert*, too). Type *non-expert* imitates if (3) holds (i.e. if $1 \geq p \geq p^* + c^*$) and doesn't imitate if (3) doesn't hold (i.e. if $0 \leq p < p^* + c^*$). This is because, if (3) holds, he a) believes that at least one other player can predict better than himself and b) also believes that this improvement will outweigh the imitation costs (at least on average). Who to imitate, if given the choice between several alternatives, is then a matter of maximizing the expected payoff, which depends heavily on these players' believed probability of yielding a true prediction (whether directly or indirectly), which is precisely the left-hand side of (3). Finally, if predicting by oneself, none of the player types has any reasons to make a prediction he doesn't believe to be true, because doing otherwise directly decreases his expected payoff based on predictive success. In the scenario presented here, a rational player shouldn't care about whether or not he is imitated, since this doesn't affect his payoff

Furthermore, for games with long duration (or resp. a large number of periods) the following convergence statement also applies:

Corollary 2 (imitation costs – convergence):

Suppose that *j* is the only *expert* in the game.Let $\hat{P}_t[j = exp]$ be a consistent estimator for the *expert's* identity, so that $\lim_{t\to\infty} \hat{P}_t[j = exp] = 1$, if and only if *j* actually is of type *expert*. Further, assume that imitation costs arenon-prohibitive$(c < c_H)$. Then, for t approaching infinity, the frequency of imitation by *non-experts* approaches 1 almost certainly, i.e. all *non-experts* will eventually start to imitate.

[12] Of course, the latter only refers to the hypothetical case that this player actually makes his own prediction.

Proof: See Appendix 2.

The general idea that is expressed in this corollary is quite straightforward: The longer the gameis repeated, the more information and higher certainty regarding the*expert*can be gained. In the short term there might be some misleading occurrences, with *non-experts* having winning streaks that make them seem better than the other players, but in the long run (assuming that *non-experts* use every bit of information they get) *non-experts* can eventually almost certainly identify an *expert*. And since imitation costs aren't so high as to prevent imitation, *non-experts*should always imitate the identified *expert* after this point. By the way, a similar result should also hold if there is more than one *expert*, with the single difference that one can't say which one of the *experts* will be imitated, only that every *non-expert* will eventually imitate any one of them.

In some cases, e.g. if $\hat{P}_t[j = exp]$ is continuous in t, a particular period t^*can be determined after which all *non-experts* will optimally imitate. For example, using the following estimator (which is also strictly increasing):

$$\hat{P}_t[j = exp] = 1 - [1 - \frac{1}{\#players}]^t \quad (4)$$

as well as a cost factor of $c = 20$ and otherwise the parameter values used in our experiment (see next section), we get $t^* = 7.63$, meaning that starting with period 8 every *non-expert* should imitate. In the first rounds of the game, this estimator actually undervalues the probability of imitating the *expert*, because of indirect imitation, but with increasing time this aspect matters less and less until in the end only the *expert* is considered a target for imitation.

These theoretical results are the foundation for our experimental investigation as described in the following section.

4. The Experiment

4.1. Experimental Procedure

The experiment was conducted in October 2010 at the Cologne Laboratory for Economic Research (Germany). Subjects were students from the University of Cologne and were recruited using the online recruiting software ORSEE (Greiner, 2004). Experimental sessions were computerized using the software z-Tree (Fischbacher, 2007). In total, 40 subjects,of whom 52.5% were female,participated in the experiment, 20 in each treatment. After arrival in the laboratory, each subject drew a card, randomly assigning them a seat in the lab. At the beginning of the experiment, subjects were asked to do a real effort task by determining the correct number of sevens in blocks of random numbers. According to their results in the real effort task,subjects were assigned to a role, either *expert* or *non-expert*.[13] Before the experiment started, subjects read the instructions explaining the procedure, the incentives, and the

[13] In each session, the five subjects with the highest scores in the real effort task were assigned a role of *expert* and the other 15 subjects the role of *non-expert*. This procedure was applied in order to ensure a fair allocation of the roles.

rules of the game. To ensure understanding of the experiment, participants had to answer several control questions about the comparative statics of the game. After all participants had answered all questions correctly, the experiment started. At the end of the experiment, subjects had to fill out a short questionnaire and finally were confidentially paid out their earnings in cash. A typical experimental session took about 90 minutes in which subjects earned on average 14.21 Euros (approx. 19.89 USD).

4.2. Experimental Design

The underlying decision situation of our experiment is a modified prediction game in which subjects are asked to give an as-precise-as-possible prediction of the future development of a randomly generated index. Subjects are randomly assigned to groups of four (one *expert* and three *non-experts*), playing the prediction game for three stages. Each stage consists often consecutive periods where subjects play with the same individuals (partner-matching design). After each stage, the groups are re-matched in such a way that each subject interacts with the same individuals in only one of the three stages.[14] Before the experiment starts, the participants are informed about their own type which remains the same during the entire experiment.

At the beginning of eachperiod, the *experts*receive a private message about the direction in which the index will deflect (*rise* or *fall*). The *non-experts* don't receive this information or any other private message. Thereafter, all subjects are asked to either make their own prediction about the index's future development by choosing one of four alternatives (*rises by more than 10%*; *rises by up to 10%*; *falls by up to 10%*; *falls by more than 10%*) or choose one of the other three group members whose prediction they want to imitate. The development of this index value is determined randomly each period, with each of the four alternatives obtaining with the same probability of 0.25. In the experiment this is achieved by first letting the computer generate a random number between 0 and 1 for each period to determine the respective alternative, with each alternative being assigned a quartile of equal size (0.01 to 0.25, 0.26 to 0.5, 0.51 to 0.75, 0.76 to 1). Depending on which of the four alternatives is chosen in this manner, the new index value is calculated by multiplying the previous index value by 1.15 (rises by more than 10%), 1.05 (rises by up to 10%), 0.95 (falls by up to 10%), or 0.85 (falls by more than 10%), respectively. For example, if the index value after Period 4 was 85 and the index was determined to rise by up to 10% in Period 5, the new index value displayed after Period 5 would be 89.25.[15]

[14] After every stage the groups are re-matched with the intention that no participant should encounter a given other participant more than once, so that every session generates 15 independent observations.

[15] The index development was actually calculated in advance to save computing power. Furthermore, the same 30 index values were used for both treatments (10 for each of the three stages), with the idea of increasing the compatibility between the two treatments. Yet since there wasn't enough data for a stage-based comparison between the two treatments, we didn't include this aspect any further in our analysis.

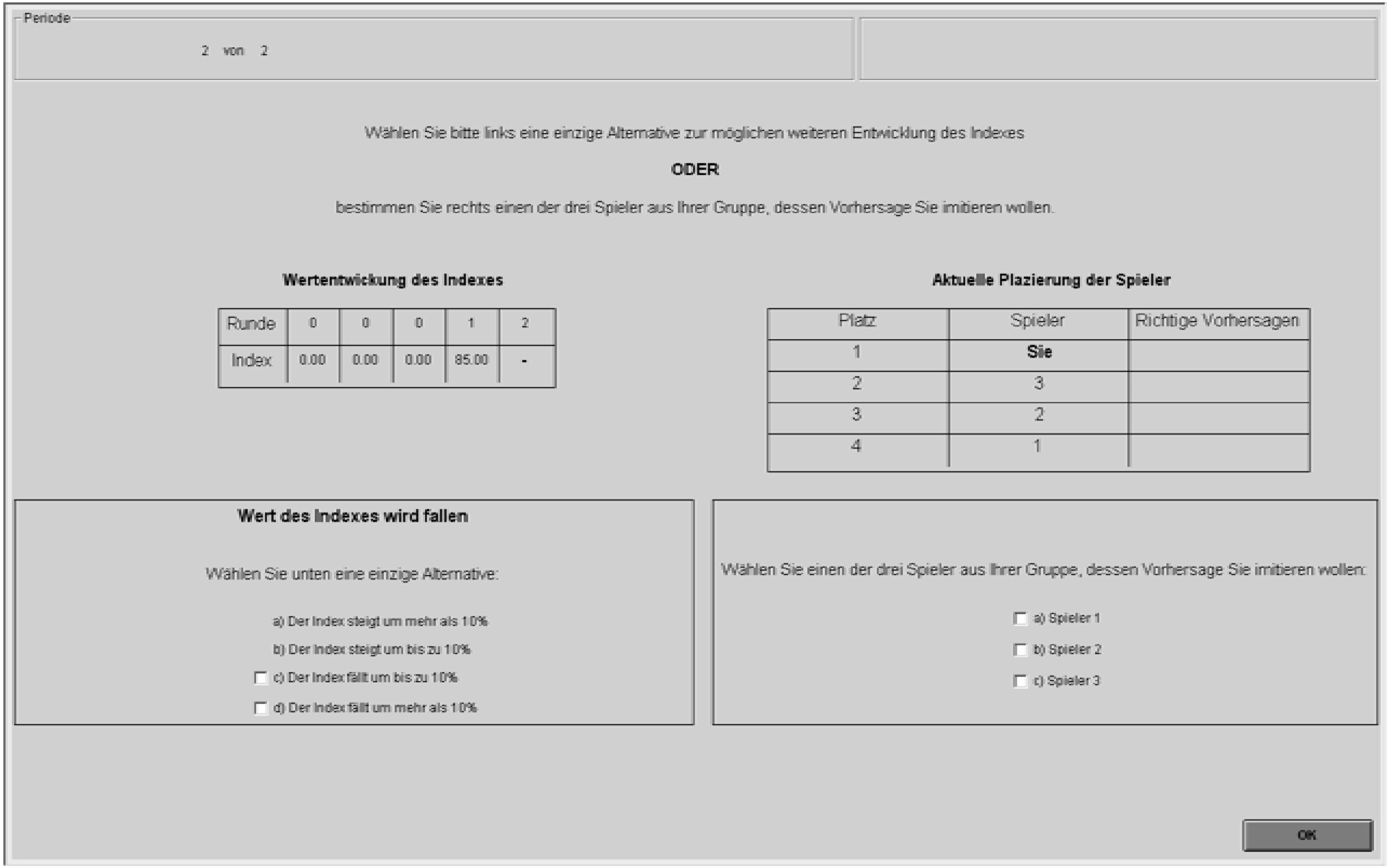


Figure 1, Screenshot, *expert* type player

While making their decision, in every period but the first one, all players (*experts* and *non-experts*) see both a history of the previous index values and a performance ranking of all players in their group, showing the number of correct predictions up to the current period (see Figure 1). The subjects receive 120 points if their prediction is true, otherwise 30 points. Furthermore, imitating subjects incur imitation costs. If two or more players imitate each other, they are all assigned the same alternative, which is determined randomly as above (by a randomly generated number between 0 and 1, each alternative represented by a quartile).

In order to investigate the impact of imitation costs on imitative behavior, two separate treatments are conducted which differ in the level of imitation costs (5 points in the *low-cost* treatment and 25 points in the *high-cost* treatment[16]). In addition, after every period all *non-experts* are asked for an assessment regarding which of their fellow playersthey think is the *expert* in the group. A correct assessment is rewarded with 15 points.[17] At the end of each period, the true value of the index is announced. Furthermore, the performance ranking is updated and displayed again (see Figure 2).

Due to a programming error, other than originally intended and also explained in the instructions to the experiment given to the test subjects, this ranking was not calculated based on the number of correct predictions up to this point, but on the total score including point deductions because of imitation.

In the *low-cost* treatment, the only effect of this additional information is that two players would occasionally switch their place in the ranking shown, although both has apparently made a true prediction (or both a false one, respectively), because one of them gained somewhat less in score due to imitation. In the *high-cost* treatment, the variations in respect to score were much larger than in the *low-cost* treatment, so that in several cases players with less correct predictions were actually shown ranking higher than others with more correct predictions.[18] In at least one instance, a test subject needed substantially more time for his decision after realizing this peculiarity.

We don't believe, however, that this mistake had a significant influence on the results, although it might have led to some confusion. In response, players might have either ignored the information provided by the ranking as unreliable or conceivably even have identified the *experts* more quickly, because these players should never lose points to imitation. Ignoring the ranking would probably mean imitating less, but for the most part in the *high-cost* treatment, which is already designed to be imitation-unfriendly. And an earlier identification of the *experts* would only affect the point in time at which imitation became worthwhile and not the decision whether or not to imitate a given player at all.

[16] In regard to the calculation following Corollary 2, the values for t* are 0.87 and ∞ respectively.
[17] This procedure is applied to ensure that a decision against imitation is not due to a failed identification of the *expert*. This additional reward for guessing the *expert* is set in such a way that on an average over all treatments these correspond more or less to the imitation costs. This balances the advantage of being a *non-expert* somewhat and thus ensures fair conditions in the experiment.
[18] For this to happen, one test subject needed to imitate at least 4 times more often than a second one and to make only one more true prediction than this one during these rounds. Then the 100 point deduction (4 times 25) due to imitation costs would exceed the 90 point gain for the additional true prediction, leading to a lower place in a score-based ranking, despite the higher number of correct predictions. The same could not happen in the *low-cost* treatment, because there it would take at least 19 rounds to accumulate more than 90 points worth of imitation costs (19 times 5 = 95).

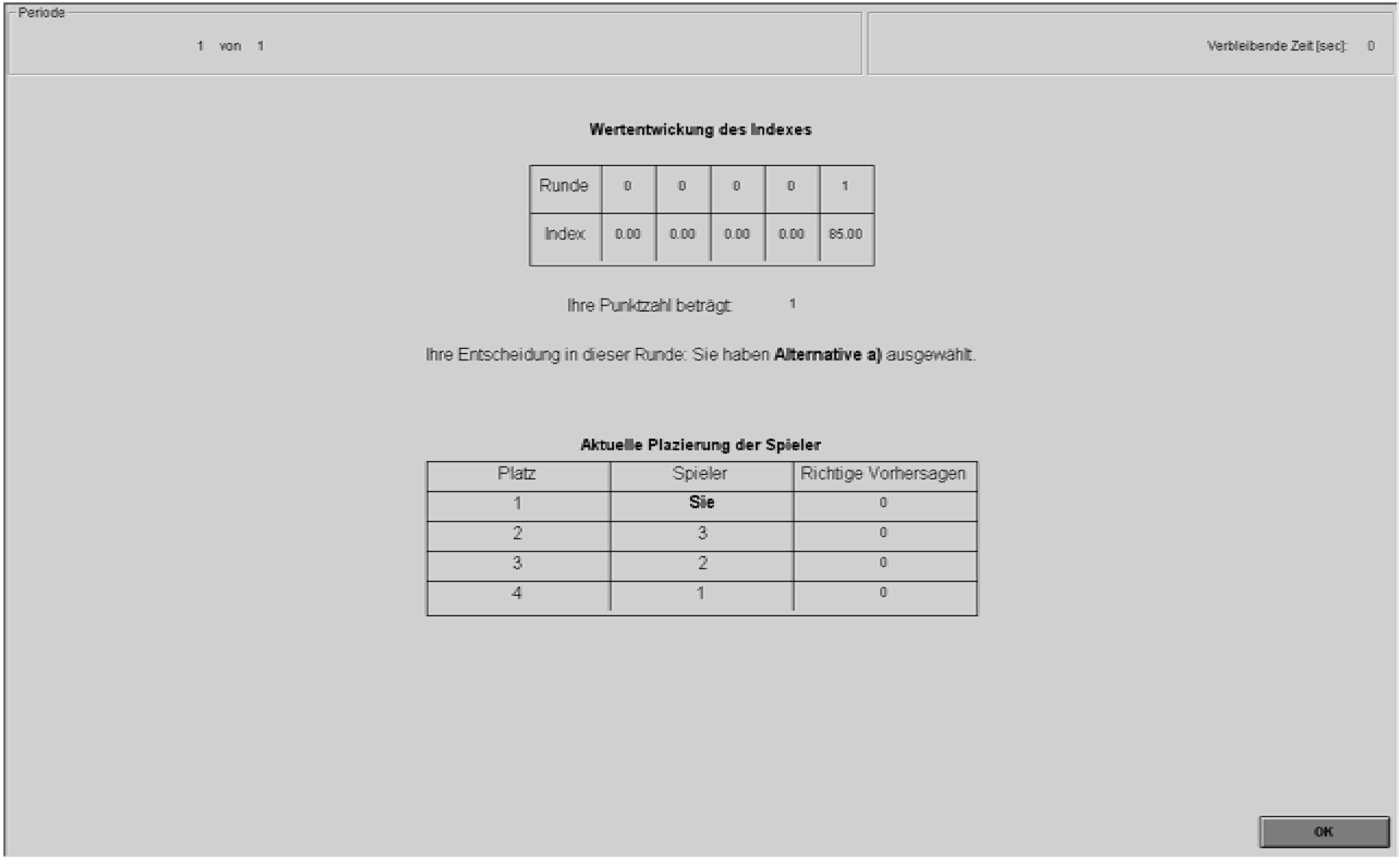


Figure 2, Screenshot at end of period

4.3. Behavioral Predictions

In connection with the theoretical results from Section 3, we derive the following hypotheses for our experimental investigation:

Hypothesis 1: The higher imitation costs are, the later non-experts start to imitate.

Generally, the higher imitation costs c in (3) are, the higher also needs to be the certainty that the imitated player j is the *expert*. The *expert's* identity can be discovered by observing the decisions and results of the other players. In particular, the more rounds have been played, the more information regarding the other players' types can be gathered, and thus, the lower the uncertainty, respectively. In order to test *Hypothesis 1*, we compare the first period of each stage,beginning with which *non-experts* choose imitation, between the *low-cost* and the *high-cost* treatments. As imitation costs in the *low-cost* treatment are significantly lower than in the *high-cost* treatment, it is straightforward to expect that the *non-experts* in the first treatment should start imitating earlier than their counterparts in the latter treatment.

Moreover, given that imitation costs are sufficiently low,to *non-experts*imitation always yields a higher expected profit than making one's own prediction.Therefore, in such a case, fully rational and risk-neutral players of type *non-expert*should always choose imitation. In particular, they should already start to imitate in the very first period, even though the *expert* is still unknown. The parameters in the *low-cost* treatment are set in such a way that a *non-expert*'s expected profit when imitating in the first period of each stage is always strictly higher than when making a prediction on his own, even if no other player imitates (55 vs. 52.5 points). Hence, we derive the following hypothesis:

Hypothesis 2: In the low-cost treatment non-experts always imitate.

However, if imitation costs are too high, imitation is no longer profitable even if the players' types are known with certainty. In this case, imitation should never be observed in even a single period. In the *high-cost* treatment, imitation costs are chosen in such a way that a *non-expert's* expected profit when imitating is always strictly lower than when making a prediction on his own. In particular, this is still true even after the *expert* has been identified (50 vs. 52.5 points). Hence, we derive the following hypothesis:

Hypothesis 3: In the high-cost treatment no one imitates.

Hypotheses 2 and *3* are tested by evaluating the participants' imitative behavior in the related treatments.

In order to construct solid counter-hypotheses, in particular to *Hypotheses 2* and *3*, an expansion of the theoretical model would be required first, which goes beyond the scope of this paper. Instead, we will at this point only give some motivation for possible alternative outcomes of the experiment.

For example, a boundedly rational decision-maker might also decide to imitate in the case of high costs, as long as $w_H > c > c_H$ is true. Moreover, if $w_L > c$ is also true, this tendency should even be stronger, since loss aversion is of no effect then. That's because in this case the players receive a positive payoff in every period, no matter the action they choose. If, conversely, $w_L < c$, then a possibly negative payoff due to a false prediction of the imitated player might act as a deterrent to imitation.

Risk-loving players might also prefer imitation in spite of high costs, because "transferring" the act of forecasting to another player adds to the uncertainty of their decision. Instead of just one lottery (making the correct prediction), imitating players face two lotteries: first they have to choose the correct *expert*, and then the chosen *expert* has to make a correct prediction.

Unfortunately, a distinction between risk-loving and boundedly rational decision-makers can only be investigated in more extensive studies, in which a sufficient number of test subjects with the respective traits are present. That's why in the treatments conducted here, only a short questionnaire will be used to hopefully discover at least certain vague trends regarding these issues.[19]

5. Results

In almost all instances, the *experts* make their predictions themselves, i.e. do not choose to imitate other players. This leads to comparably similar score results for these players in both treatments (average total scores of 701.67 and 696 in the *low-cost* and *high-cost* treatment, respectively). Yet, while most of the *experts* finish first in their respective group, not all of them do: In the *high-cost* treatment, all *experts* finish first (though three of them have to share this rank). Yet, in the *low-cost* treatment, in 3 out of 15 cases the *experts* fare worse than at least one other player in their group; in one additional case the first place is shared. Consequently, in several groups *non-experts* have a hard time identifying "their" *expert*, when this one ranks only

[19] Since this questionnaire didn't generate any conclusive data, its results are omitted from this paper.

second or even worse. However, the knowledge about the distribution of player types in each group, or, more precisely, the fact that there is exactly one *expert* in each group, might have made sure that *experts* weren't bothered too much by their misfortune to actually start to imitate a guaranteed *non-expert*.

The aggregated data show a significant effect of the cost factor on imitative behavior. The *non-experts* are on average less successful than the experts, both in terms of total score and number of correct predictions (average total scores of 561.33 and 460.33 and average number of correct predictions of 3.2 and 2.58 in the *low-cost* and *high-cost* treatment, respectively). So higher imitation costs apparently decrease the predictive success of the *non-experts* and accordingly lower their total score. The additional earnings due to a correct identification of the *expert* are about the same in both treatments (average values of 89.67 and 92.67 points in the *low-cost* and *high-cost* treatment, respectively).

In the *low-cost* treatment, subjects imitate significantly more frequently than in the *high-cost* treatment (see Figure 3). A Wilcoxon rank-sum test comparing both samples on a round-by-round basis (see also Table 1a) by and large refutes the hypothesis of both samples being equal by giving it a probability of being true of $p < 0.05$ in rounds 2 to 6 and 10. While the fact that the difference between the two treatments is not significant in rounds 7 to 9 might vanish in a larger sample size, this appears unlikely for round 1. Instead, in the first round of a prediction game like this the value of the cost factor seems to be indeed irrelevant, perhaps because the subjects underrate the chance of receiving a high score by means of imitating a random other player rather than trusting "their own" luck.

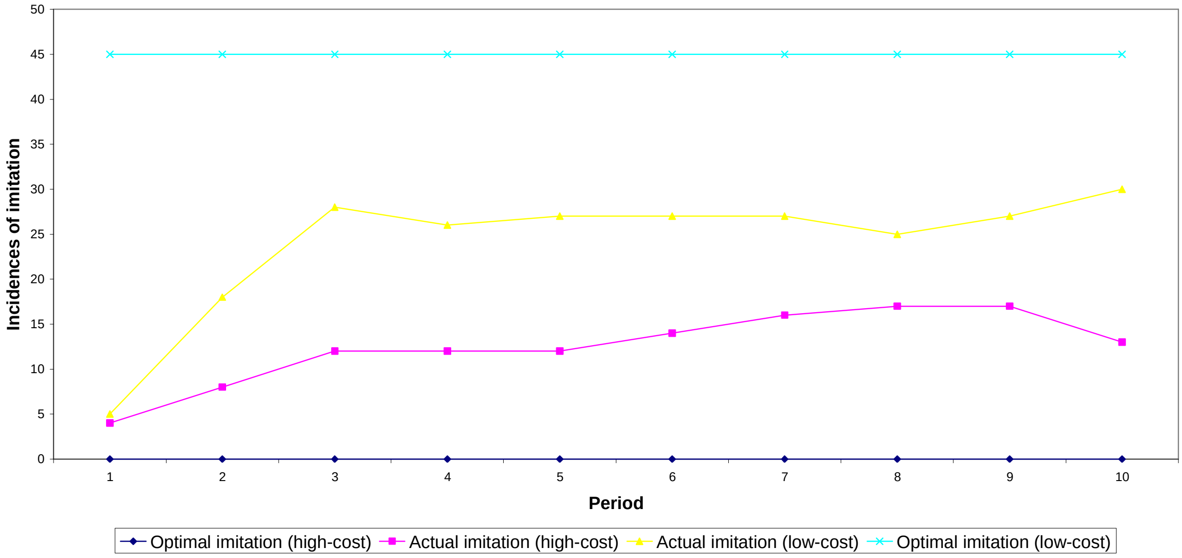


Figure 3, Aggregate frequency of imitation of *non-experts* – meaning of the graphs (from top to bottom): 1) Expectation of the model for *low-cost* treatment, 2) Actual frequency of imitation in *low-cost* treatment, 3) Actual frequency of imitation in *high-cost* treatment, 4) Expectation of the model for *high-cost* treatment. The maximum number of imitation events per period amounts to 45 (15 *non-experts* in 3 stages).

Even clearer is the refutation of *Hypotheses 2* and *3* derived from the model. In the *low-cost* treatment all participants were expected to imitate as early as the first period, in order to maximize their expected payoff (*Hypothesis 2*). Yet in fact, hardly anyone imitated in this period.[20] Conversely, in the *high-cost* treatment subjects imitate significantly more frequently (see Table 1b) than is predicted by *Hypothesis 3*, according to which with these parameters no imitation should be observed at all.

a) Low-cost vs. high-cost

	round 1	round 2	round 3	round 4	round 5	round 6	round 7	round 8	round 9	round 10
z value	0.089	2.280	2.689	2.009	2.408	1.995	1.615	1.162	1.502	2.519
p value(2-sided)	0.9289	0.0226	0.0072	0.0445	0.0160	0.0460	0.1063	0.2453	0.1331	0.0118

b) Hypothesis 3

	round 1	round 2	round 3	round 4	round 5	round 6	round 7	round 8	round 9	round 10
# subjects imitating	4	8	12	12	12	14	16	17	17	13
# subjects not imitating	41	37	33	33	33	31	29	28	28	32
z value(after clustering)	1.730	2.226	2.787	2.787	2.615	2.936	2.780	2.930	2.930	2.614
p value(2-sided)	0.0837	0.0260	0.0053	0.0053	0.0089	0.0033	0.0054	0.0034	0.0034	0.0090

Table 1, Results of data analysis – a) A comparison of the number of imitations in the *low-cost* treatment and the *high-cost* treatment using a Wilcoxon rank-sum test. b) Wilcoxon signed-rank test of *Hypothesis 3* using data from the *high-cost* treatment. In both cases, the z values have been calculated after clustering observations from the same group to arrive at the mentioned number of independent observations (n = 15).

[20] A Wilcoxon signed-rank test almost certainly refutes *Hypothesis 2* (p = 0.0003 of both samples being equal).

Instead, in both treatments an increase from an at first small frequency of imitation to a then more or less constant level can be observed. This development can apparently be explained quite well with players first trying to determine the *expert's* identity and at least in part basing their decision on this knowledge. As Figures 4, top and bottom, show, in both treatments the frequency of imitation moves practically in parallel to the sum of correct *expert* guesses. The bottom part of Figure 4 also makes apparent that the low imitation activity in the *high-cost* treatment can't at all be traced back to these subjects just being worse in discovering the *experts*.

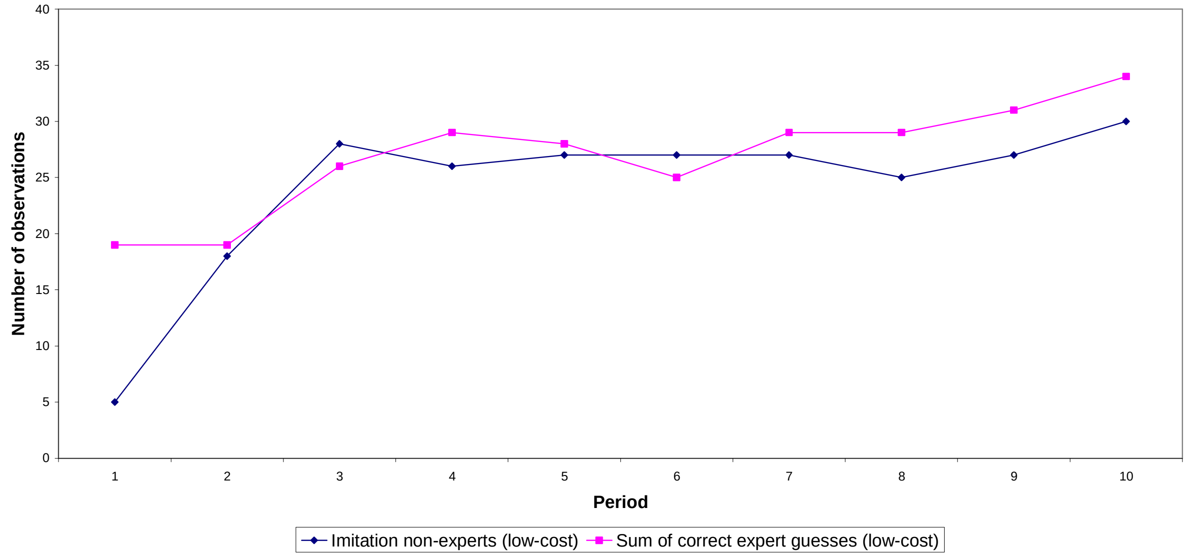
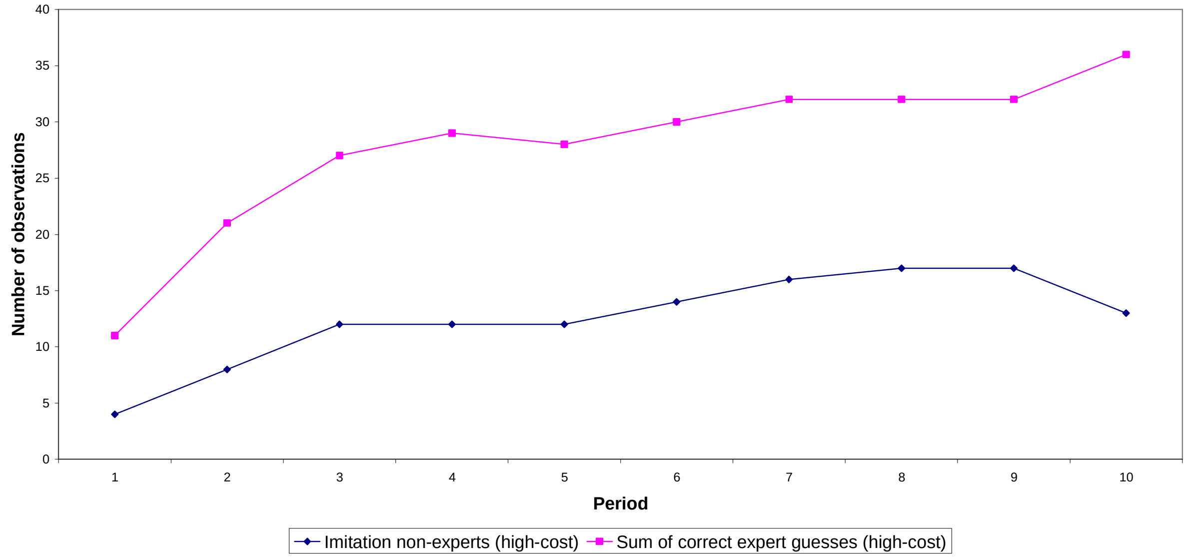


Figure 4, Matching of the *non-experts'* frequency of imitation and the *experts'* recognition level – A comparison of the frequency of imitation (both total and only those actually imitating the *expert*) and the guesses for the *expert's* identity made after every period. The top figure displays results from *low-cost* treatment, the bottom figure results from *high-cost* treatment.



But how can this deviation from the hypotheses be explained? Some additional clues for this are provided by a detailed analysis of the game's course in the individual groups (see Table 2). There it becomes apparent that quite frequently a change of action (no matter whether from "no imitation" to "imitation" or vice versa) was preceded by a false prediction in the previous round, namely in 175 of 562 cases, or resp. 31.1 % of all cases. This percentage is somewhat larger in the first treatment (37.2 % compared to 25.6 % in *low-cost* and *high-cost* treatment resp.) underlining again the influence of the cost factor.[21] What also strikes the eye is that in the *low-cost* treatment players also switch between imitated "*experts*" more frequently (26.0 % vs. 18.7 % of instances of changed actions following false predictions).

	low-cost **treatment**			***high-cost*** **treatment**			**total**		
Opportunity to switch after false prediction	**269**		**100.00%**	**293**		**100.00%**	**562**		**100.00%**
no switch	169		**62.83%**	218		**74.40%**	387		**68.86%**
total number of switches after false prediction	100		**37.17%**	75		**25.60%**	175		**31.14%**
of these from don't imitate to imitate	47	47.00%	17.47%	35	46.67%	11.95%	82	46.86%	14.59%
newly imitated player not on rank 1		*3*	*6.38%*		*3*	*8.57%*		*6*	*7.32%*
of these from imitate to don't imitate	27	27.00%	10.04%	26	34.67%	8.87%	53	30.29%	9.43%
of these from imitate to imitate (other player)	26	26.00%	9.67%	14	18.67%	4.78%	40	22.86%	7.12%
newly imitated player not on rank 1		*9*	*34.62%*		*1*	*7.14%*		*10*	*25.00%*

Table 2 Detailed analysis of imitative behavior– action chosen by *non-experts* in response to wrong prediction in the previous round.

An in part observable regular alternation between "imitation" and "no imitation" might be explained by the phenomenon of "probability matching" (Grant et al., 1951; Erev and Haruvy, forthcoming, p. 9): The players know that the *experts* are right on average only with every second prediction, which is why they keep alternating between "imitation" and "no imitation". Such a behavior might possibly be avoided through a longer duration of play (with more periods and corresponding learning effects). Also relevant might be the "win-stay, lose-shift" decision rule described by Nowak and Sigmund (1993) (see also Erev and Haruvy, forthcoming, p. 49): An action which has led to a false prediction in period $t-1$ thus apparently has good chances of being replaced by another action in period t (as the data suggest), while a successful action is often retained. The recency effect investigated by Estes (1964) (and see also Erev and Haruvy, 2010, p.15) describes a similar behavior, to wit, that the action which has led to the highest payoff in period $t-1$ is (also) used in period t. A more general analysis of "impulsive" decisions in reaction to previous outcomes is given by Avrahami and Kareev (2010).

[21] Hereby, the statistical population only consists of the instances of wrong predictions from periods 1 to 9, as after the 10th period no change of action is possible anymore.

6. Discussion and Conclusion

Along with the (probably to be expected) effects of imitation costs, the investigations conducted here point out another factor able to influence the decision about imitating other players, to wit, the predictive success in the preceding round. The next step should therefore be to find a (preferably rational) explanation of this behavior and to describe it theoretically in an expanded model. Which factors exactly make this behavior better than the "optimal" strategy postulated by the original model? Further experimental investigations might also given some indication of whether this behavior persists in the longer term or whether it is brought to extinction sooner or later by other strategies (e.g. via learning processes).

Moreover, it has to be noted that the possible explanations mentioned in the previous section, i.e. probability matching, the "win-stay, lose-shift" rule, or a recency effect, cannot be tested conclusively by means of the data on hand. However, it is certainly possible to focus the design of future treatments on these effects, e.g. by giving less information to the test subjects at the beginning of the game and instead allowing them to learn any necessary additional information (e.g. payoff distributions or the other players' predictive accuracy) in the course of the experiment, just as is done in the experiments described by Erev and Haruvy (forthcoming). Such a design should preclude probability matching, while at the same time amplifying decision-making based on previous outcomes.

References

Ashiya, M. / Doi, T. (2001): “Herd behavior of Japanese economists”, Journal of Economic Behavior & Organization, 46, 343–346.

Anderson, L. / Holt, C. (1997):“Information Cascades in the Laboratory”, American Economic Review, 87(5), 847-862.

Avrahami, J. / Kareev, Y. (2010): “The Role of Impulses in Shaping Decisions”, Discussion Paper No. 552, The Hebrew University of Jerusalem.

Bikhchandani, S. / Hirshleifer, D. / Welch, I. (1992): “A Theory of Fads, Fashion, Customand Cultural Change as Informational Cascades”, Journal of Political Economy, 100(5), 992-1026.

Bikhchandani, S. / Hirshleifer, D. / Welch, I. (1998): “Learning from the Behavior of Others:Conformity, Fads, and InformationalCascades”, Journal of Economic Perspectives 12(3), 151-170.

Cesa-Bianchi, N. / Lugosi, G. (2006): “Prediction, Learning, and Games”, Cambridge, Cambridge University Press.

Conlisk, J. (1980) “Costly Optimizers versus Cheap Imitators”, Journal of Economic Behavior and Organization, 1(3), 275-293.

DeLong, J. B. / Froomkin, A. M. (2000): “Speculative Microeconomics for Tomorrow’s Economy”, First Monday, 5(2)

DeSantis, A. / Markowsky, G. / Wegman, M. N. (1988): “Learning Probabilistic Prediction Functions”, in: “Proceeding of the 1st Annual Workshop on Computational Learning Theory”, 312-328; Morgan Kaufman, San Mateo, CA.

Drehmann, M. / Oechssler, J. / Roider, A. (2004): “Herding and Contrarian Behavior in Financial Markets – An Internet Experiment”, Discussion Paper No. 7, University of Mannheim, Governance and the Efficiency of Economic Systems (GESY).

Erev, I. / Haruvy, E. (forthcoming): “Learning and the Economics of Small Decisions“, In Kagel, J. H. / Roth, A. E. (forthcoming) (Eds.): “The Handbook of Experimental Economics”, Volume 2, Princeton, Princeton University Press.

Estes, W. K. (1964): “Probability learning”, in: A. W. Melton (ed.), “Categories of human learning”, 89-128, New York, Academic Press.

Fischbacher, U. (2007): “z-Tree: Zurich Toolbox for Ready-made Economic Experiments”, Experimental Economics, 10(2), 171-178.

Gigerenzer, G. / Goldstein, D. G. (1996): “Reasoning the Fast and the Frugal Way: Models of Bounded Rationality”, Psychological Review, 103(4), 650-669.

Graham, J. (1999): “Herding among Investment Newsletters: Theory and Evidence”, Journal of Finance, 54(1), 237–268.

Grant, D. A. / Hake, H. W. / Hornseth, J. P. (1951): “Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement”, Journal of Experimental Psychology, 42, 1-5.

Greiner, B. (2004): “The Online Recruitment System ORSEE 2.0 – A Guide for the Organization of Experiments”, Working Paper Series in Economics 10, University of Cologne, Department of Economics.

Grossman, S. J. /Stiglitz, J. E. (1976): “Information and Competitive Price Systems”, American Economic Review, 66(2), 246-253.

Harabi, N. (1991): “Innovation versus Imitation: Empirical Evidence from Swiss Firms”, MPRA Paper 26214, University Library of Munich Germany

Heemeijer, P. / Hommes, C. / Sonnemans, J. / Tuinstra, J. (2009): “Price Stability and Volatility in Markets with Positive and Negative Expectations Feedback: An Experimental Investigation”, Journal of Economic Dynamics and Control, 33(5), 1052-1072

Hirshleifer, D. / Teoh, S. H. (2003): “Herd Behaviour and Cascading in Capital Markets: a Review andSynthesis”, European Financial Management, 9(1), 25–66

Kirchkamp, O. / Nagel, R. (2003): “No imitation – on local group interaction, learning and reciprocity in prisoner’s dilemma experiments”, Working Paper No. 03-04, SFB 504, University of Mannheim

Laurent R-A (2008): “Product innovation and imitation in a duopoly with differentiation by attributes”, PSE Working Papers 2008-04, Paris School of Economics (Ecole normale supérieure)

Mansfield, E. / Schwartz, M. /Wagner, S. (1981): “Imitation Costs and Patents: An Empirical Study”, Economic Journal, 91, 907-918

Nowak, M. / Sigmund, K. (1993): “A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner’s Dilemma game”, Nature, 365, 56-58.

Offerman, T. / Potters, J. / Sonnemans, J. (2002): “Imitation and Belief Learning in an Oligopoly Experiment”, Review of Economic Studies, 69(4), 973-997

Plott, C. R. / Sunder, S. (1982): “Efficiency of Experimental Security Markets with Insider Information: An Application of Rational-Expectations Models”, Journal of Political Economy, 90(4), 663-698.

Schlag, K. H. (1998): “Why Imitate, and if so, How? A Bounded Rational Approach to Multi-Armed Bandits”, Journal of Economic Theory 78(1), 130-156

Schurz, G. (2008): “The Meta-inductivist’s Winning Strategy in the Prediction Game: A New Approach to Hume’s Problem”, Philosophy of Science 75, 278-305.

Vovk, V. (1998): “A game of prediction with expert advice”, Journal of Computer and System Sciences, 56(2), 153-173.

Welch, I. (1989): “Seasoned Offerings, Imitation Costs, and the Underpricing of Initial Public Offerings”, Journal of Finance, 44(2), 421-449

Welch, I. (1992): “Sequential Sales, Learning, and Cascades”, Journal of Finance, 47(2), 695-732.

Wolfers, J. / Zitzewitz, E. (2004): “Prediction Markets”, Journal of Economic Perspectives, 18(2), 107-126.

Appendix 1 – Instructions for the test subjects[22]

Welcome to this experiment!

You will now participate in a scientific experiment. Please read the following instructions carefully. Here you are told everything you need to know for the participation in this experiment. Please also note the following:

- Today's experiment consists of a **preliminary experiment** and a **main experiment**
- From this point on as well as during the entire experiment **no communication** is permitted. If you have any questions, please raise your hand.
- Absolutely all **decisions occur anonymously**, i.e. none of the other participants gets to know the identity of the person who has made a particular decision.
- For arriving on time you receive an amount of **2.50 euros** (show-up fee). In the course of the experiment you can earn an additional amount of money. Both amounts will be paid to you at the end of the experiment in cash.
- The **payment likewise occurs anonymously**, i.e. no participant is told how high the payment of any other participant is.

Procedure of the preliminary experiment

- The preliminary experiment takes three minutes. Your task is to determine the correct number of sevens in blocks of random numbers and to handle as many number blocks as possible within the three minutes. In doing so, the following applies:
  - Every correctly handled number block is worth 2 points.
  - Every wronglyhandled number block is worth 0 points.
- At the end of the preliminary experiment you are told:
  - the number of correctly and wronglyhandled number blocks
  - your score
- For every point of your point of your score in this preliminary experiment you receive 10 cents.

Procedure of the main experiment

- In total, the main experiment consists of **three** stages with 10 rounds respectively.
- Your task in this main experiment is to give an as-precise-as-possible prediction of the future development of an index.
- The value of the index is generated **randomly**.
- In every stage you will play together in a group with three other players. Please note that in every stage you will play together with different players. It is made sure that during the entire experiment you will never play together with the same player in different stages.
- There are two types of players with a different state of information among all players.

[22] Translated from German. Here only the version for the treatment with high imitation costs is given, as the text is virtually identical except for the value of this cost factor.

- At the beginning of the main experiment, the five best players from the pre-experiment (with the five highest scores) are assigned the role of "**expert**". During the entire experiment this type assignment does not change. You will be informed about your own type at the beginning of the main experiment. Please note that in every group there is **exactly one** "expert".
- In addition, all players are told the starting value of an index (100).
- At the beginning of each round, the experts receive a private message about the direction in which the index will deflect ("rises" or resp. "falls"). All other players don't receive any information.
- Afterwards all players make a prediction about the future development of the index. In doing so, you can choose **one of two possible actions**:

  (1) **Either** you choose an alternative from a list of alternatives for the possible future development of the index. Every alternative obtains with the same probability:

    a) The index rises by more than 10%.

    b) The index rises by up to 10%.

    c) The index falls by up to 10%.

    d) The index falls by more than 10%.

  (2) **Or** you designate one of the three players from your group, whose prediction you want to imitate. Whatever this player chooses also applies for you as a prediction. If two or more players from a group mutually imitate themselves (e.g. Player 1 imitates Player 2, Player 2 imitates Player 3, andPlayer 3 imitates Player 1), all players are assigned the same randomly determined prediction. Please note that the imitation of another player involves **costs of 25 points[23] each time**, which are deducted from your score in the respective round.
- The game will only be continued, when all players have made their decisions.
- Your payoff in every round occurs according to the accuracy of your predictions. You receive:
  - **120 points, if the prediction is true**
  - **30 points, else**
- In addition, in each round you – unless your type is "expert" – will be asked for the assessment, which of your fellow playersis the "expert" in your group. If your assessment is correct, you receive **15 points, otherwise zero points**. You will be told the points you receive for your assessment **only at the end of each stage**.
- At the end of each round the new value of the index is announced. Furthermore, a performance ranking with the number of the correct predictions of all players in your group is published.
- Please note that **solely the number of correct predictions** is relevant for the performance ranking (and not the respective score).[24]
- After the first 10 (or resp. 20) rounds the second (or resp. third) stage begins, in which you will play together with three other players in a new group.

[23] These are the costs for the high-cost treatment. The costs were given as "5 points" in the low-cost treatment (see footnote above).

[24] Due to a programming mistake this was actually implemented exactly the other way around. See Section 4.2 for more information on this problem and possible effects on the players' behavior.

- Please note that at the end of the experiment **three rounds from each stage** respectively are chosen randomly, which then determine your payment. That is, only those points, which you have gained in the chosen rounds, are of relevance to your payment.
- For every point you have gained in the chosenrounds you receive 1.7 cents.

NOTE

Please think hard about any one of your decisions, as these will possibly determine the amount of your payment at the end of the experiment.

PAYMENT

The amount earned by you in the preliminary and main experiment will be paid to you in cash at the end of the entire experiment.

END OF THE EXPERIMENT

- At the end of the experiment we will ask you to carefully fill in a short questionnaire. Any personal data provided by you will be treated confidentially and used only for this single scientific purpose. Afterwards, all personal data are deleted.
- Please remain seated after filling in the questionnaire until we call up your booth number. Please bring this instruction and your booth number with you to the front. Only then the payment of your game result can occur.

Thanks for your participation and good luck!

Appendix 2 – Proofs of the propositions from Section 3

Proof of Proposition 1

Player i imitates a given player j only if this generates a higher expected payoff than making his own prediction, that is:

$$E\left[\pi^{imitate\ j}\right] \geq E\left[\pi^{no\ imitation}\right] \tag{A.1}$$

In the case assumed here, with a high score value w_H for a correct prediction and a low score value w_L otherwise, this calculates as follows:

$$\hat{P}\left[p_{j,t} = y_t\right] \cdot w_H + \left(1 - \hat{P}\left[p_{j,t} = y_t\right]\right) \cdot w_L - c \geq P\left[p_{i,t} = y_t\right] \cdot w_H + \left(1 - P\left[p_{i,t} = y_t\right]\right) \cdot w_L$$
(A.2)

Some transformations yield (3) ■

Proof of Proposition 2

1. We look at the extreme case, in which no other player imitates, so that indirect imitation does not contribute to the utility from choosing imitation. Consequently, in period 1 only the information known prior to the game's start is relevant to the decision. By means of the law of total probability, the left-hand side of (3) can then be written as follows, using the knowledge about the player types' probability of predicting correctly:

$$\hat{P}\left[p_{j,t} = y_t\right] = \hat{P}[j = exp] \cdot P\left[p_{j,t} = y_t | j = exp\right] + \left(1 - \hat{P}[j = exp]\right) \cdot P\left[p_{j,t} = y_t | j \neq exp\right]$$
(A.3)

The probability of *j* being the *expert* thus corresponds to the share of *experts* in the group, meaning that (3) can be modified as follows (after first substituting (A.3) for the left-hand side of the inequality):

$$\frac{\#experts}{\#players - 1} \cdot P\left[p_{j,t} = y_t | j = exp\right] + \left(1 - \frac{\#experts}{\#players - 1}\right) \cdot P\left[p_{j,t} = y_t | j \neq exp\right]$$

$$\geq P\left[p_{i,t} = y_t\right] + \frac{c}{w_H - w_L} \tag{A.4}$$

This inequality is always true for $c = 0$, as long as player i isn't the *expert*, meaning that imitation without a cost factor is always worthwhile, even if the *expert* hasn't been identified yet and no other player "helps" via indirect imitation. However, for certain parameters of the prediction game one can also calculate a lower cost bound $c_L > 0$, being of particular interest for an experimental investigation, since in this case imitation is just as worthwhile in spite of low costs as without imitation costs.

2. Follows directly from Proposition 1.

3. In the worst case, imitation costs are so high that inequality (3)in Proposition 1 isn't fulfilled, even if an *expert* has been identified with certainty. In this case, indirect imitation is of no importance anymore and (3) can be written as follows, replacing the subjective belief with an objective probability:

$$P\left[p_{expert,t} = y_t\right] \geq P\left[p_{i,t} = y_t\right] + \frac{c}{w_H - w_L} \tag{A.5}$$

Above an upper cost bound c_H, imitation is not worthwhile under any circumstance from the perspective of a rational risk-neutral decision-maker. ■

Proof of Corollary 2

After again substituting term (7) into (3), let $P[j = exp]$ approach 2, yielding (9). Suppose (9) is not fulfilled. This is only the case if costs are prohibitively high or j is not of type *expert*, both of which were assumed to be untrue. ■