WZB

Wissenschaftszentrum Berlin
für Sozialforschung

Daniel Friedman, Steffen Huck,
Ryan Oprea, and Simon Weidenholzer

From Imitation to Collusion: Long-run Learning in a Low-Information Environment

Discussion Paper

SP II 2012–301r

August 2012 (revised October 2013)

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Affiliation of the authors:

Daniel Friedman
University of California Santa Cruz

Steffen Huck
WZB and University College London

Ryan Oprea
University of California Santa Barbara

Simon Weidenholzer
University of Essex

From Imitation to Collusion: Long-run Learning in a Low-Information Environment*

Abstract

We explore the stability of imitation in a 1,200-period experimental Cournot game where subjects do not know the payoff function but see the output quantities and payoffs of each oligopolist after every period. In line with theoretical predictions and previous experimental findings, our oligopolies reach highly competitive levels within 50 periods. However, already after 100 periods quantities start to drop and, eventually fall deep into collusive territory without pausing at the Nash equilibrium. Our results demonstrate how groups of subjects can learn their way out of dysfunctional heuristics, and thus suggest the need for a new theory of how cooperation emerges.

Zusammenfassung

Wir untersuchen die Stabilität von Imitationsverhalten in einem experimentellen Cournot-Oligopol mit 1.200 Perioden, in denen die Teilnehmer zwar keine Informationen über die Gewinnfunktion haben, aber nach jeder Periode die Output-Mengen und den Gewinn jedes Oligopolisten sehen können. Im Einklang mit theoretischen Vorhersagen und Befunden vorheriger Experimente, erreichen unsere Oligopole hoch-kompetitive Levels innerhalb von 50 Perioden. Aber bereits nach 100 Perioden fangen die Output-Mengen an zu fallen, im weiteren Verlauf sogar bis tief in kollusive Bereiche, ohne am Nash-Gleichgewicht zu pausieren. Unsere Befunde veranschaulichen, wie Gruppen von Personen ihren Weg aus dysfunktionalen Heuristiken heraus finden können, und legen daher die Notwendigkeit einer neuen Theorie über die Entstehung von Kooperation nahe.

Keywords: Cournot oligopoly, imitation, learning dynamics, cooperation.

JEL Codes: C73, C91, D43

* We are grateful to the National Science Foundation for support under grant SES-0925039, James Pettit for programming the software, Luba Petersen for assistance in running the sessions, and conference participants at the 2012 American Economic Association Meetings, the 6th Nordic Conference on Behavioral and Experimental Economics in Lund 2011, the 2012 Santa Barbara Conference on Experimental and Behavioral Economics, seminar audiences at Alicante, Cambridge, Edinburgh, Essex, Frankfurt, Carlos III, CERGE-EI, Cologne, Nottingham, Pompeu Fabra, Queen Mary, Stirling, UCL, Baruch College, the University of Zürich and the WZB. Weidenholzer acknowledges financial support from the Vienna Science and Technology Fund (WWTF) under project fund MA 09-017.

1 Introduction

Imitation is an attractive heuristic when players have little information about the strategic environment but can observe others' choices and success. Compared to popular learning models that focus on own payoffs, imitation makes more comprehensive use of available information — but not necessarily *better* use, as first shown by Vega-Redondo (1997) for the case of Cournot games where imitation generates the perfectly competitive Walrasian outcome. Within the broad class of aggregative games (Alós-Ferrer and Ania 2005), Cournot games are, similarly to public good games or common pool resource games, notable for their tension between social efficiency and individual optimization.[1] The efficient collusive profile contrasts with the less efficient Nash equilibrium, and with the still less efficient fully competitive or Walrasian outcome where price is equal to marginal cost. Vega-Redondo showed that imitating quantity choices of the more profitable players leads precisely to that least efficient profile, where in linear settings economic profits are zero.

Of course, this unfortunate outcome arises from a blind spot in the imitation heuristic — it ignores the fact that prices fall with greater quantities. Nevertheless, the heuristic has been quite descriptive of laboratory behavior in low-information environments where players observe other players' quantity choices and profits but not the underlying payoff function. Most of these studies (including, among others, Huck, Normann, and Oechssler 1999, Offerman, Potters, and Sonnemans 2002, Apesteguía, Huck, and Oechssler 2007 or Apesteguía, Huck, Oechssler, and Weidenholzer (2010)), feature what has been considered "long horizon" repeated interaction of around 50 periods.

Our point of departure is to examine a much longer horizon. We employ the new ConG software (Pettit, Friedman, Kephart, and Oprea 2012) which allows for periods to be so short that human subjects perceive action as taking place in continuous time. Here we instead use the software to implement discrete 4-second periods — rather short by recent standards, but perceived by our subjects as comfortable stop-action in discrete time. This enables us to increase the number of periods to 1,200.

The results are dramatic — what looked like stable long-run behavior in earlier studies turns out to be transient. In the first 50 periods of our experiment we replicate the very competitive outcomes observed in previous studies. However, soon thereafter the trend reverses and quantity choices start to drop. Quantities often approach the Cournot-Nash level after 100 periods but they do not halt there, or even pause. Rather they continue to drop until they reach almost fully collusive levels in

[1] Since consumers are not considered players in such games, social efficiency refers to the players' joint payoff maximum at the cartel profile.

duopolies and reach, on average, deep into collusive territory in triopolies. These collusive levels are reached despite our use of a hyperbolic demand function that creates much stronger incentives to deviate from collusion than the linear demand functions seen in most Cournot oligopoly experiments.

The primary contribution of the present paper is to document this transition in outcomes — from very competitive to collusive. The transition demonstrates how players can learn to abandon dysfunctional heuristics and find better ways to reconcile group interest with self interest. Interestingly, cooperation does not seem to be supported by Nash reversion or similar strategies; indeed the evidence suggests that our subjects never even learn the best response function. Instead, they appear to gravitate to alternative heuristics that align the players' incentives and enable a form of punishment and forgiveness. We present some suggestive evidence on how these heuristics operate, but apparently new theory must be developed to complete our understanding of long-run learning and the emergence of cooperation.

The literature on oligopoly experiments in the laboratory is too vast to be surveyed here. Besides the papers already mentioned on "long run" behavior in low information Cournot games, we perhaps should note very early work on posted price oligopoly by Friedman and Hoggatt (1980) and Alger (1987). Some of their oligopolies lasted over 100 periods, but the results were inconclusive and hard to compare to Cournot oligopolies.

Section 2 describes the basic theory relevant to our investigation, and Section 3 lays out our laboratory procedures including the user interface as well as the treatments and matching protocol. Section 4 summarizes aggregate results. It shows that initially play becomes very competitive, consistent with Vega-Redondo's imitation model, but that eventually behavior changes and overall profits rise towards collusive levels. Section 5 analyzes individual level behavior, and finds that while subjects begin by imitating rivals with higher payoffs, they eventually find different heuristics that lead them toward and then support cooperation. Although we observe clear end-game effects that demonstrate that subjects are aware of last rounds, and other evidence shows that they are aware of profitable deviations from cooperation, we find that our subjects do not understand Nash reversion. Indeed, they never learn crucial parts of the best-reply correspondence of the stage game, let alone its Nash equilibrium. Nevertheless, subjects enjoy ever-longer spells of collusive play with more effective and shorter "punishment" episodes. Section 6 discusses our findings, which do not seem to vindicate standard repeated game theory. We do note connections to several existing approaches, including Tit for Tat and Win-Continue, Lose-Reverse, as well as Conjectural Variations and standard learning models. Although suggestive, none of these theoretical approaches seems able to explain the emergence of cooperation that we see in our data. Section 7 summarizes our contribution and points out promising directions for future research.

Appendix A contains supplemental data analysis, Appendix B reproduces instructions to subjects. Appendix C collects supplementary mathematical derivations.

2 Basic Theory

We study a repeated Cournot game played by a fixed finite number $n \geq 2$ of strategically identical players with constant marginal cost $c \geq 0$. Each period, each player i chooses a quantity x_i in a finite interval $[x_L, x_U]$. Price P is a decreasing function of the aggregate quantity $X = \sum_{j=1}^{n} x_j$, and player i's profit that period is $\pi_i = a + (P(X) - c)x_i$, including an exogenous additive constant a that captures benefits from other activities net of fixed cost. Our experiment uses $n = 2$ or 3, the interval $[x_L, x_U] = [0.1, \frac{12}{n}]$, $a = c = 10$, and unit elastic demand with $XP(X) = 120$, so

$$\pi_i(x_i, x_{-i}) = 10 + \left(\frac{120}{\sum_j x_j} - 10 \right) x_i. \tag{1}$$

Maximal quantity choice $x_i = x_U = \frac{12}{n}$ by every player i yields the minimal price $P = \frac{120}{nx_U} = 10$ equal to marginal cost. Associated minimal profits are $\pi_i^{PCW} = a + 0 = 10$ for every player. We refer to this action profile as the perfectly competitive Walrasian outcome (PCW).

At the other extreme of the action space, minimal quantity choice $x_i = x_L = 0.1$ by every player i yields the maximal price $P = \frac{120}{nx_L} = 1200/n$ and indeed maximal total profits $n\pi_i^{JPM} = 9n + 120$. We call this profile the joint profit maximum (JPM).

The best response of player i to $X_{-i} = \sum_{j \neq i} x_j$ is the unique solution $x_i^* = b(X_{-i}) \in [0.1, \frac{12}{n}]$ to the first-order condition

$$0 = \frac{\partial \pi_i}{\partial x_i} = \frac{120}{x_i + X_{-i}} - 10 - \frac{120 x_i}{(x_i + X_{-i})^2}, \tag{2}$$

and is given by

$$b(X_{-i}) = 2\sqrt{3X_{-i}} - X_{-i}. \tag{3}$$

Imposing the relevant symmetry condition $x_i + X_{-i} = nx_i$ in (2) and solving for x_i, we obtain the Cournot-Nash equilibrium profile (CNE) as $x_i^{CNE} = \frac{12(n-1)}{n^2}$. The corresponding price is $P^{CNE} = \frac{10n}{n-1}$, and the resulting equilibrium profit for each player is $\pi_i^{CNE} = a + \frac{10}{n-1} \cdot x_i^{CNE} = 10 + \frac{120}{n^2}$. Table 1 summarizes these static predictions for the duopoly $(n = 2)$ and triopoly $(n = 3)$ cases.

Compared to a linear demand specification, the unit elastic demand embodied in payoff function (1) has three important advantages for experimental work. First, as shown in Table 1, it gives a

Table 1: Static outcomes for payoff function (1)

	Duopoly			Triopoly		
	x_i	P	π_i	x_i	P	π_i
JPM	0.1	600	69	0.1	400	49
CNE	3	20	40	$2.6\bar{6}$	15	$23.\bar{3}$
PCW	6	10	10	4	10	10

clean separation between the three static outcomes of interest. Second, it creates a much stronger temptation to defect at the JPM. Finally, for $n < 6$, the payoff function is not as flat around the best response. See Appendix C for details on the limitations arising from a linear specification of the demand function.

When players have little a priori information but can observe their competitors' actions and profits they may resort to imitation; in particular they might simply copy the action of the player who was most successful in the previous period. This "imitate-the-best" heuristic was introduced into the theory literature by Vega-Redondo (1997). Vega-Redondo's model also allows agents from time to time to make mistakes and choose a quantity different from the one prescribed by the imitation rule. He shows that as the error rate goes to zero, the limit of the dynamic process spends almost all time in the PCW profile.

Under imitation, the PCW profile is robust in several senses. It can be reached rapidly: if a single player ever chooses x_U, mistakenly or otherwise, she will immediately be imitated by all players thus achieving the PCW the next period, absent other mistakes. Moreover, once the PCW is achieved, single deviations will never be imitated under Vega-Redondo's (1997) rule. Apesteguía, Huck, and Oechssler (2007) show that PCW is also the unique stochastically stable outcome for a wide range of other imitation rules, including Schlag's (1998) proportional imitation rule, and the imitate-the-best-average rule of Eshel, Samuelson, and Shaked (1998).

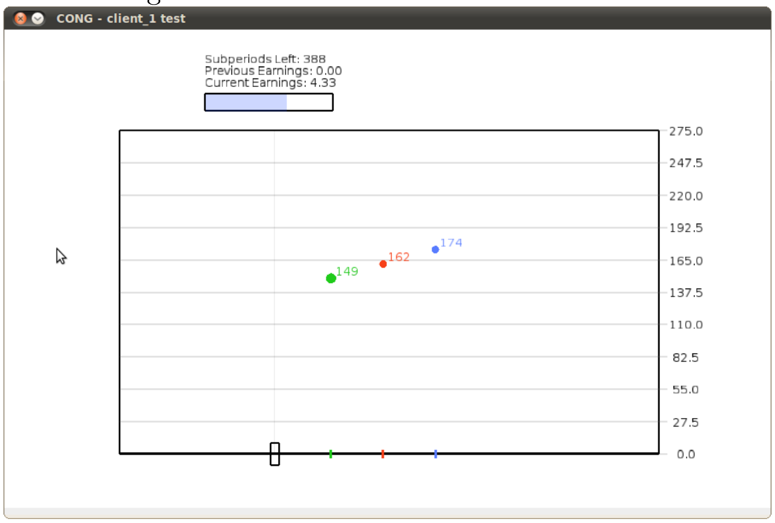
Alós-Ferrer and Ania (2005) show that stochastic stability of the PCW outcome follows also from the fact that it is a strict finite-population ESS in the sense of Schaffer (1988). That is, unilateral deviations from the PCW profile $(x_U, x_U, \ldots, x_U)$ satisfy the strict payoff inequality

$$\pi_i(x' | \overbrace{x_U, \ldots, x_U}^{n-1}) < \pi_i(x_U | x', \overbrace{x_U, \ldots, x_U}^{n-2})$$

for all $x' \neq x_U$, i.e., the deviator earns a lower payoff than the non-deviators.

The intuition behind these stability results is simple. All firms in Cournot oligopoly face the same price, and as long as that price is above marginal cost, the most profitable firm is the one with

Figure 1: Screenshot from ConG software.



the largest quantity. Imitation will therefore lead firms to increase quantities, driving price down to marginal cost. (Price below marginal cost is not possible with our restricted strategy space, but even if it were, the firm with the smallest quantity would be the most profitable, and once again imitation would drive the price back towards marginal cost.) In our game the PCW is the unique profile where price equals marginal cost. At any other feasible profile, a deviation towards the PCW choice x_U will give the deviator higher profits than the non-deviators. Moreover, as just noted in the ESS discussion, any single deviation from PCW earns the deviator smaller profit than the non-deviators. Thus the PCW outcome is the only stochastically stable state, and is relatively robust to mistakes.

Vega-Redondo's theory has been highly predictive in previous experimental studies. Subjects seem attracted to the imitation heuristic and experimental markets do swiftly become very competitive and stay so for the time horizons previously explored. In the process, subjects persistently earn sub-Nash payoffs — though intuitive and simple, the imitation heuristic is highly destructive to earnings. The primary question we ask in the present study is whether, given enough experience, subjects will learn to abandon the imitation heuristic and thereby collectively escape from the low earnings PCW trap.

3 Laboratory procedures

The experiment used new ConG software (Pettit et al. 2012), with the user interface illustrated in Figure 1. Three key features allow us to run hundreds of Cournot periods in a single session.

- An intuitive graphical interface displays previous-period actions and payoffs, conveying key feedback information in a glance. Color-coded tick marks on the x-axis show each subject's previous-period quantity choice, e.g., the subject's own choice is shown in green. The y-axis measures profit, so the heights of color-matched dots show everyone's previous-period profits; exact amounts can be read from small font text next to each dot.

- Subjects make quantity choices by simply clicking on the screen, or dragging the hollow-box slider at the bottom of the screen. The set of available quantities is nearly continuous, with a granularity of less than 0.007 units over the interval $[0.1, 6]$ in the Duopoly treatment and $[0.1, 4]$ in the Triopoly treatment. A subject who wants to retain the current action into next period can do so simply by not clicking or dragging.

- Periods are time limited at four seconds. A timer bar above the quantity/profit graph fills in over the course of the period; once it is filled the period is over. During the period each subject can adjust her action as often as she likes; the payoff-relevant action is that seen when the period ends. Immediately thereafter subjects see the actions and resulting payoffs achieved in that period by themselves and their fellow oligopolists.

The four second time limit was shown in pilots to steer safely between the twin pitfalls of time pressure and boredom. Subjects did not seem hurried or frantic during game play and, in informal post-experiment interviews, expressed comfort with the pacing of the game. We believe that this comfort arose from the highly visual graphical interface, which allows fine adjustment of actions in a single click and information dissemination in a glance. The default carry-over of previous actions also allowed subjects, at very low cost, to stand still for several periods while thinking about their decisions, further reducing time pressure.

We will see in section 4 that behavior in the first 50 periods of our experiment is very similar to that seen in previous experiments, reassuring us that our new design features do not drastically reshape behavior. The new features do, however, allow us to run 1,200 periods in less than two hours.

We employed 72 subjects in six sessions of twelve subjects each at the LEEPS laboratory at the University of California, Santa Cruz in April 2011. In half of the sessions we matched subjects exclusively into duopolies and in the other half into triopolies, i.e. we ran two treatments using a completely between-subject design. Our matching algorithm grouped subjects into independent "silos" of six subjects each. Subjects interacted only with subjects in their own silos, thereby giving us six completely independent groups in each treatment. Each 1,200 period session is divided into three 400 period blocks. At the beginning of each block, subjects are rematched to new counterparts in their silo, and no subject interacts more than once with the same counterpart(s).

Because our focus is on adaptation to low-information environments, we told subjects very little about their payoff functions. Using clear but non-technical language, we told them only that the functions were symmetric, time-invariant and determined uniquely by the [quantity] choices of the group members. Subjects were students from all majors and recruited online via ORSEE (Greiner 2004). Instructions, read aloud to the subjects at the beginning of the session, are reproduced in Appendix B. Subjects were paid their average earnings in each of the three blocks at the rate of 12 US cents per point in Duopoly and 18 cents in Triopoly. We paid an additional a show-up fee of $5. On average, sessions lasted just under two hours and subjects earned $21.00.

4 Aggregate results

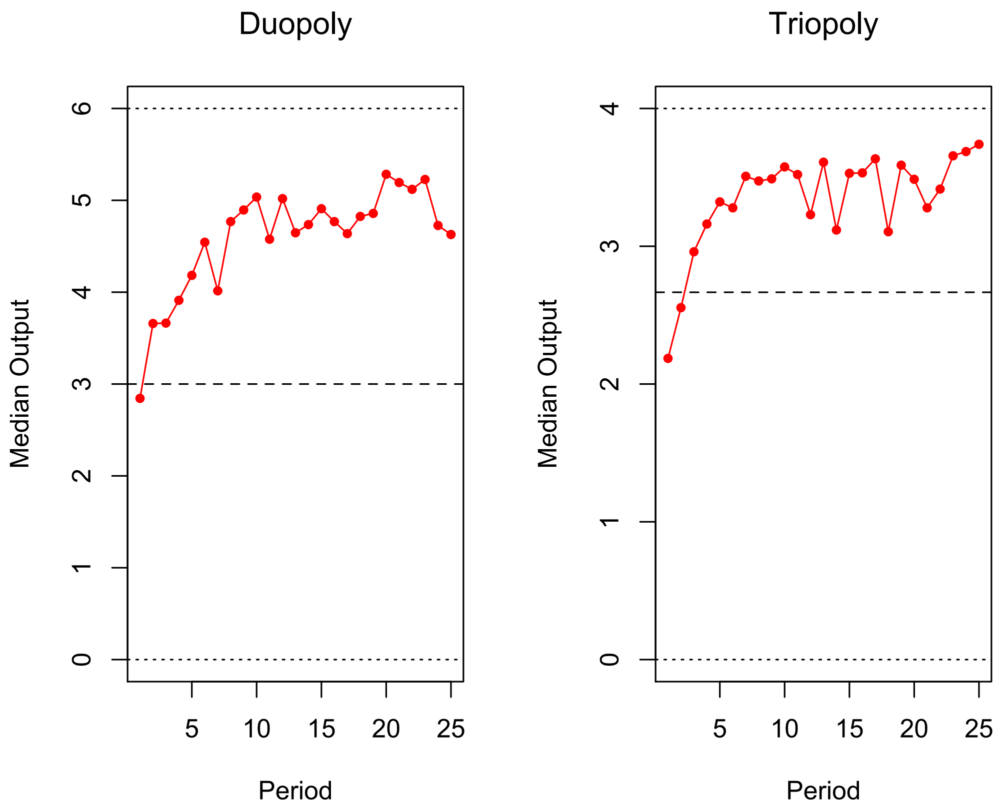
We first examine the early periods of each session to see whether there are any obvious qualitative differences from earlier Cournot experiments in low-information environments. The left-hand panel of Figure 2 plots median quantities from the Duopoly treatments while the right-hand panel does the same for Triopoly in the first 25 periods of the experiment. In Appendix A we plot the evolution of median profits in the same manner.

Markets become very competitive within just a few periods and settle into the competitive region between Cournot-Nash (CNE) and Walras (PCW). This is not only true for the overall medians but also for every single observed oligopoly in both treatments. There are slight differences between duopolies and triopolies with the latter being closer on average to PCW than the former. Indeed, in some triopolies price is equal to marginal cost for sustained periods of time.

To document the initial rise in quantities, note that median quantities increase from the first to the 25th period for each of the six independent matching groups ("silos") in each treatment. The increase is statistically significant at the one percent level in both duopoly and triopoly according to a paired Wilcoxon Signed-Rank test. Over the next 25 periods, median quantities continue to fluctuate in the competitive region above CNE with little sign of systematic trend.

Thus, over the first 50 periods we see essentially the same behavior as in earlier studies. This is despite the fact that in those studies it took over an hour to run 50 periods, versus a little over 3 minutes in our experiment.

Figure 2: Median quantities in the first 25 periods

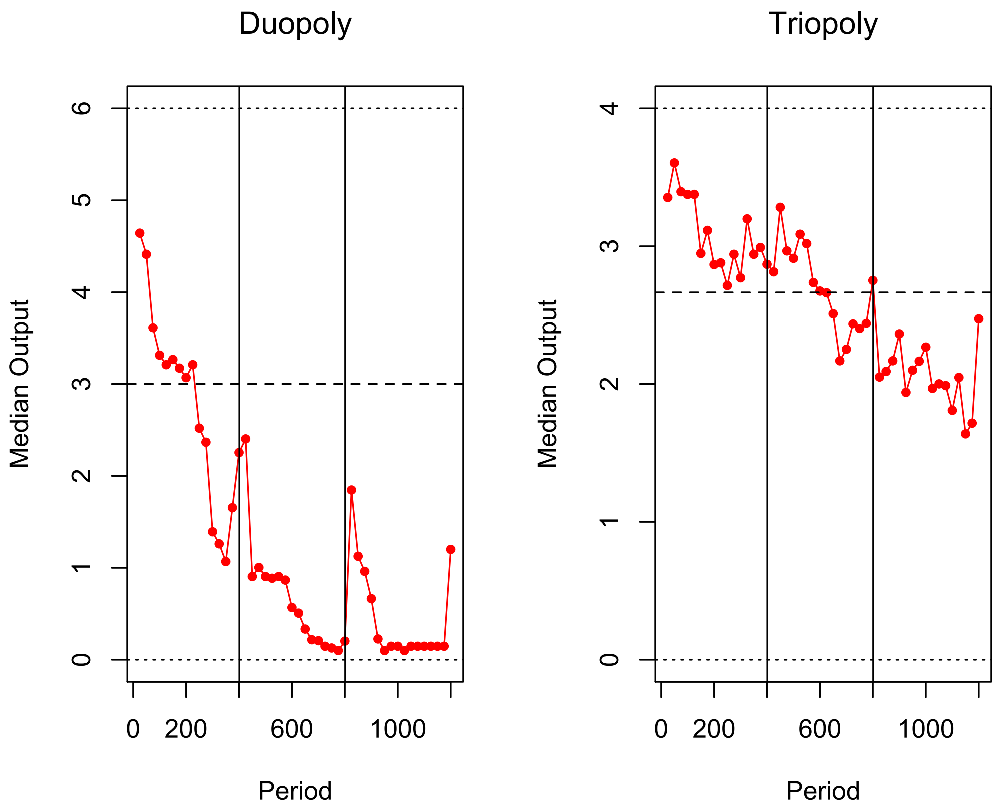


Result 1 *In both duopolies and triopolies, median output quantities initially trend upwards, and settle between CNE and PCW levels during the first 50 periods.*

The next figure is central to our study, and gives an overview of behavior in the long run. Figure 3 plots output choice over the full 1,200 periods of our experiment, with each dot representing the median quantity in a 25-period bin. The three blocks are demarcated by solid vertical lines. Analogous profit graphs can be found in Appendix A.

In Duopoly, there is a stark contrast between the first fifty periods (two dots) and the long-run. Highly competitive outcomes as predicted by Vega-Redondo's imitation model are only observed in those first 50 periods. After that average quantity choices start to drop sharply. Quantities continue to fall even after crossing the Cournot-Nash (CNE) level, and in periods 275-350 are much closer to full collusion (JPM) than to CNE. Of course, the median could hide some interesting heterogeneity. However, inspection of individual groups reveals that none of our matching groups spent any significant time systematically close to the CNE. (More on this below and in Appendix A.2.)

Figure 3: Median quantities in all periods, plotted in 25 period bins.



In the second Duopoly block, collusion becomes prevalent much more quickly; in some duopolies it is nearly perfect and remarkably stable for long intervals of time. Collusive tendencies are even more pronounced in the third block.

In Triopoly, quantities again start to trend downwards after intense competition in the first 50 periods. However, the decline of quantities (and the rise of profits) is much slower than in Duopoly and never approaches full collusion on average (although there is one group of subjects that colludes perfectly in the last block). Also, heterogeneity across groups is much greater than in Duopoly, especially in the last block. Nevertheless there is a systematic trend that takes subjects deep into the collusive territory between CNE and JPM.

To document the drop in quantities, note that the median output choice falls from the first to the final block in each of the six silos in each treatment. For each treatment, this decrease is statistically significant at the (two-tail) five percent level according to a paired Wilcoxon Signed-Rank test.

Table 2 summarizes our aggregate results. It shows median quantities, prices, and profits for the three blocks and also for the first and last 50 periods only. An analogous table reporting means

Table 2: Median quantities, prices, and profits

	Duopoly			**Triopoly**		
Periods	Quantity	Price	Profit	Quantity	Price	Profit
$1-50$	4.54	13.98	23.74	3.46	12.52	16.52
$1-400$	3.17	18.43	35.45	3.11	13.59	18.66
$401-800$	0.57	90.01	63.11	2.74	14.60	21.20
$801-1,200$	0.28	107.36	68.53	2.08	18.70	26.73
$1151-1200$	0.40	91.30	68.51	2.03	19.44	23.03

can be found in Appendix A.

Result 2 *After peaking in the first 50 periods, quantities in both Duopoly and Triopoly begin a long decline towards the collusive JPM level. Median quantities closely approximate JPM by the final block in Duopoly, while in Triopoly median quantities fall nearly by half, and remain well below the CNE level.*

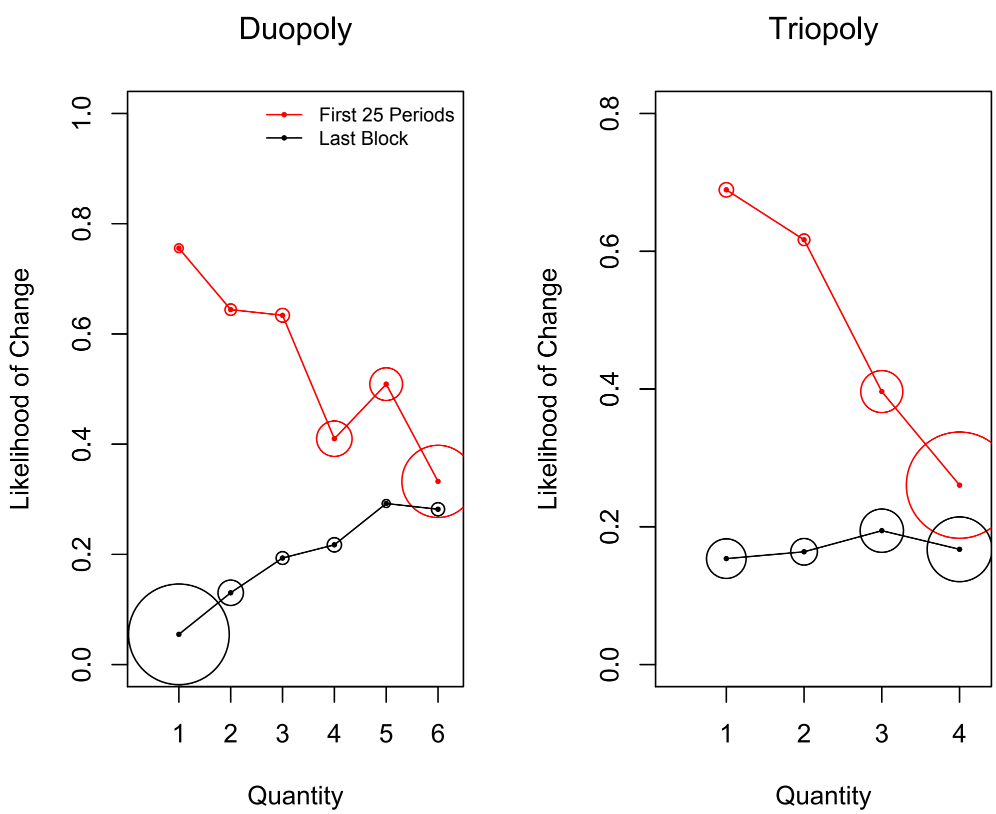
Figure 3 also shows clear end-game effects; evidently subjects are aware of the finite nature of the game. There are also clear restart effects: after rematching, subjects take a while before reducing quantities to the cooperative levels seen in the previous block.

Figure 4 provides an alternative perspective on aggregate behavior. It plots the likelihood that a subject adjusts quantity in period $t+1$ as a function of her quantity choice in period t. Each point in the Figure represents averages from a quantity bin $[0,1],(1,2],(3,4]$... Data from the first 25 periods (in red) are plotted separately from data from the final block (in black).[2] The radius of the circle around each point is proportional to the number of subjects producing that quantity in period t.

The large red circles show that highly competitive output choices are most persistent as well as most common in the first 25 periods in both duopolies and triopolies. This is consistent with the imitation models described earlier. But the pattern changes dramatically by the final block. In Duopoly persistence completely reverses, with collusive quantities becoming most stable and

[2] The last time bin contains strong end-game effects, so we expand the window for capturing behavior late in the session to include the entire 400 period block. The results are qualitatively the same for all plausible specifications, and indeed would be sharper if we excluded the first few time bins, which contain restart effects. See Appendix A for finer details.

Figure 4: Stability of quantities.



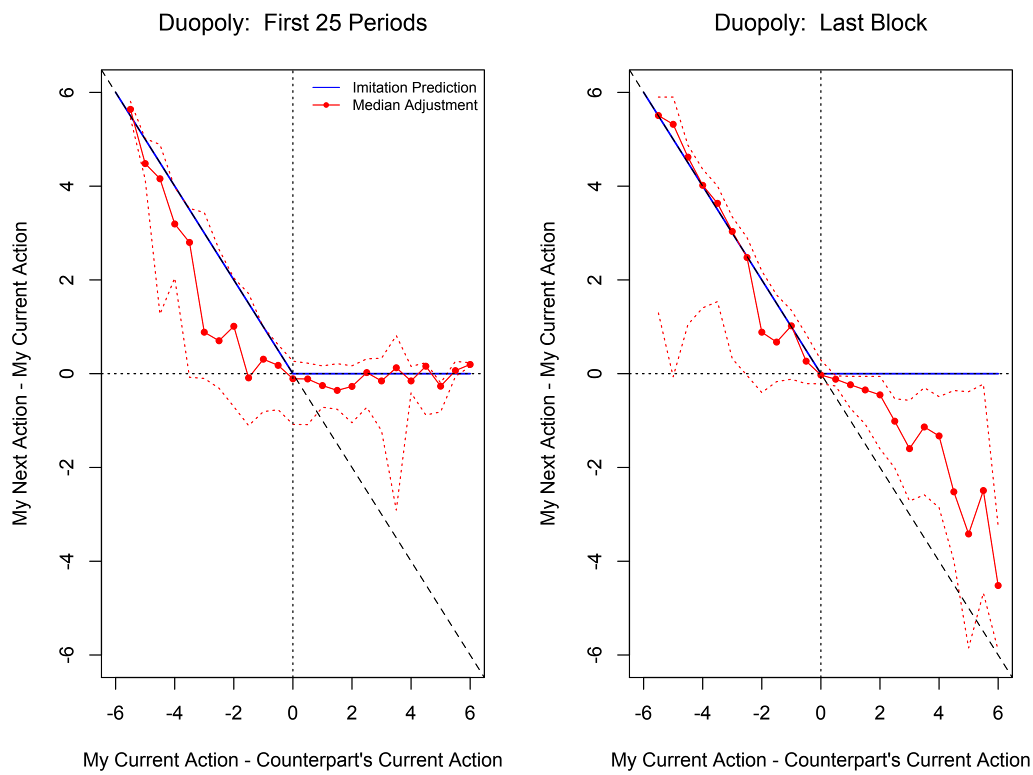
Walrasian quantities least stable. In Triopoly, where groups are more heterogeneous, the relationship becomes almost flat. Still, the change from early period behavior is striking.

Figure 4 illustrates another important finding. The black lines lie well beneath the red lines, indicating that subjects are considerably less likely to change their quantities later in the experiment. This suggests subjects approach a behavioral equilibrium with experience, particularly in Duopoly where colluding subjects rarely change their quantities.

Result 3 *In early periods PCW outputs are most stable and JPM outputs least stable. By the final block, this pattern has disappeared, and has reversed in duopoly. Overall, subjects adjust less frequently by the final block.*

Thus the aggregate data suggests that subjects do learn out of the mal-adapted imitation heuristic. Indeed, it suggests that the behavioral change begins just after 50 periods, the number available to previous studies within the usual two-hour session. While earlier studies may have been tantaliz-

Figure 5: Median quantities in Duopoly



ingly close to detecting the behavioral change, our data also show that it may take several hundreds of periods to achieve behavioral equilibrium in a low information environment.

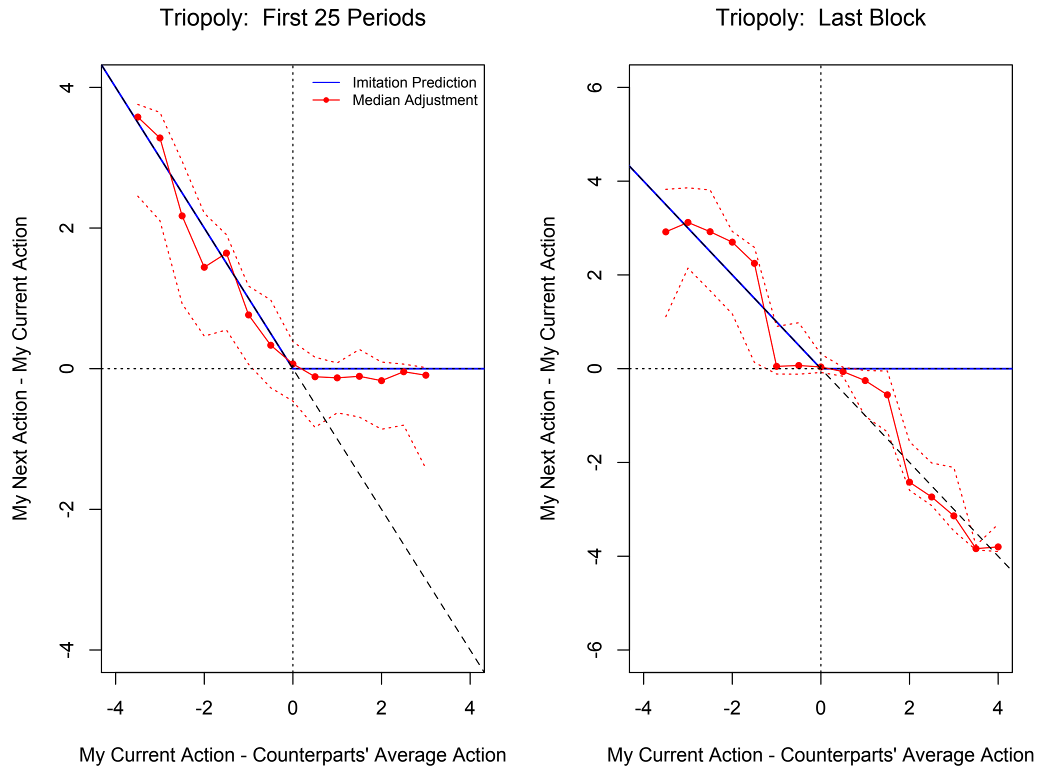
5 Individual behavior

What sort of individual adaptation lies behind the change in aggregate behavior? We now compare individual quantity adjustments early on to those in later play, and see whether subjects eventually employ a (myopic) best response, or some other rule, once they gain experience with the payoff function.

Figures 5 and 6 plot adjustments from period $t-1$ to period t as a function of the difference between own quantity and counterparts' average quantity in period $t-1$. Dots connected by solid lines show binned medians[3] while dotted lines show the 25th and 75th percentiles, i.e., the central 50 percent confidence interval. The noiseless prediction from Vega-Rodondo's imitation rule is shown

[3] Finer action bins are possible here than in the previous figure, because the difference data are more evenly dispersed than the level data.

Figure 6: Median quantities in Triopoly



in blue. That rule prescribes matching higher counterpart output, thus a point on the off-diagonal for negative output differences, and no change when own output is higher, thus following the positive x-axis above zero.

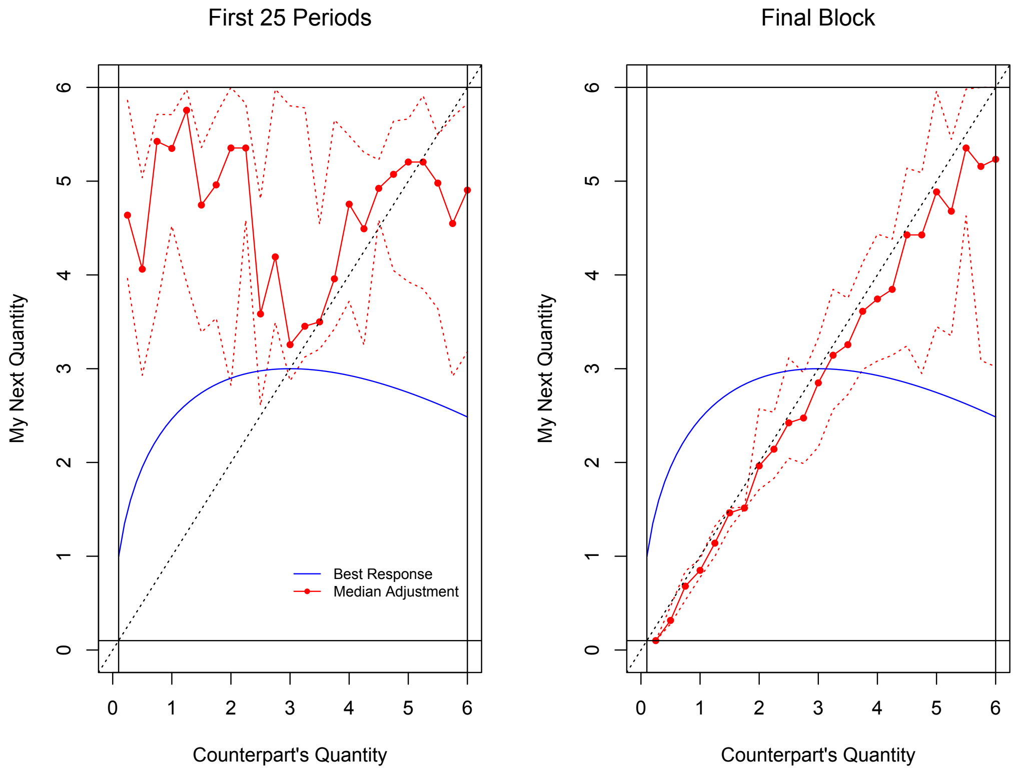
The left hand panels of the figures show that output adjustments conform strikingly well to this prediction at the beginning of the session. The prediction lies everywhere within the central confidence interval and the empirical noise amplitude generally seems reasonable. We conclude that during the first 25 periods, subjects do indeed tend to "imitate the best."

The right hand panels tell a much different story. The median adjustment is now close to the off-diagonal line regardless of whether subjects have lower or higher quantities than their counterparts. Otherwise put, in the last block the predominant mode of adjustment is to match others' output choices whether or not they earned higher profits last period.

Result 4 *In early periods, subjects tend to imitate the most profitable player (including self). In later periods, by contrast, subjects tend to match counterparts' actions regardless of their profitability.*

We conclude that subjects indeed escape the destructive imitate-the-best heuristic. But do they

Figure 7: Median quantities in Duopoly



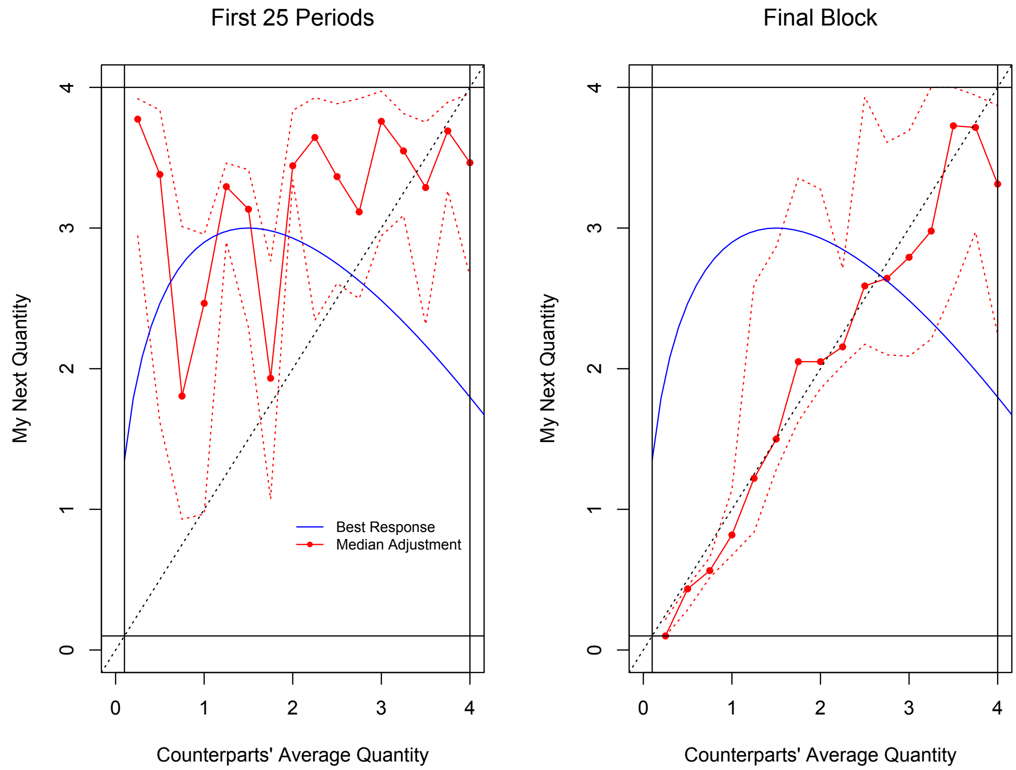
move towards (myopic) best response? Figures 7 and 8 plot opponents' (average) quantities in period $t-1$ on the x-axis and own quantity in period t on the y-axis. Again, connected red dots show median responses and dotted lines show the 25th and 75th percentiles. The blue line shows the best-reply prediction and the main diagonal shows perfect imitation.

The left panels again show data from the first 25 periods. Quantities virtually never coincide with best response. This is not surprising as subjects are given no initial information about their payoff functions. Consistent with the previous graphs and with the imitation heuristic, quantities tend to roughly follow the diagonal at high quantities and exceed it at low quantities.

The right panels again show data from the final block of 400 periods. Quantities (except at the sparsely populated upper end) are tightly bunched along the diagonal, again showing that subjects tend to indiscriminately match their counterparts' quantities. Moreover, the data show no tendency to move towards the best response line, except where it intersects the main diagonal.

Additional evidence comes from from post-experimental questionnaires, reported in Appendix A.3. Fully incentivized elicitations of subjects' beliefs regarding the direction of better replies in the stage game revealed that the vast majority of subjects were aware that one could profitably

Figure 8: Median quantities in Triopoly



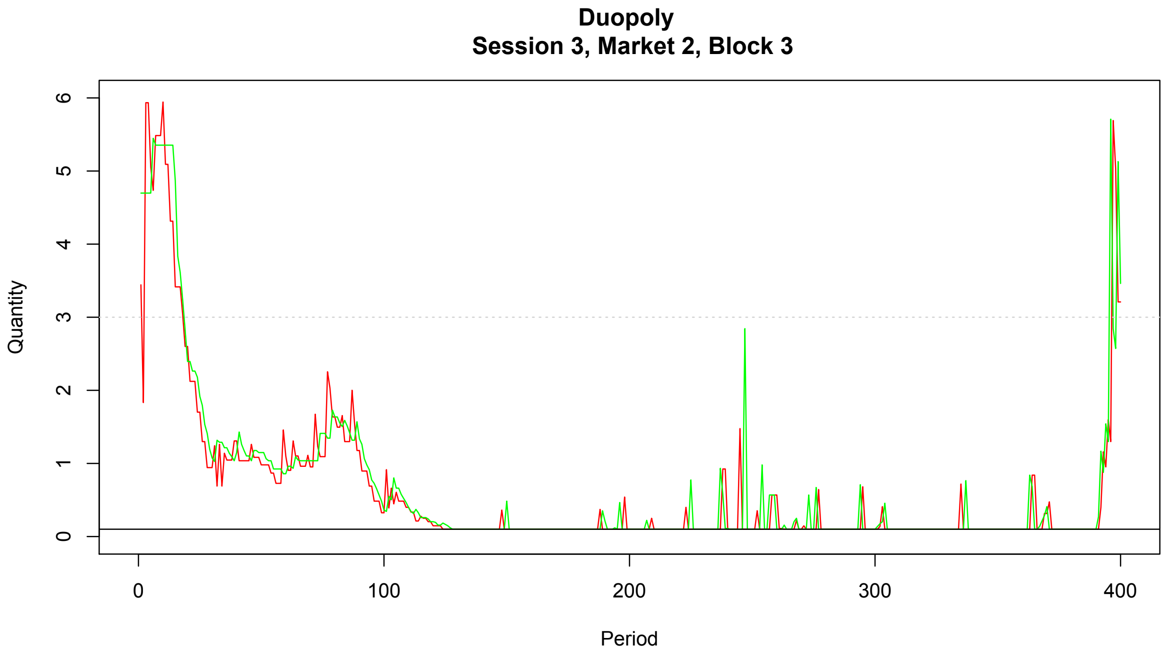
deviate from the JPM. However, they never acquired systematic knowledge of the rough shape of the best-reply correspondence. For example, very few subjects realized that the best reply against the CNE profile is the CNE action. Rather they believed that higher quantities would be more profitable. Yet more evidence on the irrelevance of BR in explaining subjects' behavior can be found in Appendix A.2.

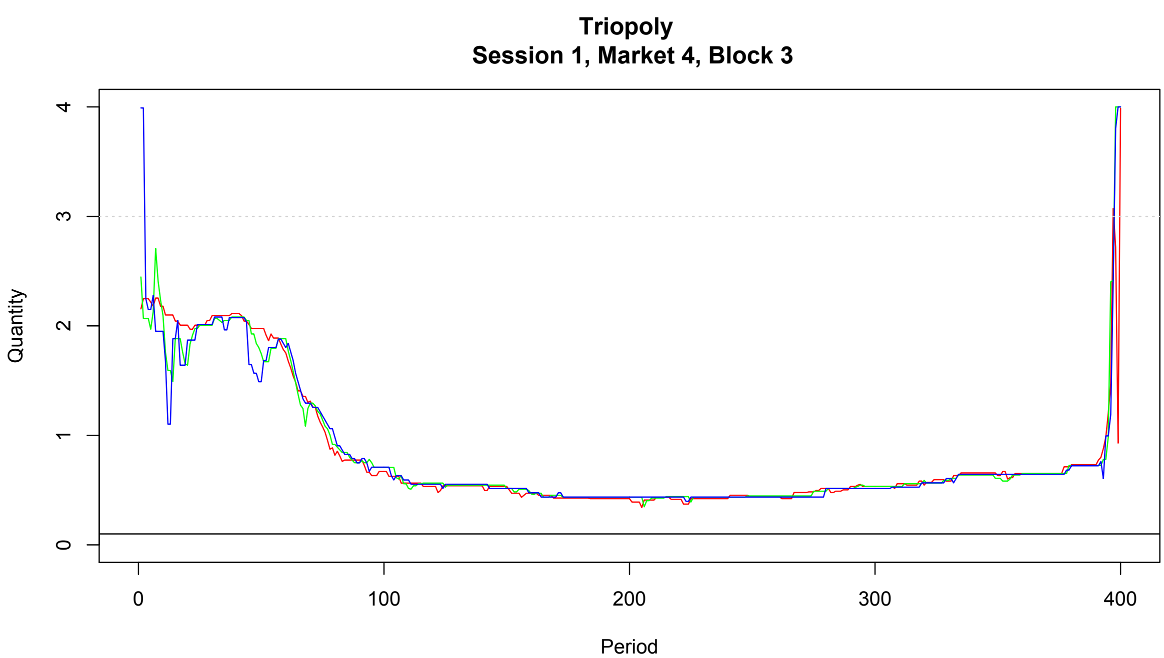
Result 5 *Subjects show no tendency towards employing the best response in either the short run or the long run.*

Granted that, after abandoning the imitate-the-best heuristic, subjects mainly adopt an unconditional strategy of matching others' choices and not an approximate best response, one would then like to know the implications. How does unconditional matching play out over time?

Some insight can be gleaned from inspecting examples of the last 400 period block, as in Figure 9. Individual subjects' output choices are plotted in red and green and (in triopoly) also blue. In the duopoly example, after an initial flurry, subjects closely track each other on a steady decline towards collusion, but at around period 20 (that is, period 820 in aggregate plot numbering) they

Figure 9: Examples from the data illustrating matching behavior.





level out. They then test each other, remaining in the middle of the collusion zone between CNE and JPM until around period 80, at which point they resume a slow downward march with one small interruption. After about 120 periods of their duopoly, they achieve full collusion at the JPM, and remain there for most of the remaining periods. We see a sprinkling of brief episodes in which one duopolist defects for a period or two, and the other immediately retaliates, usually proportionately. Here the gains from unilateral deviations clearly are fleeting, and collusion collapses only in the last few periods. Similar patterns can be seen in the triopoly example, but with less exploration of the action space. Perhaps for that reason, quantities stall out slightly above the JPM, until cooperation again fails in the last few periods. Inspecting other example blocks shows that subjects sometimes start reducing their quantities more or less simultaneously, while in other oligopolies a single subject takes the lead, reducing her own quantity to demonstrate to her counterparts that higher payoffs are available.

Appendix A includes "bar code" diagrams that compactly summarize behavior over time in all 48 duopoly blocks and all 36 triopoly blocks. The diagrams partition action profiles into three color-coded categories: competitive (all players' payoffs are below the CNE level), collusive (all payoffs above CNE), and other (some earn more and some earn less than in CNE). These figures show that the vast majority of deviations from collusion (to other) pass through the competitive region before returning back to collusion. This is unconditional matching at work which (as discussed below) reminds us of Tit for Tat.

The diagrams also show that the spells in "other" tend to get shorter and the collusive spells get longer. In the first block, the average collusive spell lasts 24.4 periods in duopoly and 2.8 periods in triopoly. This increases to 139.2 and 38 in the second block and, finally, reaches 174.2 and 67.2 in the last block. On the other hand, the average number of consecutive periods spent in non-collusive regions, conditional on a defection from collusion, drops in Duopoly from 68 to 56.6 to 13.8.[4] The pattern is slightly different in Triopoly where the respective figures are 86, 142, and 90.1, illustrating how much more complicated it is with three players to coordinate behavior.

6 Discussion

At first we were quite puzzled by our results. We expected to see some movement away from the self-defeating imitate-the-best heuristic found in earlier work, but conjectured that subjects, sampling

[4] Besides indicating better coordination in later periods, this finding suggests to us that most of our subjects are not tired or bored even after 1000 periods.

more and more information about the payoff functon over time, would learn to best respond and eventually converge to the Cournot-Nash equilibrium. We thought it barely possible that they might learn Nash reversion strategies and move into the collusive region.

Instead, we saw subjects move from the strategically naive imitate-the-best heuristic to an apparently even more naive heuristic of unconditional matching, irrespective of how profitable the matched strategy was. Even stranger, the upshot was that subjects then earned profits higher than in our most optimistic conjecture!

Slowly we came to realize that, although matching others' quantities unconditionally seems simpleminded, it in fact is very effective in sustaining collusion. The process is not unlike tit-for-tat. A cooperative move (reducing one's own output) is profitable if it is soon matched by counterparties, i.e., if it provokes cooperation on their part. A defection move (increasing output) is not very profitable if counterparties match it next period, i.e., if it provokes others to defect. Indeed, by aligning outputs unconditionally, subjects experience that better collective outcomes are available, and that incentivizes further downward adjustments.

The adjustment process whereby subjects reduce their quantities in parallel small steps is not well described by any prominent learning model. We have documented that subjects do not switch to (myopic) best replies when more information about the payoff matrix emerges. Nor do we see traces of fictitious play, which would require knowledge of the best-reply correspondence and predict convergence to the CNE. Selten's directional learning model predicts gradual adjustments towards better replies, but on the contrary, once subjects are below the CNE, they systematically chose worse replies. Nor can reinforcement learning explain our subjects' systematic exploration of new (lower) quantities.[5]

The behavior we observe is reminiscent of the "win-continue, lose-reverse" learning algorithm suggested by Huck, Normann, and Oechssler (2003, 2004). The 2003 paper analyzes a class of dilemma games where agents move on a grid. Each agent determines the direction of the next step on the grid by examining their change in payoff. As long as the payoff increases, an agent continues to move into the same direction. Once the agent's payoff drops, the direction is reversed. The 2004 paper considers a continuous-time version of this process for Cournot games. In both cases, it is shown that behavior converges globally to the symmetric JPM. The basic logic for this result is that the process first aligns agents' actions. They cannot move systematically away from each

[5] In Appendix C.2. we show that the CNE is the only profile that survives the iterated elimination of strictly dominated strategies. As shown by Beggs (2005), a consequence of this is that reinforcement learning also converges to the CNE.

other. (Just consider a Cournot duopoly with the large firm moving up and the small firm moving down; then the price stays constant; and it is impossible that output increases for large firms and output reductions for small firms are, both, profitable.) Once actions are aligned, agents essentially search for the JPM. The CNE cannot be stable as parallel output reductions increase all payoffs and eventually the JPM is reached.

While the “win-continue, lose-reverse” model has a similar flavor to what we observe, it is hard to take that model to our data. Our subjects are not bound to move stepwise on a grid and actual movements are sometimes quite large, especially late in many blocks when subjects test their opponents and occasionally defect. It is not obvious how to modify the model to accommodate such jumps, much less endgame effects.

Although our subjects learn to play a repeated game effectively they do not acquire the rationality assumed in folk theorems. In fact, they never learn to best reply, not even for the most relevant of strategy profiles. In some sense, of course, this does not matter. Subjects do not play the one-shot game; they play a repeated game. And what they learn about the repeated game is just enough to achieve collectively rational outcomes.

7 Conclusion

We believe that our study makes three fundamental contributions. First, it shows the relevance of long horizons. It sheds light on the relative importance of the amount of experienced feedback as opposed to the mere passing of time. Previously, 50 periods was generally considered sufficient to observe settled behavior. Now we see that the technical limitations of earlier software (for which implementation of longer horizons was impractical) meant that important aspects of learning in the long run were simply missed. Interestingly, time as such (providing subjects with the opportunity to analyze the game through cognition) turns out not to be the major bottleneck. Behavior in the first 50 periods of our experiment nicely mirrors behavior observed in earlier studies although in our experiment 50 periods take less than four minutes while in previous studies over an hour would have passed. In other words, multiplying the clock time for consideration by a factor of ten to twenty seems (in the case of Cournot games at least) not to make a difference. Conversely, increasing the amount of feedback through sheer repetition changes the picture dramatically.

Second, we see how additional repetitions help subjects to learn their way out of a superficially attractive but ultimately fallacious heuristic. Eventually imitation of successful others ceases to be attractive. Subjects learn that they are hurting themselves and are able to overcome their initial impulse to copy what has made others relatively more successful. Escape is possible even from a devilishly baited trap.

Third, we offer a new perspective on the emergence of cooperation. Subjects replace mal-adapted imitation by other heuristics. Interestingly, these other heuristics are neither more complicated nor more sophisticated: they are just better suited to the repeated-game setting. Subjects learn that it is in their collective interest to produce small quantities. They move into collusive territory through alignment of actions and a local ("win-continue, lose-reverse") search heuristic. By mutually matching quantities, subjects teach one another that their actions will be shadowed by others in the future, encouraging search for high collective payoffs (rather than search for individual best response). This is reminiscent of the old literature on conjectural variations (Friedman 1977). In our experiment, subjects do not merely conjecture that others will match their output adjustments; they actually experience it first hand. Consequently, they learn over time that deviations from cooperation do not pay. The ever increasing length of collusive spells in our data confirms this sort of emerging sophistication.

While we are not able to identify and estimate a precise structural model of the underlying learning process — our experiment was simply not designed for the task — our results call for new theoretical efforts to capture the long-run emergence of cooperation through the adaptation of heuristics. Matching of others actions and a gradual slow search process appear to be desirable ingredients. In fact, we have simulated such a process with some noise and were able to generate dynamics not complectly unlike those that we observed.

In larger classes of games and under different informational conditions there will, of course, be other heuristics that subjects might find initially attractive, and other heuristics to which they might eventually converge. The heuristics identified in this paper are particularly suited for symmetric low-information settings with ordered strategies. The literature dealing with better a priori information about demand and cost functions, for example, has shown that myopic best replies are of immediate attraction to subjects, leading them into Nash equilibrium outcomes. Again, one might ask whether in long-horizon settings like ours, subjects would learn out of such an inefficient heuristic. In general, the set of relevant heuristics might be large and in some games it might be harder to overcome mal-adaption than in others. Studying long-run learning of heuristics in different circumstances may emerge as an attractive new agenda in experimental economics. That agenda would also open new avenues for economic theory.

References

ALGER, D. (1987): "Laboratory tests of equilibrium predictions with disequilibrium data," *Review of Economic Studies*, 54, 105–145.

ALÓS-FERRER, C., AND A. B. ANIA (2005): "The Evolutionary Stability of Perfectly Competitive Behavior," *Economic Theory*, 26, 179–197.

APESTEGUÍA, J., S. HUCK, AND J. OECHSSLER (2007): "Imitation—theory and experimental evidence," *Journal of Economic Theory*, 136, 217–235.

APESTEGUÍA, J., S. HUCK, J. OECHSSLER, AND S. WEIDENHOLZER (2010): "Imitation and the evolution of Walrasian behavior: Theoretically fragile but behaviorally robust," *Journal of Economic Theory*, 145(5), 1603–1617.

BEGGS, A. (2005): "On the convergence of reinforcement learning," *Journal of Economic Theory*, 122(1), 1–36.

ESHEL, I., L. SAMUELSON, AND A. SHAKED (1998): "Altruists, Egoists, and Hooligans in a Local Interaction Model," *The American Economic Review*, 88, 157–179.

FRIEDMAN, J. W. (1977): *Oligopoly and the Theory of Games*. North Holland, Amsterdam, New York.

FRIEDMAN, J. W., AND A. C. HOGGATT (1980): *An experiment in noncooperative oligopoly*. Jai Press, Greenwich, CT.

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

HUCK, S., H.-T. NORMANN, AND J. OECHSSLER (1999): "Learning in Cournot Oligopoly - An Experiment," *Economic Journal*, 109, C80–C95.

——— (2003): "Zero-knowledge cooperation in dilemma games," *Journal of Theoretical Biology*, 220(1), 47–54.

——— (2004): "Through Trial and Error to Collusion," *International Economic Review*, 45(1), 205–224.

OFFERMAN, T., J. POTTERS, AND J. SONNEMANS (2002): "Imitation and Belief Learning in an Oligopoly Experiment," *Review of Economic Studies*, 69(4), 973–97.

Pettit, J., D. Friedman, C. Kephart, and R. Oprea (2012): "Continuous Game Experiments," .

Schaffer, M. (1988): "Evolutionarily Stable strategies for a Finite Population and a Variable Contest Size," *Journal of Theoretical Biology*, 132, 469–478.

Schlag, K. (1998): "Why Imitate, and if so, how? A Boundedly Rational Approach to Multi-armed Bandits," *Journal of Economic Theory*, 78, 130–156.

Vega-Redondo, F. (1997): "The Evolution of Walrasian Behavior," *Econometrica*, 65, 375–384.

Table 3: Mean quantities, prices, and profits

	Duopoly			**Triopoly**		
Periods	Quantity	Price	Profit	Quantity	Price	Profit
$1-50$	4.22	17.26	27.81	3.07	13.97	19.32
$1-400$	2.95	82.58	40.53	2.80	18.84	22.05
$401-800$	1.54	259.33	54.57	2.60	33.57	23.98
$801-1,200$	1.34	286.50	56.61	2.01	74.66	29.92
$1151-1200$	1.48	276.75	55.16	2.03	85.51	29.71

Appendices: For On-line Publication

A Additional Analysis

A.1 Profit Time Series

Figures 10 and 11 plot profits over time and are analogous to Figures 2 and 3. Top, middle and bottom dotted horizontal lines represent Cartel, Nash and Walrasian profit levels, respectively. The plots suggest that subjects' profits fall well below Nash levels in the first 50 periods and rise above Nash levels in the long run.

A.2 Mean quantities, prices, and profits

Table 3 parallels Table 2, but shows mean (instead of median) quantities, profits, and prices in each of the three blocks and in the first and last 50 periods.

A.3 Failure of Best Response Over Time

In this subsection, we provide evidence that subjects never in the aggregate experience a period of consistent best response. Figure 12 provides 6 panels. Each corresponds to a 1-point range of counterparts' previous period average quantity (ranges are listed above each plot). In each range the range of best responses is demarcated by dashed horizontal blue lines. Dashed horizontal red lines provide the bounds for imitating average quantity. The x-axis of each panel plots period. Data is binned into 50 period intervals and the black line plots medians. Figure 13 provides analogous

Figure 10: Median profits in early periods

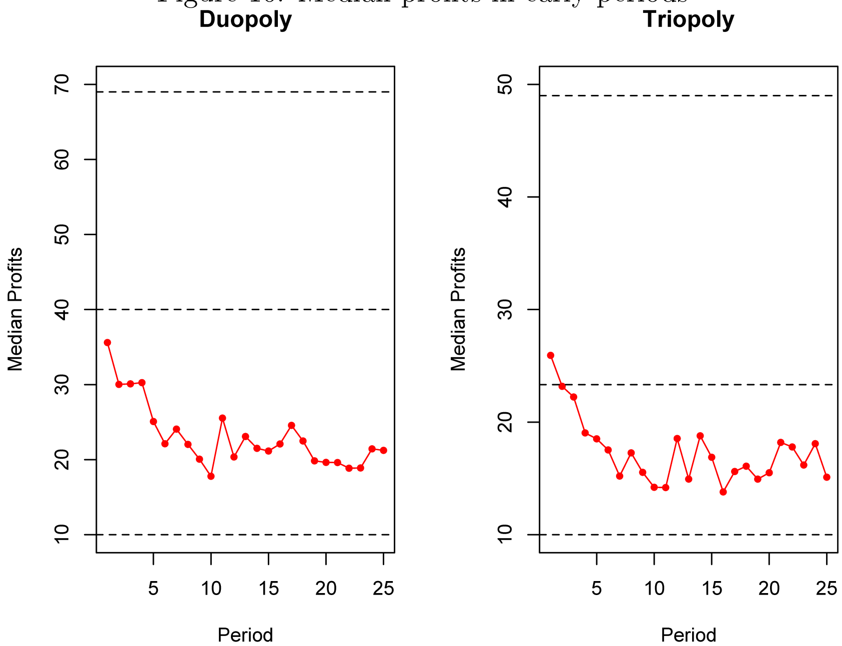
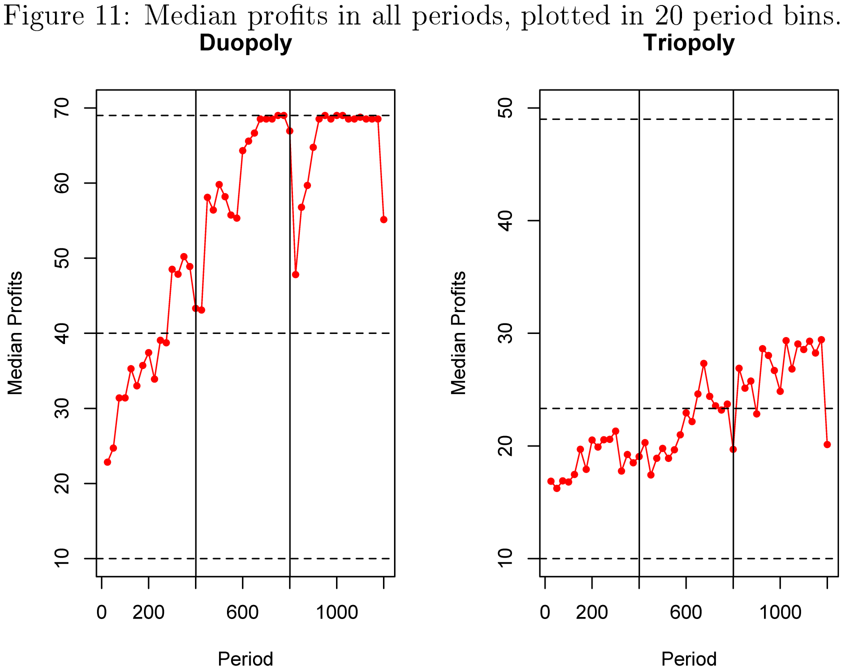


Figure 11: Median profits in all periods, plotted in 20 period bins.



Duopoly					
	Choice other		Alternatives		
Q.	x_2		x_1^1	x_1^2	correct
D1	3		1	3*	12/12
D2	3		3*	6	0/12
D3	1.15		1.15	2.31*	12/12
D4	6		2.49*	6	6/12
Triopoly					
	Choices others		Alternatives		
Q.	x_2	x_3	x_1^1	x_1^2	correct
T1	4	4	1.8*	4	4/12
T2	0.1	0.1	0.1	1.35*	9/12
T3	0.1	0.1	1.35*	4	5/12
T4	2.66	2.66	0.75	2.66*	11/12
T5	2.66	2.66	2.66*	4	0/12

Table 4: Best response quiz. Correct answers are denoted by an asterisk.

data for Triopoly.

It is evident from these figures that median quantities very seldom enter the blue bounds of best response, and that the exceptions are isolated, not bunched. The data therefore are inconsistent with subjects entering a phase of best response at the aggregate level. Instead, plotted data tend to increase from panel to panel after early periods, consistent with unconditional imitation.

A.4 Incentivized Quiz Results

At the end of some of the later sessions, subjects were shown printouts of screens similar to the ones used in the experiment. Markers denoted the counterparts' strategies and two slider positions indicated two possible strategies available. Subjects were asked to circle the slider that would earn the higher payoff in the one-shot game given the counterparts' strategies, and they received a cash payment of $0.50 for each correct answer. Table 4 summarizes the questions and reports on the fraction of correct answers.

Questions D1 and T4 asked whether the CNE quantity or a lower quantity gives a higher profit against the other(s) choosing the CNE quantity. Almost everybody had this question correct,

Figure 12: Response to counterpart actions over time in Duopoly.

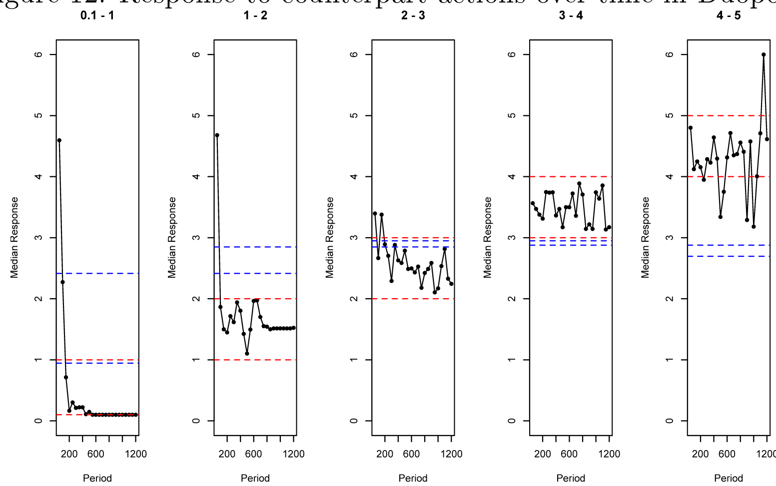
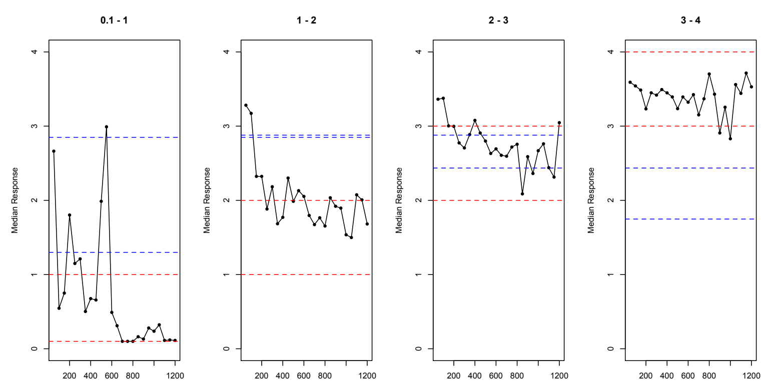


Figure 13: Response to counterpart actions over time in Triopoly.



indicating that subjects are aware that downward deviations from the CNE are not profitable. D2 and T5 asked a similar question: is an upward deviation from the CNE profit increasing? Strikingly, nobody had this question correct. D4 and T1 asked whether subjects would choose a best response to the PCW-outcome or would go for the PCW outcome themselves. The message that emerges is somehow mixed: in Duopoly half of the subjects belief that the PCW-quantity earns higher profits than the best response and in Triopoly 3/4 of the subjects held this belief. D3 and T2 asked whether individual profits are higher at a (rather) collusive outcome or when deviating to a higher quantity. Everybody had this answer correct in Duopoly and 3/4 had this answer right in Triopoly. Thus, almost everybody was aware that it pays off to deviate from the collusive outcome. Finally, T3 asked whether subjects think that the PCW outcome gives a high payoff than the best response when the others collude. 7/12 subjects had this question wrong. The overall message that emerges from this exercise is that subjects at best have a rather blurred picture of their optimal strategy choice in this oligopoly game.

A.5 Bar codes and Punishment

We partition the state space into three regions: competitive (if all players' payoffs are below the CNE payoff), collusive (if all players' payoffs are above the CNE payoff), and other (where some earn more and some earn less than the CNE payoff). We color-code these regions red (competitive), green (collusive) and black (other). Figures 5 and 6 plot transition probabilities over time for movements between these regions. Figures 7 to 12 show bar codes where every period is represented by a single color-coded bar indicating in which region subjects stayed in every period. These figures show one of the more remarkable features of the data — namely how, after a deviation from the collusive region occurs (that is after a change from green to black), play almost always moves into the competitive region (that is into the red) before returning back to collusive play.

Subjects' reaction speeds get faster from block to block and punishment phases get shorter and shorter in duopolies. For triopolies, we see how this process is nosier and slower, refecting the more difficult coordination problem.

The transition probabilities demonstrate several features of the data set: They show the increasing stability of collusion for both duopolies and triopolies. And they show how rare are direct transitions from collusive to competitive and vice versa. Almost all changes occur via "other", reflecting individual defections (rather than common dissatisfaction with collusive outcomes) and demonstrating that forgiveness and repentance occur subsequently rather than simultaneously.

Figure 14: Transition probabilities, Duopoly.

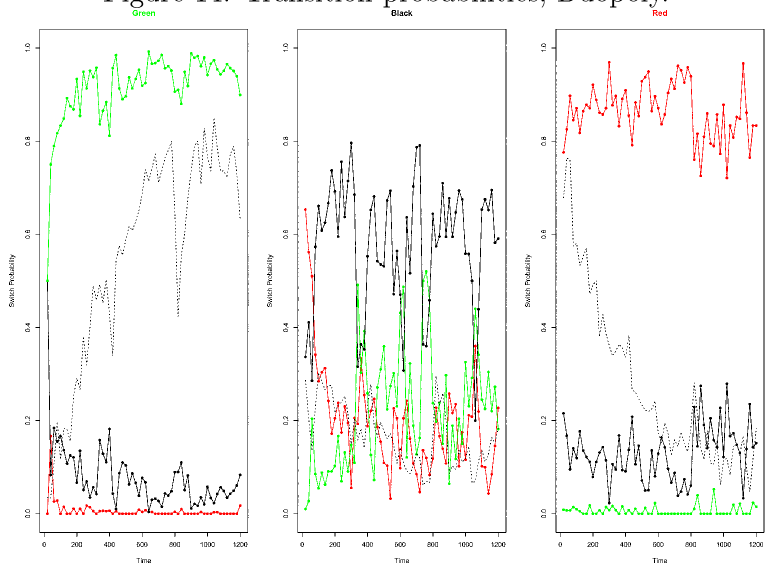


Figure 15: Transition probabilities, Triopoly.

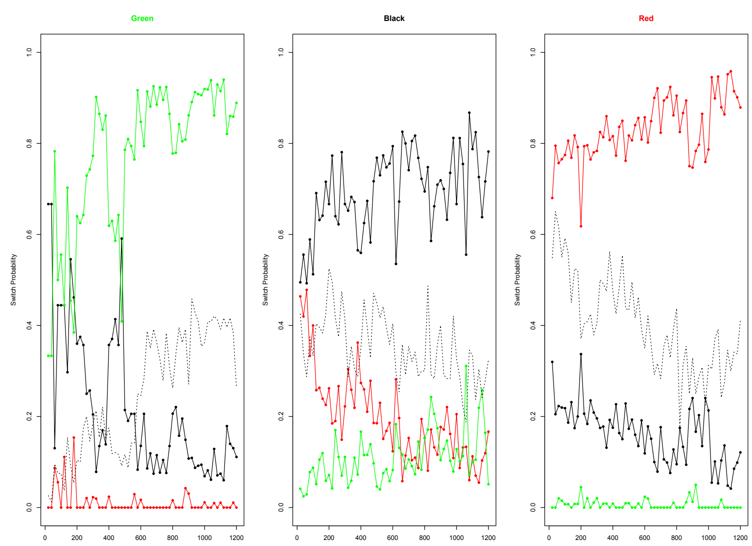


Figure 16: Bar codes from Block 1, Duopoly.

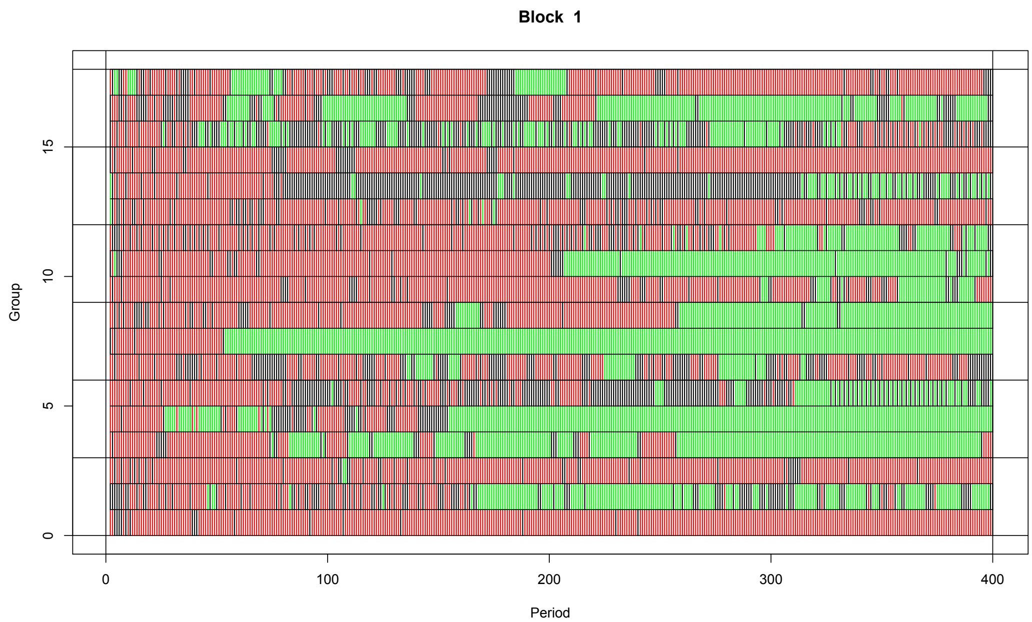


Figure 17: Bar codes from Block 2, Duopoly.

Figure 18: Bar codes from Block 3, Duopoly.

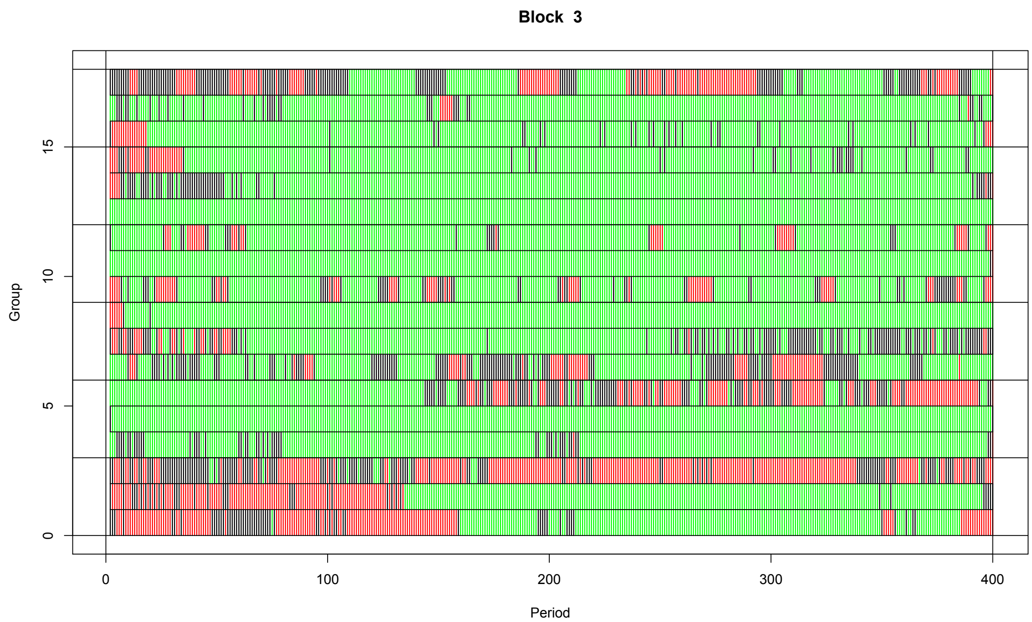


Figure 19: Bar codes from Block 1, Triopoly.

Figure 20: Bar codes from Block 2, Triopoly.

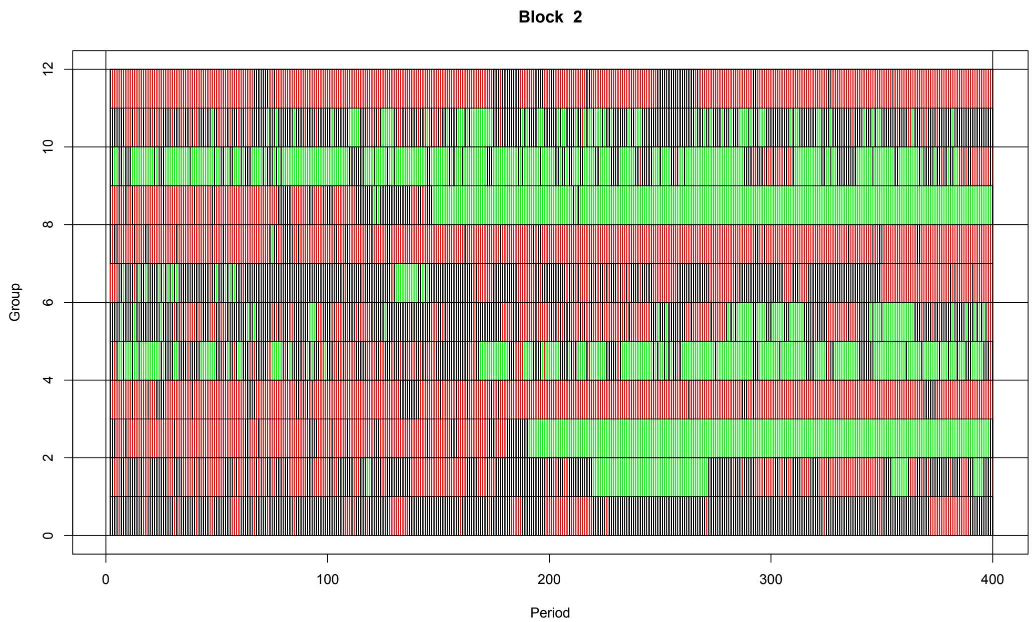
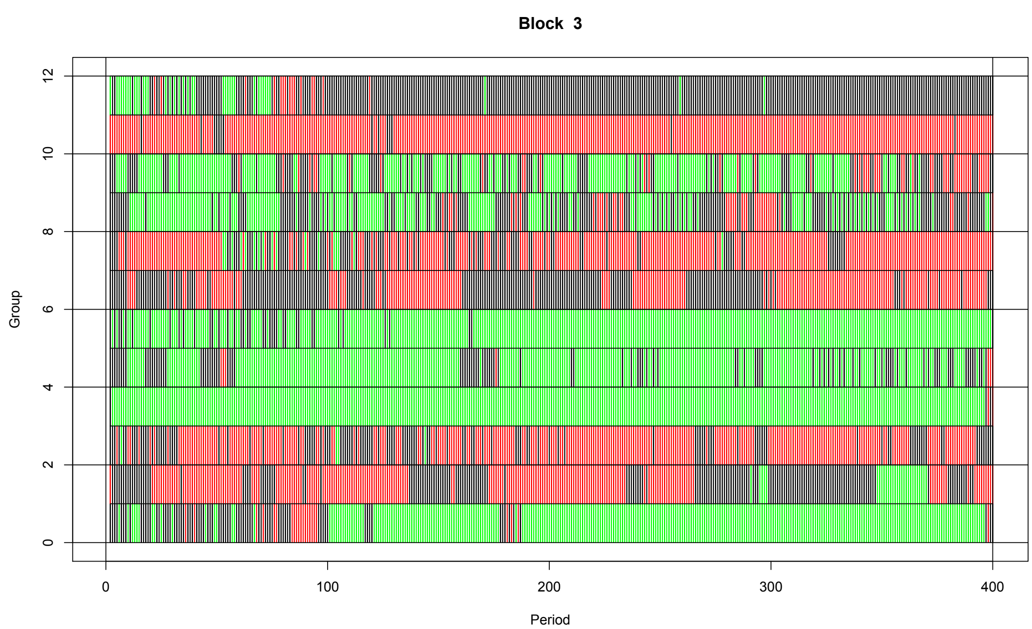


Figure 21: Bar codes from Block 3, Triopoly.



B Instructions

These are the instructions used in both Duopoly and Triopoly sessions. In the instructions we used the term “period” to refer to what the paper calls “blocks” and “subperiods” to refer to what the paper calls “periods.”

Instructions

Welcome! This is an economics experiment. If you pay close attention to these instructions, you can earn a significant sum of money, which will be paid to you in cash at the end of the last period. Please remain silent and do not look at other participants’ screens. If you have any questions, or need assistance of any kind, please raise your hand and we will come to you. If you disrupt the experiment by talking, laughing, etc., you may be asked to leave and may not be paid. We expect and appreciate your cooperation today.

The Basic Idea

The experiment will be divided into a number of periods and in each period you will be anonymously matched with one or two other players via the computer. Each period will be further divided into a number of subperiods. In each subperiod you and your counterparts will secretly select strategies and at the end of the subperiod the combination of your and your counterparts’ strategies will determine your earnings for the subperiod.

We will not tell you exactly how earnings are determined but here are a few facts:

- Your earnings in each subperiod depend entirely on your strategy and your counterparts’ strategies, and nothing else.
- The function that determines your earnings will not change over the course of the experiment. That is, if you and your counterparts use the same strategies at time A as at time B, you both will all have the same earnings at time A as at time B.
- Your earnings are symmetric with your counterparts’. In particular, if you and your counterparts all choose the same strategy, then you all will earn the same amount.

The screen display

Figure 1 [*identical to Figure 1 in the paper*] shows the computer display you will use to make decisions and interact with your counterpart. At the top of the screen is a bar showing elapsed time in the current subperiod. When the bar fills up the subperiod is over and a new subperiod will immediately begin. Your strategy is the location (from left to right) of the black square slider at the bottom of the screen. During each subperiod you can freely adjust your tentative strategy by clicking on the screen or dragging the slider. Your actual strategy for the subperiod is the location of your slider **at the end** of the subperiod.

When the subperiod is over you will be shown a **green dot** visualizing your payoff rate from that subperiod. The higher the dot, the higher the payoff earned. The precise payoff number is shown floating next to the dot. You will also be shown **blue and red hash marks** at the bottom of the screen showing the location of your counterparts' strategies in the last subperiod and **blue and red dots** representing your counterparts' payoffs from the subperiod that just ended (if you are matched with only one other participant you will only see blue hash marks and dots).

It is important to keep in mind that your counterparts' strategies, your payoff dot and your counterparts' payoff dots always display **outcomes from last subperiod**. You will not learn payoffs or your counterpart's strategy from the current subperiod until after the subperiod is over.

Earnings

Your earnings will be given in points. Point totals reported after each subperiod are given as payoff **rates**, i.e., the payoff you would receive for the entire period if you acted the same way each subperiod. Your actual point earnings for a single subperiod can be calculated by dividing the payoff number reported by the number of subperiods in the current period. For example, if the period contains 50 subperiods and your payoff dot shows earnings rate of 200 in the last subperiod, then you actually earned $200/50 = 4$ points in that subperiod.

Your points will accumulate over the course of the experiment. The screen will always display your "Current Earnings" during the period so far and "Previous Earnings" accumulated over previous periods. You will be paid cash for points earned at a rate written on the white board at the front of the room.

Frequently asked questions

Q1. Is this some kind of psychological experiment with an agenda you haven't told us?

Table 5: Static outcomes for the linear payoff function

	Duopoly			Triopoly		
	x_i	P	π_i	x_i	P	π_i
JPM	3	6	28	2	6	22
CNE	4	4	26	3	3	19
PCW	6	0	10	4	0	10

Answer. No. It is an economics experiment. If we do anything deceptive or don't pay you cash as described then you can complain to the campus Human Subjects Committee and we will be in serious trouble. These instructions are meant to clarify the game and show you how you earn money; our interest is simply in seeing how people make decisions.

C Theoretical Details

C.1 Comparison to linear demand

To document the comparison to linear demand consider the inverse demand function $P = 12 - n\bar{x}$. We summarize the relevant benchmarks for this case in Table 5.

Under our unit elastic demand function, switching to the best response to the JPM-quantity of the other player yields an increase of profits by 58.9% in Duopoly. In Triopoly this temptation is even higher, as the best response to the JPM quantities increases profits by 106.2% Note that the temptations to deviate from the JPM-outcome are much lower in the corresponding linear demand case where a deviator can expect only a 8% rise in profits in Duopoly and a 18.2% increase in Triopoly.

To see that for the unit elastic demand function the payoff function is not as flat around the best response as in the case of a linear demand function for $n < 6$ note the following. Under linear demand the FOC is $0 = \frac{d\phi_i}{dx_i} = 12-(n-1)\bar{x}_{-i}-2x_i$ and payoff curvature is determined by $\frac{d^2\phi_i}{dx_i^2} = -2$. By contrast, for our constant elasticity specification, FOC is $0 = \frac{d\pi_i}{dx_i} = \frac{120}{\sum_j x_j} - 10 - \frac{120x_i}{(\sum_j x_j)^2}$, and payoff curvature is determined by $\frac{d^2\pi_i}{dx_i^2} = \frac{-240}{(\sum_j x_j)^2} + \frac{240x_i}{(\sum_j x_j)^3}$. Substituting for the last term from the FOC and simplifying yields $\frac{d^2\pi_i}{dx_i^2} = \frac{-20}{nx^*}$, where the symmetric NE quantity is $x^* = 12\frac{n-1}{n^2}$. Hence for $n = 6$ we have $\frac{d^2\pi_i}{dx_i^2} = \frac{-20}{(12)\frac{5}{6}} = -2$, the same as for $\frac{d^2\phi_i}{dx_i^2}$, but for lesser n we have $\frac{d^2\pi_i}{dx_i^2} < -2 = \frac{d^2\phi_i}{dx_i^2}$.

C.2 Iterated Elimination of Strictly Dominated Strategies

To show that the CNE is the unique point in the serially undominated set, let us first consider the derivative of the profit function. If this derivative is positive a higher quantity will lead to higher profits and if it is negative decreasing one's quantity is profit increasing. We have

$$\frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} = \frac{120}{x_i + X_{-i}} - 10 - \frac{120 x_i}{(x_i + X_{-i})^2}.$$

We have $\frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} > 0$ if $0 < x_i < 3$ and

$$\underline{X}_{-i}(x_i) < X_{-i} < \bar{X}_{-i}(x_i) \tag{4}$$

where $\underline{X}_{-i}(x_i) = 6 - 2\sqrt{3}\sqrt{3 - x_i} - x_i$ and $\bar{X}_{-i}(x_i) = 6 + 2\sqrt{3}\sqrt{3 - x_i} - x_i$. Note that (4) represents the set of quantities of the other players for which a quantity increase pays off. Likewise, we have $\frac{\partial \pi_i(x_i, X_{-i})}{\partial x_i} < 0$ if

$$x_i > 3 \tag{5}$$

or if $0 < x_i \leq 3$ and

$$X_{-i} > \bar{X}_{-i}(x_i) \tag{6}$$

The previous two inequalities capture cases where, depending on the own quantity and the quantity chosen by the others, a quantity decrease results in higher profits.

Duopoly: Consider an interval of the form $[x_L, \hat{x}_U]$. Note that by (5) we know that a slight quantity decrease will earn strictly higher profits (regardless of the quantity X_{-i} chosen by the other) if $x_i > 3$. Since at an upper bound no quantity increase is possible all upper bounds $3 < \hat{x}_U \leq 6$ are strictly dominated by a lower quantity. Iteratively applying this argument, starting from $x_U = 6$, shows that all upper bounds $3 < \hat{x}_U \leq 6$ are iteratively strictly dominated. Now consider any interval of the form $[\hat{x}_L, 3]$. The set of quantities of the other player for which an increase in the own quantity results in higher profits is given by: $\underline{X}_{-i}(\hat{x}_L) < X_{-i} < \bar{X}_{-i}(\hat{x}_L)$. We have $X_{-i} \geq \hat{x}_L$ and $X_{-i} \leq \hat{x}_U$. Thus, if $\underline{X}_{-i}(\hat{x}_L) \leq \hat{x}_L$ and $\bar{X}_{-i}(\hat{x}_L) \geq \hat{x}_U$ it pays off to increase one's quantity for any quantity chosen by the other player. Both inequalities hold for $\hat{x}_L < 3$. Thus, for any interval of the form $[\hat{x}_L, 3]$ the lower bound is strictly dominated by a higher quantity, showing that the CNE quantity $x_i = 3$ is the only serially undominated strategy.

Triopoly: Again, (6) reveals that as in duopoly all quantities $x_i > 3$ are iteratively strictly dominated by some lower quantity. Thus, we have obtained a new undominated upper bound $x_U^0 = 3$.

Now consider intervals of the form $[x'_L, 3)$. Consider (4) and note that we have $X_{-i} \geq 2x'_L$ and $X_{-i} \leq 6$. We have $\underline{X}_{-i}(x'_L) < 2x'_L$ whenever $x'_L \leq 3$ and we have $\bar{X}_{-i}(x'_L) > 2x_U^0 = 6$ whenever

$x'_L < 6(\sqrt{2}-1) = x^0_L$. Thus, for all lower bounds $x'_L < x^0_L$ we can find a profit increasing deviation if the others choose their quantities in the interval $[x'_L, 3)$. Thus, we have obtained a new lower bound $x^0_L = 6(\sqrt{2}-1)$.

Consider now an interval $[\hat{x}_L, \hat{x}_U]$ with lower bound $\hat{x}_L$ and upper bound $\hat{x}_H$ with $\frac{3}{2} < \hat{x}_L < \hat{x}_U \leq 3$. By (6), it pays off to further reduce one's quantity for each upper bound $\hat{x}'_U$ that satisfies $X_{-i} > \bar{X}_{-i}(\hat{x}'_U)$. We know that $X_{-i} \geq 2\hat{x}_L$. Thus, it pays off to further reduce one's quantity if $2\hat{x}_L > \bar{X}_{-i}(\hat{x}'_U)$. Provided that $\hat{x}_L > \frac{3}{2}$, this can be written as $\hat{x}'_U > f(\hat{x}_L)$ where

$$f(x) = 2\sqrt{6x} - 2x.$$

Thus, we have found a new upper bound $\hat{x}''_U = f(\hat{x}_L)$.

By (4) it pays off to further increase one's quantity for each lower bound $\hat{x}'_L$ if $\underline{X}_{-i}(\hat{x}'_L) < X_{-i} < \bar{X}_{-i}(\hat{x}'_L)$. Since $X_{-i}(\hat{x}'_L) \geq 2x'_L$, the first inequality holds whenever $\hat{x}'_L < 3$. Further, we have $\bar{X}_{-i}(\hat{x}'_L) \leq \hat{2}x_L$ if $\hat{x}_U > \frac{3}{2}$ and $\hat{x}'_L < f(\hat{x}_U)$. Hence, we have found a new lower bound $\hat{x}''_L = f(\hat{x}_U)$.

The previous argument establishes that, for $\frac{3}{2} < \hat{x}_L < 3$ and $\frac{3}{2} < \hat{x}_U \leq 3$, given an undominated interval $[\hat{x}_L, \hat{x}_U]$ we can obtain a new undominated interval $[f(\hat{x}_L), f(\hat{x}_U)]$. We can now iterate the function $f(\cdot)$ on this interval. By the intermediate value theorem, a sufficient condition for the function f to be a contraction mapping is that $|f'(x)| < 1$ which is the case whenever $\frac{2}{3} < x < 6$. Thus f is a contraction mapping which, by the Banach fixed point theorem, assures convergence to the unique fixed point $x = f(x) = \frac{8}{3}$. This, together with the previous observations that $x^0_L = 6(\sqrt{2}-1)$ and $x^0_U = 3$, shows that the CNE is the only quantity in the serially undominated set.

Discussion Papers of the Research Area Markets and Choice 2013

Research Unit: **Market Behavior**

Nadja Dwenger, Dorothea Kübler, Georg Weizsäcker Preference for Randomization: Empirical and Experimental Evidence	SP II 2013-201
Kai A. Konrad, Thomas R. Cusack Hanging Together or Being Hung Separately: The Strategic Power of Coalitions where Bargaining Occurs with Incomplete Information	SP II 2013-202
David Danz, Frank Hüber, Dorothea Kübler, Lydia Mechtenberg, Julia Schmid 'I'll do it by myself as I knew it all along': On the failure of hindsight-biased principals to delegate optimally	SP II 2013-203
David Hugh-Jones, Morimitsu Kurino, Christoph Vanberg An Experimental Study on the Incentives of the Probabilistic Serial Mechanism	SP II 2013-204
Yan Chen, Onur Kesten From Boston to Chinese Parallel to Deferred Acceptance: Theory and Experiments on a Family of School Choice Mechanisms	SP II 2013-205
Thomas de Haan, Roel van Veldhuizen Willpower Depletion and Framing Effects	SP II 2013-206
Christine Binzel, Dietmar Fehr Giving and sorting among friends: evidence from a lab-in-the-field experiment	SP II 2013-207
Volker Benndorf, Dorothea Kübler, Hans-Theo Normann Privacy Concerns, Voluntary Disclosure of Information, and Unravelling: An Experiment	SP II 2013-208
Rebecca B. Morton, Marco Piovesan, Jean-Robert Tyran Biased Voters, Social Information, and Information Aggregation Through Majority Voting	SP II 2013-209
Roel van Veldhuizen The influence of wages on public officials' corruptibility: a laboratory investigation	SP II 2013-210

Research Unit: **Economics of Change**

Luisa Herbst, Kai A. Konrad, Florian Morath Endogenous Group Formation in Experimental Contests	SP II 2013-301

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice

Kai A. Konrad, Florian Morath SP II 2013-302
Evolutionary Determinants of War

Armin Falk, Nora Szech SP II 2013-303
Organizations, Diffused Pivotality
and Immoral Outcomes

Maja Adena, Steffen Huck, Imran Rasul SP II 2013-304
Charitable Giving and Nonbinding Contribution-Level Suggestions.
Evidence from a Field Experiment

Dominik Rothenhäusler, Nikolaus Schweizer, Nora Szech SP II 2013-305
Institutions, Shared Guilt, and
Moral Transgression

Dietmar Fehr, Steffen Huck SP II 2013-306
Who knows it is a game? On rule understanding, strategic awareness
and cognitive ability

Maja Adena, Michal Myck SP II 2013-307
Poverty and Transitions in Health

Friedel Bolle, Jano Costard SP II 2013-308
Who Shows Solidarity with the Irresponsible?

Kai A. Konrad SP II 2013-309
Affection, Speed Dating and Heart Breaking

Maja Adena, Ruben Enikolopov, Maria Petrova, SP II 2013-310
Veronica Santarosa, and Ekaterina Zhuravskaya
Radio and the Rise of the Nazis in Pre-war Germany

Perdro Dal Bó, Guillaume R. Fréchette SP II 2013-311
Strategy Choice In The Infinitely Repeated Prisoners' Dilemma

Steffen Huck, Gabriele K. Lünser, Jean-Robert Tyran SP II 2013-312
Price Competition and Reputation in Markets
for Experience Goods: An Experimental Study

Rune Midjord, Tomás Rodríguez Barraquer, Justin Valasek SP II 2013-313
Over-Caution of Large Committees of Experts

WZB Junior Research Group: **Risk and Development**

Lubomír Cingl, Peter Martinsson, Hrvoje Stojic, Ferdinand M. Vieider SP II 2013-401
Separating attitudes towards money from attitudes towards probabilities:
stake effects and ambiguity as a test for prospect theory

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice