Topic 13 – Individual decision-making I

Martin Kocher
University of Munich

Course in Behavioral and Experimental Economics

Preview of topic 13

EUT vs. PT

+ Kahnemann and Tversky (1979)
+ Starmer (2000)

Risk

+ Rabin (2000): presentation Johannes Maier

Time pressure

+ Trautmann and Kocher (2007)

Some fun

+ Kocher et al. (2007), Sutter and Kocher (2003, 2006)

The main questions treated in this section

- Systematic deviations from Expected Utility Theory (EUT).
- The importance of Prospect Theory (PT).
- Decision-making under risk and ambiguity.
- Time pressure and fun applications of individual decision-making

EUT – A refresher I

Daniel Bernoulli (1738):

- He transforms payoffs of players in the St. Petersburg game (the price for the following gamble: a coin is flipped repeatedly until a head is produced; the player gets € 2^n) into utilities.
- The idea is revived in economics in the 1940ies and 1950ies.
- Neumann and Morgenstern (1947): axioms.

EUT – A refresher II

Following the presentation of Starmer (2000):

- Preferences are defined over prospects.
- Outcomes and probabilities are known (decision-making under *risk*).
- Prospects denoted ***q*** and can be represented by a probability distribution ***p*** = (p_1…, p_n) over a fixed set of pure consequences ***X*** = (x_1,..., x_n).
- Standard assumptions.

EUT – A refresher III

EUT can be derived from 3 axioms

- Ordering (includes completeness and transitivity).
- Continuity

Those two together imply that preferences can be represented by a real-valued function.

- Independence (if $\mathbf{q} \geq r$, then ($\mathbf{q}$, p; $\mathbf{s}$, 1-p) $\geq$ ($\mathbf{r}$, p; $\mathbf{s}$, 1-p) for all p)

Then preferences can be represented by

$V(\mathbf{q}) = \sum_i p_i u(x_i)$ which is a monotonic function.

EUT – Descriptive limitations I

Systematic violations of the independence axiom

Common consequence effect (Allais paradox):

Two sets of prospects (Allais, 1953):

- s_1 = (1 Mill., 1) and r_1 = (5 Mill., 0.1; 1 Mill., 0.89; 0; 0.01)
- s_2 = (1 Mill., 0.11; 0, 0.89) and r_2 = (5 Mill., 0.1; 0; 0.9)

EUT predicts choosing either s_1 and s_2 or r_1 and r_2, but many people choose s_1 and r_2.

EUT – Descriptive limitations II

Common consequence effect more general:

- Two sets of prospects:
- $\boldsymbol{s^*}$ = (y, p; c, 1-p) and $\boldsymbol{r^*}$ = ($\boldsymbol{q}$, p; c, 1-p), where $\boldsymbol{q}$ = (x, λ; 0, 1-λ) with c, x and y being non-negative and $x > y$.
- Both prospects give outcome c with probability 1-p (the "common consequence"), and the independence axiom would imply irrelevance. Many studies, however, show that the size of c matters ($\boldsymbol{s^*}$ chosen when $c = y$, and $\boldsymbol{r^*}$ chosen when $c = 0$).
- Allais: x = 5 Mill., y = 1 Mill., p = 0.11, λ = 10/11.

EUT – Descriptive limitations III

Systematic violations of the independence axiom

Common ratio effect:

Suppose you had to choose in a lottery between a sure prospect of € 3,000 and a prospect with an 80% chance of winning € 4,000 and a 20% chance of getting nothing.

Now choose in a lottery between a 25% chance of winning € 3,000 (otherwise nothing) and a 20% chance of winning € 4,000 (otherwise nothing).

Many people go for € 3,000 in the first lottery and € 4,000 in the second (which is inconsistent with EUT).

EUT – Descriptive limitations IV

Common ratio effect more general:

- $\boldsymbol{s^{**}} = (y, p; 0, 1\text{-}p)$ and $\boldsymbol{r^{**}} = (x, \lambda p; 0, 1\text{-}\lambda p)$, where $x > y$.
- Assume that the ratio of "winning" probabilities is constant, then for pairs of prospects of this structure, EUT implies that preferences should not depend on the value of p.
- Yet, numerous studies reveal a tendency of individuals to switch from $\boldsymbol{s^{**}}$ to $\boldsymbol{r^{**}}$ as p falls.

EUT – Descriptive limitations V

- Deviations from EUT are not idiosyncratic, but systematic.
- Two newer approaches (since the 1970ies) to cope with the evidence:
  (i) Keep most of the axioms and assumptions untouched but adjust the utility function in a way to allow for deviations from the independence axiom.
  (ii) More fundamental changes regarding the assumptions over preferences.

EUT – Descriptive limitations VI

Preferences:

- Implicit assumptions: procedure invariance, description invariance.
- One challenge of description invariance: the preference reversal:

  Subjects have to choose between two risky prospects (one with a small probability and a high price (the $-bet) and one with a high probability and a smaller price (the P-bet)); two tasks: (i) plain choice, and later (ii) assign minimum selling prices *M*($) and *M*(P) to the prospects.

EUT – Descriptive limitations VII

- Studies show that for many people P > $ but M($) > M(P), i.e. the ordering depends on the elicitation procedure [which is actually more often the case that one would expect for many effects in individual decision-making under risk; see also next slide].
- Explanation: (i) Choice and valuation tasks involve different mental processes; (ii) Violation of the transitivity assumption.

EUT – Descriptive limitations VIII

"Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:"

Each group then faced a choice between two policy options.

Options presented to group I:

"If program A is adopted, 200 people will be saved

If program B is adopted, there is a 1/3 probability that 600 people will be saved, and a 2/3 probability that no people will be saved."

Options presented to group II:

"If program C is adopted, 400 people will die.

If program D is adopted, there is a 1/3 probability that nobody will die, and a 2/3 probability that 600 people will die."

Tversky and Kahneman (1981):

72% of subjects prefer A over B, but only 22% of subjects prefer C over D.

Framing effect

Conventional Non-EUT I

- Theories playing around with preferences over risk: EUT imply parallel and linear indifference curves. Those theories relax this assumption in different appropriate ways (fanning out – Machina (1982), weighted utility – Chow and MacCrimmon (1979), disappointment aversion – Loomes and Sugden (1986), aspiration level theories, and many more).
- Theories with decision weights: Humans seem to underestimate large probabilities and overestimate low probabilities (applications: life threats, insurances etc.)

Conventional Non-EUT II

- Edwards (1955, 1962): $V(\mathbf{q}) = \sum_i w_i u(x_i)$ with a probability weighting function $\pi(p_i)$.
- The easiest representation: $V(\mathbf{q}) = \sum_i \pi(p_i) u(x_i)$ but such utility functions are not monotonic anymore.
- Two ways forward: (i) Ensure that probability weighting function has a form that guarantees monotonicity; (ii) go beyond traditional economics (PT).
- Rank-dependent EUT (Quiggin, 1982): Probability weighting depends on ranks; monotonic.

Conventional Non-EUT III

- Where is p^*?

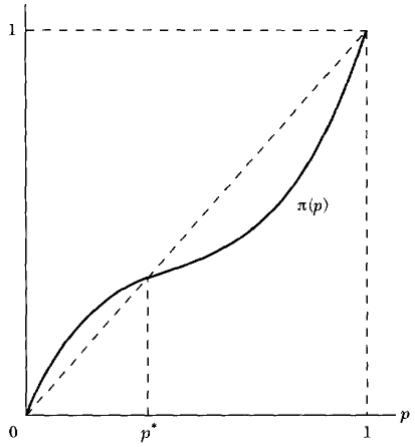


Figure 6. An (inverted) S-shaped Probability Weighting Function

17

Non-conventional theories I

- Conventional theories implicitly assume preference maximization and that humans behave as if optimizing some underlying preference function: basic economic assumptions.
- Non-conventional theories have their roots in psychology: seek to describe the behavior that leads to choice.
- They use decision heuristics and rules of choice (bounded rationality).

Prospect Theory I

- Kahneman and Tversky (1979).
- Choices are a two-phase process:
  (i) Prospects are "edited" using a variety of decision heuristics.
  (II) Choices among edited prospects are determined by a preference function (with probability weighting)
- Two distinctive feature compared to what has been discussed in economics until, then: (i) editing process, and (ii) the use of a reference point for determining gains and losses.

Prospect Theory II

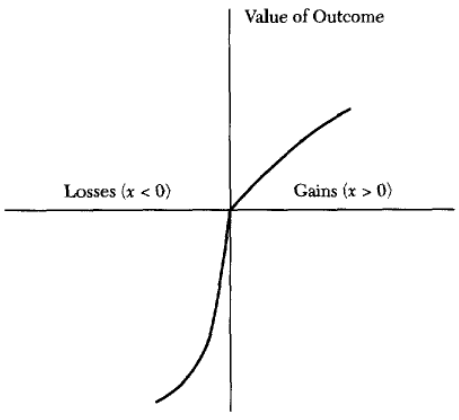


Figure 7. The Valuation of Outcomes in Prospect Theory

- Concave for gains and convex for losses.
- Steeper in the domain of losses.
- Behavioral reasons: diminishing sensitivity and loss aversion ($u'(x) < u'(-x)$).

Prospect Theory III

- In the original paper KT use a probability function with undefined values around the end points.
- Later on, they adopt the inverted S-shaped probability weighting function (cumulative prospect theory).
- PT can explain a wide variety of existing puzzles and evidence.
- The editing process is somewhat less formal. It involves determining the reference point (usually, the status quo), combining prospects, cancellation of common prospects, and the dominance heuristic (if prospects are *obviously* dominated).

Prospect Theory IV

- Some undesirable features: like intransitivity.
- But combining parts of prospect theory and rank-dependent EUT can produce variants of PT that satisfy monotonicity and transitivity (without the need to invoke an editing phase).
- How to, in general, incorporate framing and context-specific aspects in economic decision-making (under risk)?

Experimental evidence

- Decision weighting models fare best (Is this surprising? They have another degree of freedom.).
- Inverted s-shaped weighting seems to be correct.
- Prelec (2000) proposes a specific weighting function with inflection point *p* at 1/*e* and also provides an axiomatization for it.
- A case for the unconventional: violations of monotonicity, event-splitting effects, violations of transitivity in experiments.
- No single theory can account for all effects observed in experiments.

Field evidence

- Behavior on insurance markets can only be explained by a large non-linear overweighting of small probabilities (decision weighting theories).
- The equity premium puzzle (Benartzi and Thaler, 1995: myopic (i.e., boundedly rational) loss aversion); nice field experiment: Gneezy et al. (2003).
- Loss aversion as a concept that is difficult to get hold off: related to many other issues such as the endowment effect (disparity between willingness to pay and accept), opportunity costs, sunk costs, search behavior etc.). But is it robust?

Application: Time pressure and risky decisions - Trautmann and Kocher

- Many economic and financial decisions with stochastic outcomes have to be made under considerable time pressure.
- Not aware of any existing study in the economics literature that analyzes the effects of time pressure on decisions under risk in a systematic way.
- Provide empirical evidence on the impact of time pressure on a variety of risky decisions both in the gain and in the loss domain under different informational environments.

Application: Time pressure and risky decisions - Trautmann and Kocher

- Applications: financial markets, bidding in auctions, lots of other examples.
- Cues become more important; we learn more about economic decision-making (and theories of decisions-making) under risk in general.
- Added benefit: Consistency of risk measurement across different elicitation methods.

Time pressure: Related literature

- Literature on trade-off between speed and accuracy of decision-making (starting with Woodworth, 1899; Payne et al., 1993).
- Directly applicable literature with inconclusive results: Ben-Zur and Breznitz (1981) were the first to report less risk-taking under time pressure. However, Busemeyer (1993), for instance, has shown that this result depends heavily on the lottery's characteristics with respect to (the variance of) payoffs and probabilities.

Time pressure: Related literature

- In economics: Wilcox (1993), McDaniel and Rutström (2001), Sutter, Kocher and Strauß (2003), Kocher and Sutter (2006), or Rubinstein (2004). Influence of time pressure on negotiations (see, e.g., de Dreu 2003) or the effects of deadlines on economic decision-making in various contexts (for instance, auctions).

Time pressure: Basic experimental design

- 2x2-factorial design: time pressure and information provision.
- Laboratory experiment (why?).
- Sets of binary lottery choices in three independent parts of the experiment:
  - Part I: a set of independent risky gain prospects.
  - Part II: a set of simultaneous risky gain prospects according to the procedure used by Holt and Laury (2002).
  - Part III: a set of prospects including pure loss as well as mixed prospects which consisted of two subparts (Part IIIA and Part IIIB).

Time pressure: Basic experimental design

- Each choice consisted of two alternatives, denoted A and B. Prospects in the decisions problems were in the range of € -20 to € +200.

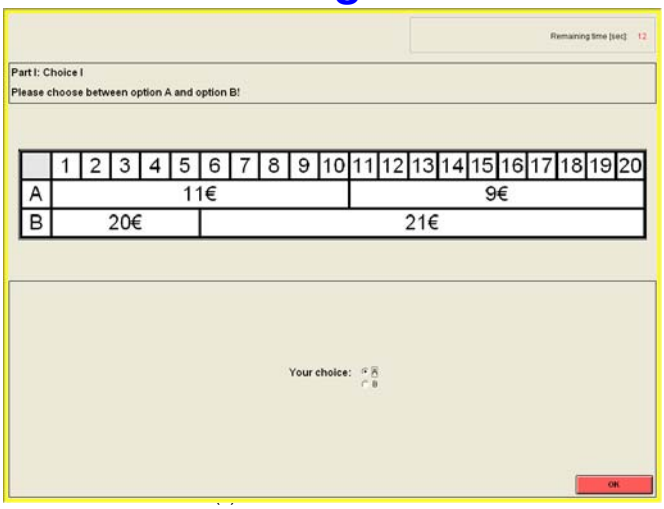
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	€20										€0									
B	€10																			

- At the end of the experiment, independently for each subject, one of the three parts was randomly selected for determining monetary payoffs. One choice was played for real through throwing a 20-sided die (random incentive lottery system).

Time pressure: Basic experimental design

- Part III: Decision problems in Part IIIA involved losses, pure gain prospects and mixed prospects. Part IIIB contained decision problems in the gain domain that yielded at least € 20 for any alternative. If Part III was chosen for real payoff, the subject would have to throw the die twice.
- Makes losses more salient and avoid a house money effect.
- Everything common knowledge.

Time pressure: Basic experimental design



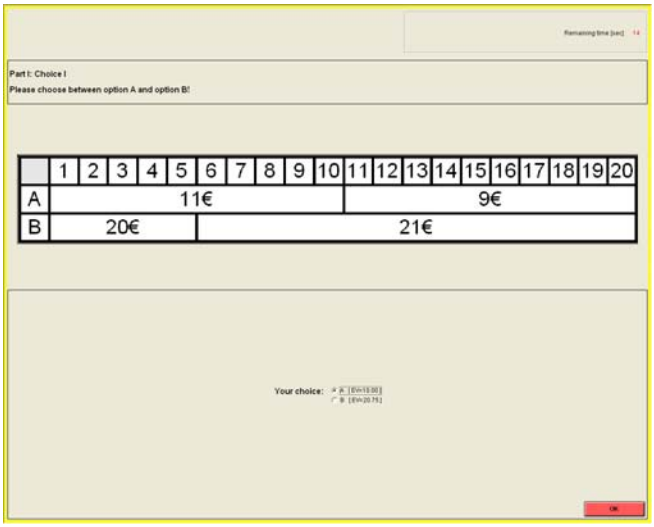
Time pressure: Treatments

Table 1: Overview of the Treatments

	No time pressure	Time pressure
No EV information	NTP-NEV N = 42	TP-NEV N= 41
EV information	NTP-EV N = 45	TP-EV N = 48

EV = expected value; N = number of observations.

Time pressure: Treatments



Time pressure: Treatments

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	12€				9.60€															
B	23.10€				0.60€															

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	12€							9.60€												
B	23.10€							0.60€												

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	12€											9.60€								
B	23.10€											0.60€								

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	12€														9.60€					
B	23.10€														0.60€					

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	12€																9.60€			
B	23.10€																0.60€			

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	12€																		9.60€	
B	23.10€																		0.60€	

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	12€																			9.60€
B	23.10€																			0.60€

Time pressure: Treatments

Table 1: Maximum and Actual Median Decision Time in Seconds

	No time pressure			Time pressure		
		No EV info	EV info		No EV info	EV info
	Max	Actual	Actual	Max	Actual	Actual
Part I	60	5.64	5.95	4	2.38	2.05
Part II[a]	150	59.5	71	30	29	26
Part IIIA[b]	60	5.87	5.95	4	2.42	2.47

The numbers are average medians in Part I as well as Part III and medians in Part II; EV info = expected value information.

[a] Part II decision time refers to total time for seven choices of the Holt and Laury-list.

[b] Data for Part IIIB (gain prospects which ensure that subjects cannot make losses in the end) are not provided because we do not use them in the results section. The time limit in this part was, however, identical to the time limit in Part I and Part IIIA.

What amount of time pressure? Why EV information?

Time pressure: Theoretical predictions

- Expected utility theory: increase in error; if a systematic effect on utility curvature, then symmetric in gain and loss domain.
- Prospect theory: Since PT involves (i) utility curvature, (ii) probability weighting, and (iii) gain-loss attitude several effects of time pressure are conceivable.

Time pressure: Theoretical predictions

- Aspiration level theories: $V(X) = \sum_{j=1}^{n} p_j u(x_j) + \mu P^+ - \lambda P^- \qquad \mu, \lambda \in R^+$

  Predicts risk aversion for the gain domain and risk seeking for the loss domain; for mixed gambles, it predicts loss aversion and gain seeking.

  Aspiration levels are the consequence of simplifying decision heuristics. Phenomena predicted by the model should become more pronounced under time pressure.

Time pressure: Results – some definitions

We analyze the effects of time pressure and expected value information on seven variables:

1. RAG: risk aversion for gains (counts the number of safe choices in six decisions between pure gain prospects each involving one sure gain).
2. RAGHL: risk aversion for gains using a Holt and Laury (2002) choice list with pure gain prospects (indicates the switching point).
3. RAL: risk aversion for losses (counts the number of safe choices in decisions between pure loss prospects each involving one sure loss).

Time pressure: Results – some definitions

4. RALMPS: risk aversion for losses considering choices between prospects and mean preserving spreads of these prospects.
5. PLA: avoidance of prospects with a prominent loss (counts the number of choices of pure gain prospects over mixed prospects with higher expected value).
6. PGS: seeking of prospects with a prominent gain (counts the number of choices of mixed prospects over a pure loss prospects with higher expected value).
7. ENDOW: safe choices in Part IIIB-decisions between prospects and their expected values used to endow subjects with at least €20 for the part involving losses.

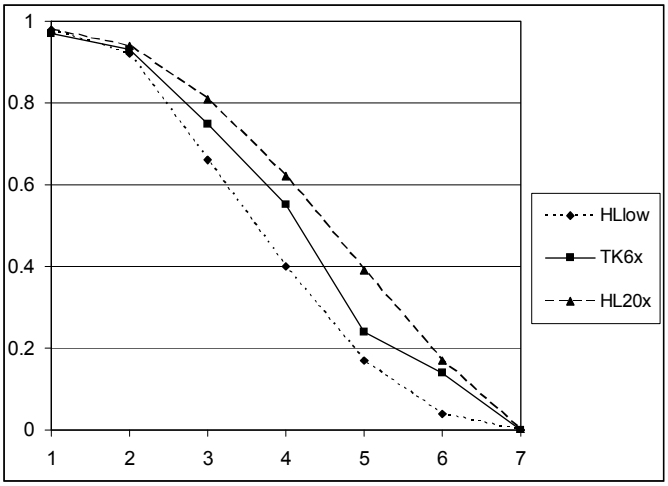
Time pressure: Results

- Time pressure manipulation: Highly significantly quicker decision-making. Correlation between our variables and decision times were practically zero for all variables in the treatments without time pressure. That is, there were not certain types of decision makers (in terms of our variables of interest) that were more constrained than others.

Time pressure: Results

- Consistency of risk attitude measures: Single choices are better than lists (like in Holt and Laury, 2002) to study the effects of time pressure.
- We find a significant positive correlation between RAG and RAGHL, which is strongest in the absence of time pressure as would be expected (no time pressure: Spearman's $\rho=0.49$, $p<0.0001$; time pressure: Spearman's $\rho=0.37$, $p<0.0038$).

Time pressure: Results



Time pressure: Results

- We consider the results of linear regressions (or ordered probit) of our variables on dummies for time pressure, the presence of expected value information and a dummy for gender.
- Gender has been found to influence risk and loss attitude. We replicate the findings of stronger risk aversion and loss aversion for females in simple gamble tasks. We do not find any effect on gain seeking.
- We also tested for interactions of time pressure and expected value information but there were no significant effects. Similarly there were no interactions with gender.

Time pressure: Results

Table 4: Linear Regression Results for Pure Gains

OLS	RAG	RAG	RAGHL	RAGHL	ENDOW	ENDOW
Time pressure	–0.011	–0.001	0.01	–0.054	–0.026	–0.032
	(0.035)	(0.051)	(0.038)	(0.055)	(0.039)	(0.057)
Expected value information	0.069	0.079	0.025	–0.024	0.011	0.006
	(0.035)	(0.05)	(0.037)	(0.048)	(0.039)	(0.055)
Time pressure × expected value		–0.019		0.121		0.011
		(0.071)		(0.075)		(0.079)
Female	**0.121**	**0.12**	0.059	0.062	**0.122**	**0.122**
	(0.037)**	**(0.038)** **	(0.039)	(0.039)	**(0.042)****	**(0.042)** **
# observations	172	172	146	146	170	170

Standard errors in parenthesis; ×: interaction; * significant at 5% level, ** significant at 1% level

Time pressure: Results

Table 5: Linear Regression Results for Pure Losses

OLS	RAL	RAL	RALMPS	RALMPS
Time pressure	**0.078**	**0.11**	**0.143**	**0.203**
	(0.039)*	**(0.056)***	**(0.057)***	**(0.082)***
Expected value information	0.0003	0.031	0.046	0.103
	(0.039)	(0.054)	(0.057)	(0.08)
Time pressure × expected value		–0.061		–0.114
		(0.077)		(0.113)
Female	**0.084**	**0.083**	0.109	0.106
	(0.041)*	**(0.041)***	(0.06)	(0.06)
# observations	171	171	173	173

Standard errors in parenthesis; ×: interaction; * significant at 5% level, ** significant at 1% level

Time pressure: Results

Table 7: Linear Regression Results for Mixed Prospects

OLS	PLA	PLA	PGS	PGS
Time pressure	**0.17** **(0.055)****	0.115 (0.08)	**0.253** **(0.043)****	**0.186** **(0.062)****
Expected value information	**– 0.12** **(0.056)***	**–0.173** **(0.078)***	**– 0.125** **(0.043)****	**–0.188** **(0.06)****
Time pressure × expected value		0.106 (0.11)		0.129 (0.086)
Female	**0.13** **(0.059)***	**0.133** **(0.059)***	–0.006 (0.046)	–0.003 (0.046)
# observations	172	172	168	168

Standard errors in parenthesis; ×: interaction; * significant at 5% level, ** significant at 1% level

Time pressure: Main findings

- No treatment effects for choices in the gain domain.
- Time pressure significantly increases risk aversion in the loss domain. No effects of expected value information with pure loss choices.
- Time pressure increases avoidance of losses and gain seeking (most salient cues) at the cost of a reduced expected value (exactly what aspiration level theories would predict). Expected value information mitigates the effect.

Three treasures about football and economics

Some fun applications of individual decision-making

1. Don’t bend it like Beckham. Myths and facts about penalty shootouts

Martin G. Kocher
Marc V. Lenz
Matthias Sutter

Introduction I

- Effort is not always increasing with stronger incentives.
- High rewards or the threat of severe punishment could impose pressure and lead to poor performance or complete failure instead.
- Such a perverse effect of incentives is denoted “choking under pressure” in social psychology.

Introduction II

- Reasons for choking: magnitude of stakes or rewards to be gained (Baumeister, 1985), the importance of achieving a success (Kleine, Sampedro and Lopes, 1988), expectations of negative consequences (Paulus, 1983), public expectations (Baumeister, Hamilton and Tice, 1985; Strauss, 1997 and1998), and the mere presence of others (Zajonc, 1965).
- Butler and Baumeister (1998) show that college students performed worst on difficult mental arithmetic tasks and computer games in front of a purportedly friendly audience as compared to performance in front of a neutral or even adversarial audience.

Introduction III

- Knowledge about the extent and the origins of choking is also interesting for labor economists in order to design optimal contracts and workplace environments.
- However, it is difficult to study in the working environment because variables are difficult to control for.
- Experimentalists would now propose a laboratory experiment (although pressure is difficult to induce), but an alternative is using field data from a non-working environment: in our case data on football.

Our approach

- We study determinants of successful penalty kicks in penalty shootouts.
- We take all penalty shootouts in the German cup competition (DFB-Pokal) from 1986 to 2006), i.e., 95 shootouts. As everywhere the German cup has a tournament structure. In case a professional club (from the first or the second league) plays against a club from the third or a lower league, the match always takes place at the stadium of the lower-ranked club.
- We have data on: round number, leagues the opponents are playing in, home-away team, order of penalties, success of penalties.

Related literature

- Dohmen (2007) examines penalties in the German premier football league (Deutsche Bundesliga).
- Two possible limitations: (i) regular penalty kicks might be subject to an endogeneity problem as subjects typically self-select into the task of kicking it; (ii) a large fraction of regular penalty kicks in the German Bundesliga are not relevant for finally winning or losing the match.
- Our data overcome these limitations to the greatest possible extent, and we can, additionally, address two prominent myths: whether it is important to shoot the first penalty in shootouts and whether there is a home team advantage.

Some infos and rules

- First, each team selects five players (out of the eleven players on the field).
- Then, it is randomly determined by the toss of a coin which team places the first penalty kick. During the whole shootout teams alternate in kicking.
- The shootout is terminated as soon as the number of penalty conversions of one team cannot be matched by the other team. For the first 10 kicks (five from each team) this means that the shootout is over whenever one team has scored more goals than the other could score, even if it were to complete its five kicks.

Some infos and rules

- If, after both teams have taken five kicks, both have scored the same number of goals, or have not scored any goals, kicks continue to be taken in the same order until one team has scored a goal more than the other from the same number of kicks. Each penalty kick during the shootout is taken by a different player and all eligible players must take a kick before any player can take a second kick.

Results for football fans

- The team that happens to kick first wins in 48%.
- Home teams win in 53% of all shootouts.
- The second finding requires some more discussion when compared to the 68% regular matches won by the home team during the season (excluding draws).
- It could be: (i) the effect of social pressure that especially effects home team players; or (ii) less influence of the referee (see third treasure).

Results on social pressure

- Overall conversion rate in cup (74.6%) vs. premier league (74.2%: Dohmen). Two effects might cancel out: Higher stakes (higher effort and better performance), higher social pressure (less performance).

Table 2. The determinants of scoring a goal in the shootout

	Model [1]	Model [2]	Model [3]	Model [4]
Included shootouts	95 (all shootouts)	91 (all with known attendance)	60 (all in 1st and 2nd round)	35 (all after 2nd round)
Home player	-0.059	-0.063	-0.029	-0.114
Goal difference	0.025	0.034	0.023	0.001
Decisive penalty	-0.128	-0.121	-0.447 **	0.533 *
Home * decisive	0.143	0.077	0.603 **	-0.707 *
Attendance (in 1000s)		0.0002		

* (**) $p < 0.1$ (0.05)

2. Shirt color and team performance in sports

Matthias Sutter
Martin G. Kocher

Introduction I

- Color of sportswear has a significant impact on the outcome of contests among individuals.
- Hill and Barton (2005) show that there is a consistent and statistically significant pattern in which contestants wearing red win more fights (Olympic Games 2004; boxing, tae kwon do, wrestling). Note that shirt colors (red or blue) are assigned randomly.

Introduction II

- The effect is stronger in close competitions and for men.
- Their explanation: Red coloration is a sexually selected, testosterone-dependent signal of male quality in a variety of animals, and in some non-human species a male's dominance can be experimentally increased by attaching artificial red stimuli. It correlates with human male dominance and testosterone levels. In humans, anger is associated with a reddening of the skin due; fear is associated with increased pallor in similarly threatening situations.

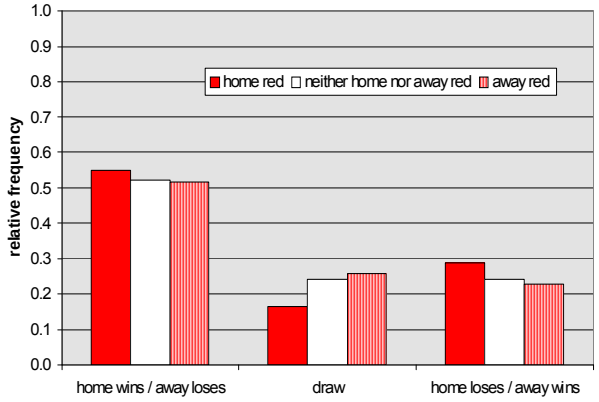
Introduction III

- Rowe et al. (2005) take the judo competition at the 2004 Olympics (blue and white coats) and find a winning bias for blue.
- Their explanation: Better visibility of white; therefore, contestants wearing blue can better anticipate attacks.
- Hill and Barton reject this argument in a reply (by referring to the gender effect).

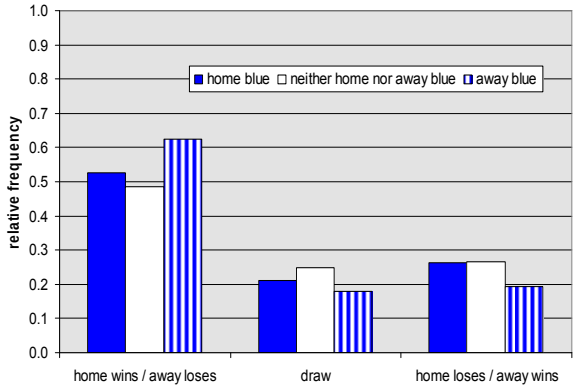
Introduction IV

- But what about team sports? Hill and Barton argue based on a very small sample (Euro 2004) that the same holds for football. We tested that claim for a sufficiently large sample.
- Season 2000/01 (306 football matches) in the German Bundesliga.
- The rule book: Home teams choose shirts first.

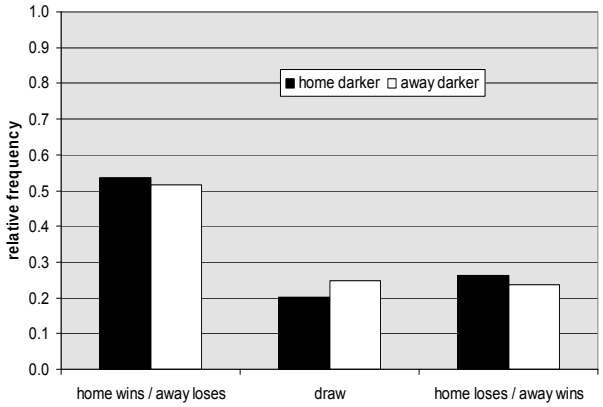
Results for shirt color I



Results for shirt color II



Results for shirt color III



Some further results

- No stronger effect in tight matches (defined by the rank in the competition before the match).
- Within-team comparison for away teams (ten out of 18 were wearing both red and non-red shirts during the season): Five out of the ten teams performed relatively better with red shirts; the other five teams performed worse with red shirts. On average, teams won 23% of their away matches in red shirts and 21% when their shirts were of a different color.
- Not enough observations for home teams (only three).

Possible explanations

- The results on individual contestants are less sound than they seem.
- The signal of a shirt's color is much less salient in team competitions.
- Team members feel much less intimidated subconsciously by an aggressive color like red when comforted at the same time by the presence and support of team-mates.
- From an endogenous perspective our result is quite obvious: Else all football clubs would be wearing red shirts at home.

3. Favoritism of agents – The case of referees' home bias

Matthias Sutter
Martin G. Kocher

Introduction I

- Football (soccer) referees are employed to interpret the rules in an impartial way.
- But they have considerable discretionary power.
- Consequently, referees have a very important influence on the final result of a football match.
- Economically speaking, it is a principal-agent relationship.

Introduction II

- Agency theory (Prendergast, 1999; Laffont and Martimort, 2002) is typically concerned with agents exerting more or less effort, supervision and information asymmetries.
- The problem of agents, like referees, making biased or impartial decisions has not been explored thoroughly so far.
- Favoritism = making biased or impartial decisions to one‘s favor.

Introduction III

- Examples in economics: hiring, promotion decision within a firm, selecting cooperation partners etc. with regard to gender, age, race, a country, a city etc.
- Often very difficult to assess because favoritism in decisions is difficult to measure in the field (largely unobservable variables like quality of job applicant).
- Remedy: Taking data from sports contests is interesting from an economic viewpoint but also interesting in itself.

Our study

- Provides an empirical test of favoritism in sports; some of the conclusions carry over to economics.
- We look at two different types of game-relevant discretionary referee decisions in the *German Bundesliga* in the season 2000/01:
  1. Amount of extra time added by referees at the end of a match (after the second half).
  2. Referee decisions on penalties.
- We test for a home bias of referees.
- Robustness checks.

Should referees exhibit a home bias?

- Well paid in professional football: in the *German Bundesliga* at that time € 3067.75 per game plus all expenses.
- Being biased in favor of the home team is not in the referee's private interest, because it leads, if detected, to be sacked by the DFB.
- Performance is monitored and evaluated by an official observer of the DFB referee committee; relegation to lower leagues is possible as a consequence of repeatedly poor evaluation.

Data I

- From the web-site of Kicker Sportmagazin.
- The web-site indicates for each match by the minute:

  the kick-off, goals, penalties, yellow or red cards, substitutions, and the final whistle of each half.
- A comment on whether awarded penalties have been legitimate and on whether a referee has refused to award a legitimate penalty (partly cross-checked).

Data II

- Season 2000/01 of the *German Bundesliga* with 306 games.

Selected average statistics per match:

- 2.93 goals, with home teams scoring 0.62 goals more than visiting teams. 161 matches won by the home teams, 76 by the visiting teams; 69 matches ended with a draw.

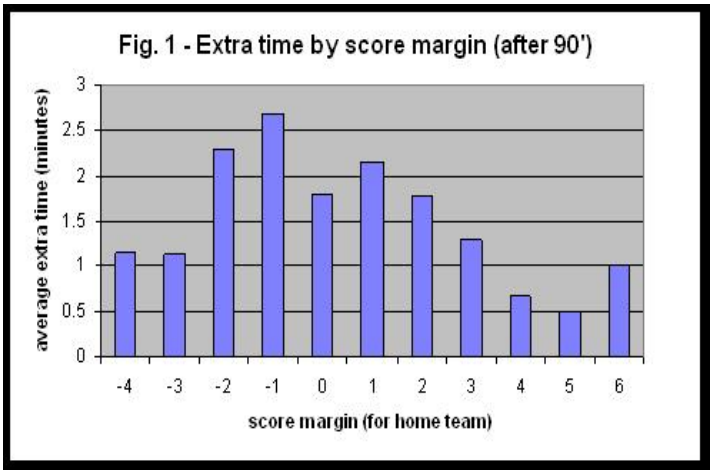
Data III

- Extra time: 1.88 minutes in the second half, 1.11 minutes in the first. Goals scored in extra time: 11.
- 76 penalties; 55 for the home teams, 21 for the visiting teams.
- Yellow cards (1.98 home teams; 2.40 visiting teams); red cards (0.1; 0.2)
- 2.65 substitutions per team; no significant difference between home and visiting teams.

Results I

- Average statistics are not very informative, because home teams are usually more aggressive and offensive than visitor teams.
- We need a more formal approach.
- Start by looking at extra time
- Lots of determinants of extra time according to the rule book: substitutions, injuries, wasting time/delaying the game etc.
- First impression: connection between extra time and score margin at the end of the 90th minute.

Results II



Results III – Controlling for other determinants

Dependent variable:	(1) extra time in **2nd half** score margin +1/–1 #	(2) extra time in **2nd half** score margin 0/–1 §	(3) extra time in **1st half** score margin +1/–1 #
Score margin	-0.53*	-0.86**	0.33~
	(0.25)	(0.28)	(0.17)
Yellow cards	0.15*	0.16*	0.04
	(0.06)	(0.07)	(0.07)
Red cards	0.01	0.45*	0.43
	(0.20)	(0.22)	(0.33)
Player substitution	0.14	0.01	0.41**
	(0.12)	(0.13)	(0.16)
Constant	1.29~	0.87	0.77**
	(0.66)	(0.69)	(0.20)
Adj. R^2	0.11	0.19	0.11
Durbin-Watson	2.17	1.64	1.85
N	109	112	134

Results IV - Interpretation

- OLS with White heteroscedasticity-consistent standard errors and covariances.
- If the home team is ahead by one goal (score margin of +1) or the result is a draw (score margin of 0), extra time is significantly shorter (about 30 to 50 seconds) than if the home team is behind by one goal (score margin of –1).

Results V - Interpretation

- If referees favor the home team, they should *not* add more extra time when the home team is behind by one goal *after the first 45 minutes.* They might even be tempted to add less extra time in that case in order to prevent more damage and to give the home team the opportunity to regroup in the break as soon as possible.
- The (ex post) probability of scoring a goal in one minute of extra time is 2%; only 10 goals have been scored in extra time in 306 games.

Size of effects and consequences

- Therefore, we looked at another potentially more important set of referee decisions: awarding penalties.
- In 2000/01 home teams were awarded 55 penalties, but visiting teams only 21. But there is no reason to assume that both instances are equally likely.
- What we need is the ratio of legitimate and awarded penalties to legitimate but refused penalties. If referees are impartial, the ratio should be the same for home and visiting teams.

Results VI

- 55 penalties for home teams; 5 classified as illegitimate; thus 50 legitimately awarded penalties; 12 refused penalties
- 21 penalties for visitor teams; 1 classified as illegitimate; thus 20 legitimately awarded penalties; 19 refused penalties.
- Whereas in 50 out of 62 cases (81%) the home team is awarded a penalty which is legitimate, visiting teams are awarded a legitimate penalty only in 20 out of 39 cases (51%). ($\chi^2 = 9.7$; $df = 1$; $p < 0.01$), indicating a clear home bias of referees with regard to awarding penalties.

Discussion I

- We cannot control for wasting time/delaying the game. If anything a home team leading by one goal should delay the game, which means that extra time should actually be longer, on average, when the home team is ahead by one goal. Since it is lower, we have a strong argument in favor of our conclusion!
- Possible reasons:

  1. Referees want to please the crowd (subconsciously).

  2. Inequity aversion.

  Or probably both.

Discussion II

- Budgets of the teams, rank or rank difference in the league before the game does not influence extra time.
- Dummies for certain clubs are also insignificant.
- Attendance influences extra time positively and significantly (the influence of the crowd).
- Similar result by Dohmen (2005): bias becomes smaller with more visiting team supporters in the stadium and in stadiums with running tracks that divide the field from the supporters.

Robustness check for extra time

	Italy: Serie A (N – 1228)	Italy: Serie B (N – 1696)	U.S.A.: MLS* (N – 159)	Spain: La Liga** (N – 750)	Germany: Bundesliga (N – 306)
av. extra time in 2nd half	3.67 min	3.65 min	3.19 min	2.93 min	1.88 min
Marginal effect of score diff -1 after 90' on extra time	0.31 min	0.38 min	0.64 min	1.03 min	0.80 min

* "Do soccer referees display home team favouritism?", Lucey and Power (2004)

** "Favoritism under social pressure", Garicano, Palacios-Huerta and Prendergast (2005).

Related evidence and open questions

- Nevill, Balmer and Williams (2002): referees are affected by the crowd's noise when deciding over tackles
- Dohmen (2005): similar findings for biased penalty decisions
- Scoppa (2007): confirms findings for Italy and can also significantly show an effect of the referee scandal.
- Reason for the referees' behavior
- Other important decisions of football referees
- Similar examinations for other sport contests

Workshop

Economics and Psychology of Football 2008

29.-30. May 2008

University of Innsbruck

Organized by: Patric Andersson, Martin Kocher, Carsten Schmidt, Matthias Sutter