Darai, Donja; Grätz, Silvia

Working Paper

Facing a dilemma: Cooperative behavior and beauty

Working Paper, No. 82

Provided in Cooperation with:
Department of Economics, University of Zurich

Suggested Citation: Darai, Donja; Grätz, Silvia (2012) : Facing a dilemma: Cooperative behavior and beauty, Working Paper, No. 82, University of Zurich, Department of Economics, Zurich, http://dx.doi.org/10.5167/uzh-63130

This Version is available at:
http://hdl.handle.net/10419/77493

University of Zurich

Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 82

Facing a Dilemma: Cooperative Behavior and Beauty

Donja Darai and Silvia Grätz

Revised version, December 2012

Facing a Dilemma: Cooperative Behavior and Beauty*

Donja Darai and Silvia Grätz†

December 7, 2012

Abstract

This paper investigates the influence of physical attractiveness on cooperative behavior. We survey data from 211 episodes of a television game show and combine it with independent facial attractiveness ratings of the show's contestants. The final of the show represents a simultaneous one-shot modified prisoner's dilemma. Our results show that facially attractive contestants provoke cooperative behavior from their counterparts, but attractive contestants themselves do not behave differently from unattractive ones. This preferential treatment or "beauty premium" rewards attractive contestants with substantially higher monetary gains of up to £2 153. This finding applies to both sexes. But, the attractiveness effect is limited to mixed-gender interactions. Contestants are only more cooperative when facing an attractive fellow player of the opposite sex. Attractiveness also has no effect in group-decisions made by the contestants prior to the prisoner's dilemma. Therefore people seem to show a preference to cooperate with someone towards whom they are personally attracted; and this preference reaches full effect when people lack other information on which to base their decision.

Keywords: beauty premium, cooperation, prisoner's dilemma

JEL classification: C71, D83, Z13

*The authors thank Ulrich Kaiser, Armin Schmutzler, Michelle Sovinsky, Roberto Weber, the seminar participants at the University of Zurich and Florida State University, the participants at the ESA International Meeting 2012 in New York (U.S.), the Workshop on Behavioral Economics Mainz (Germany), the Zurich Workshop on Economics 2012 in Luzern (Switzerland), and the ESA European Meeting 2012 in Cologne (Germany) for helpful discussions and suggestions, as well as Pascal Kappeler, Christina Richard, and Holger Scriba for excellent research assistance. Financial support of the Swiss National Science Foundation is gratefully acknowledged. The data on "Golden Balls" were provided to the authors by the television show producers, courtesy of Endemol UK plc, in May 2009.

†Both authors, Department of Economics, University of Zurich, donja.darai@econ.uzh.ch, silvia.graetz@econ.uzh.ch

"All that glitters is not gold;
Often have you heard that told."

William Shakespeare, Merchant of Venice (1596-1598)

1 Introduction

Beginning with the seminal paper of Biddle and Hamermesh (1998), which identifies a wage gap based on physical attractiveness using labor market data of the U.S. and Canada, economists started to be interested in the effects of physical attractiveness on economic decision making. For instance, Mobius and Rosenblat (2006) show that the beauty premium for attractive people is even present in a controlled laboratory labor market experiment. Eckel and Petrie (2011) provide evidence that people have a willingness to pay for being able to see a picture of their counterpart before making their decision in a trust game, suggesting that valuable information is conveyed by the physical appearance of people. However, the transmission channel of the effect of physical attractiveness on economic decision making is still only partly understood. The most prominent theory is "beauty-is-good" stereotyping. People assign a wide range of positive personality traits and abilities to physically attractive people. That is, people believe that physically attractive people are, for instance, more trustworthy, more likable, and more productive (see e.g., Dion, Berscheid, and Walster (1972)). These beliefs then alter people's behavior towards attractive people. However, *not all that glitters is gold*. There is only very limited evidence that the "beauty-is-good" stereotype is accurate.

This paper studies the relationship between attractiveness and cooperative behavior and shows that the beauty premium is also present in a high-stakes field setting with two-sided communication and previous interaction between players. Our results offer an alternative explanation of the effects of attractiveness on economic decision making. We suggest a preference based transmission channel in the sense that people's behavior is biased when facing someone attractive of the opposite sex. We confirm our hypothesis with existing experimental data from Mobius and Rosenblat (2006).

Specifically, this paper combines two independent data sets. First, to study cooperation, we survey data from 211 episodes of the British television game show "Golden Balls", in which the show's finalists play for a jackpot at the end of the show. The rules of this game follow a slightly modified version of a simultaneous one-shot prisoner's dilemma: if both contestants choose to cooperate, the jackpot is split equally; if one chooses to defect while the other cooperates, the former receives the entire jackpot and the latter gets nothing; if both defect, they both go home empty-handed. The accumulation of the jackpot takes place in two rounds of pre-play previous to the prisoner's dilemma and requires neither the contestants' effort nor cognitive ability. On average, the stake size is £12 912 indicat-

ing that decisions in the prisoner's dilemma have substantial distributional consequences for the contestants. During the two rounds of pre-play, the two finalists are selected out of four initial contestants. Second, we collect data on the contestants' facial appearance from a sample of 365 independent raters. These raters are asked to evaluate portrait photographs of the contestants along various dimensions such as physical attractiveness, sympathy, trustworthiness, or likability.

We provide evidence that physical attractiveness affects cooperative behvaior. Contestants of the show are significantly more cooperative towards a facially attractive opponent than towards a facially unattractive opponent, independent of gender, age, other demographic characteristics, stake size, and communication. Facially attractive contestants are therefore awarded with significantly higher earnings in the prisoner's dilemma. Quantifying this beauty premium, a one-standard deviation increase in facial attractiveness, at the margin, causes the contestant's expected earnings to rise by up to £2 153, as well as the contestant's probability to obtain positive earnings by up to 5.9 percentage points. This effect is not driven by non-cooperative behavior of the attractive. With minor qualifications for younger and female contestants, we find no correlation between a contestant's own facial attractiveness and cooperation.

Although the attractiveness effect likewise applies for men and women, we show that it is limited to mixed-gender interactions only. Contestants are only biased in their decision to cooperate when facing an attractive opponent of the other sex. Thus, the attractiveness effect is not present in same-gender interactions. This finding offers a new perspective on the underlying mechanism of the effects caused by attractiveness. Our results are not driven by people believing that attractive others are more likely to cooperate, but rather by people having a preference to cooperate more with someone towards whom they are personally attracted to. Furthermore, we cannot identify a beauty premium or plainness penalty during the pre-play, which clearly supports our conjecture that personal attraction serves as the underlying transmission channel of physical attractiveness on people's behavior in pairwise interactions and that attractiveness matters most when objective information is missing.

To validate our results, we repeat the analysis of Mobius and Rosenblat (2006) and uncover that their identified effects of attractiveness disappear when we parse their data by the composition of sex, that is, when we consider interactions between employers and workers of the same sex separately from those of the opposite sex. Again, "beauty-is-good" stereotyping fails to consistently explain the findings. Therefore, we offer an alternative mechanism underlying the attractiveness effects, namely a "preference-driven" mechanism. Physical attractiveness mediates people's behavior, because people want to behave differently towards someone whom they feel personally attracted to.

The remainder of this paper is structured as follows. In Section 2 we discuss related

literature. Section 3 describes the data sets and Section 4 presents the results. Potential transmission channels are evaluated in Section 5. Finally, Section 6 concludes.

2 Literature

A long history of research on physical attractiveness in psychology, sociology, and evolutionary biology and a more recent one in economics and political science shows that physically attractive people are or behave not substantially differently from physically unattractive people, but that physically attractive people receive a preferential treatment by others in many regards. This "beauty premium" rewards attractive people with substantially higher monetary gains or greater economic success. Surprisingly, the mechanism underlying the effect of physical attractiveness is still not completely uncovered.

The most prominent theory, established in sociology and psychology, is that people link beauty and positive personality traits. In particular, people attribute a variety of positive characteristics and higher abilities to physically attractive people, and negative ones to physically unattractive people (see reviews by Eagly, Ashmore, Makhijani, and Longo (1991), Feingold (1992), and Langlois, Klakanis, Rubenstein, Larson, Hallam, and Smoot (2000)). This is referred to as *"beauty-is-good" stereotyping* in the literature.[1] The theory argues that stereotype beliefs cause people to treat attractive ones more favorably, and this in turn results in higher economic gains for the attractive. Hence, beauty-is-good stereotyping can explain why attractive people generally fare better in the labor market, i.e., why they are more likely to be hired, promoted, and earn higher salaries. In their seminal work Biddle and Hamermesh (1998) use a broad household survey of the U.S. and Canada and show that physically attractive employees earn about 10-15 percent more than less attractive employees, independent of their occupation. In an experiment on the effects of beauty in experimental labor markets, Mobius and Rosenblat (2006) find that attractive employees are awarded higher wages, although they are not more productive than unattractive employees, measured by a real-effort task. This suggests that employers hold (inaccurate) stereotypical expectations about the performance of physically attractive employees. Ruffle and Shtudiner (2010) explore the value of beauty in the hiring process, examining response rates to CVs sent to companies in Israel. They uncover a gender-specific effect of attractiveness, namely a beauty premium for males, and a beauty penalty for females.

[1] In their pioneering study, Dion, Berscheid, and Walster (1972) claim "what is beautiful is good" by demonstrating that attractive people are believed to have better career prospects, to possess socially desirable traits, to lead happier lives and to be happier overall. This paper spawned a large literature on the physical attractiveness stereotype, demonstrating for both sexes a robust association between physical attractiveness and cognitive ability, competence, sociability, popularity, dominance, sexual experience, mental health, and social skills.

Also, research in political science shows that physical attractiveness has a significant impact on the evaluation of candidates and thereby on electoral outcomes (see the survey by Ottati and Deiger (2002)). For instance, attractiveness increases the candidate's number of votes and thereby the likelihood of being elected (Berggren, Jordahl, and Poutvaara (2010) and Rosar, Klein, and Beckers (2008)). Antonakis and Dalgas (2009) suggest that underlying the effect of physical attractiveness is the voters' belief that attractive candidates are more competent. Recruiting adults and children raters in Switzerland, they find that even children can predict results of the 2002 French parliamentary election retrospectively by rating the competence of the candidates from their photographs.

Furthermore, the theory of beauty-is-good stereotyping in combination with interdependent social preferences can explain findings in various experimental settings with strategic interactions. These findings suggest that people behave more generously or cooperatively towards physically attractive people. Solnick and Schweitzer (1999) show that attractive responders receive significantly higher offers than unattractive responders in the ultimatum game. In dictator games, Rosenblat (2008) shows that allocators treat physically and vocally attractive recipients more generously. Andreoni and Petrie (2008) observe a beauty premium in a repeated public goods game, which is not caused by lower contributions of attractive players, but can be attributed to an increase of the other players' contributions triggered by an attractive group member. In a repeated trust game, Wilson and Eckel (2006) demonstrate that attractive trustees are trusted more and that players expect attractive players to be more trustworthy than unattractive ones. The failure to meet these expectations leads to a beauty penalty.[2] Mulford, Orbell, Shatto, and Stockard (1998) examine a repeated prisoner's dilemma, pointing out that players are more cooperative if matched with an attractive partner. In contrast to the other studies, they use attractiveness ratings based on the perception of the player who also makes the decision in the experiment, rather than ratings elicited from independent third-party judges.

However, all above mentioned studies provide only very limited evidence for the accuracy of the beauty-is-good stereotype. Physically attractive people themselves do not behave differently than physically unattractive people.[3]

Besides the theory of stereotyping based upon peoples' visual attractiveness, Zuckerman and Driver (1989) and Zuckerman, Hodgins, and Miyake (1990) proclaim the existence

[2]In another experimental study on the trust game, Eckel and Petrie (2011) investigate the informational value of a photograph and the differential desire to acquire this information. Subjects are willing to pay to see the photograph of their partner whom they transact with, indicating that a face has a positive informational value which is used to discriminate between players in their choices.

[3]Jackson, Hunter, and Hodge (1995) show in their meta-analysis that there is a modest correlation between attractiveness and intelligence for children. Mueller and Mazur (1997) use data from a cohort of military officers and find that recruits with a high ranked facial appearance are also high ranked in their physical fitness. Concerning social skillfulness and likability, Goldman and Lewis (1977) and Erwin and Calev (1984) find that physically attractive people indeed possess better social skills and are more likable.

of a *vocal attractiveness stereotype* by showing that physical attractiveness is positively correlated with vocal attractiveness. Supporting a vocal-attractiveness mechanism, Mobius and Rosenblat (2006) find that the beauty premium even exists in treatments in which employers are able to talk to the employee without seeing the employee's picture. Rosenblat (2008) provides evidence that physically attractive people only achieve significantly better outcomes in dictator games when the dictator sees the recipient's picture in addition to hear her pre-recorded voice. This suggests voice to be one of the transmission channels of the beauty premium, at least in two-sided communication settings.
Finally, one of the oldest and most well-known theories underlying the effects of attractiveness is the one of *taste-based discrimination*, proposed by Becker (1957). Physically attractive people are favored because people enjoy being or working with them more than with plain looking people. Discrimination is based upon prejudices correlated with people's personal characteristics and is rational in the sense that interactions with such a person generate a (dis)utility for the discriminator in case of positive (negative) discrimination. Belot, Bhaskar, and van de Ven (2012) find that attractive contestants of a Dutch television game show are positively discriminated against unattractive ones in proceeding to the final stage of the show.[4] However this theory fails to explain why attractive people are also treated more favorably in one-shot interactions.

Summarizing, in line with the literature we expect physically attractive people to obtain a beauty premium in the one-shot prisoner's dilemma. Physically attractive people will be favored by others in the sense that people show greater cooperativeness towards them. But physically attractive people themselves will not behave differently compared to physically unattractive people.

3 Data

3.1 The television game show "Golden Balls"

We analyze people's behavior in a simultaneous one-shot prisoner's dilemma of 211 episodes of the British television game show "Golden Balls".[5] The dilemma is played in the final

[4] In particular attractive contestants are much more likely to reach the final round of the show, even though they are not performing better or are not more confident than unattractive contestants. Performance in this show means being the first to correctly answer trivia questions. Attractive contestants are believed to be more confident and to be more cooperative. Besides, at the end of the show a prisoner's dilemma – like the one we study – is played. Belot, Bhaskar, and van de Ven (2012) also test for effects of attractiveness on cooperative behavior, however they find no significant effects. We suspect that the reason why they are not able to identify any attractiveness effects is that their sample of finalists becomes too homogeneous with regard to attractiveness, due to the selection bias towards attractive contestants during the pre-play.

[5] The game show was aired in June 2007 and ended in December 2009. We use records of 211 episodes of the first four series of the show.

round of the show to allocate the jackpot between the two finalists: each player is assigned two balls, indistinguishable from the outside, but one contains the word "steal" and one contains the word "split". Both contestants choose one of their two balls and then open it simultaneously. If both choose the split-ball (cooperation), the jackpot is split equally; if one chooses the steal-ball (defection) while the other chooses the split-ball (cooperation), the former receives the entire jackpot and the latter gets nothing; if both choose the steal-ball (defection), they both go home empty-handed (see Figure 1). Thus, defection is a weakly-dominant strategy.

Figure 1: Prisoner's dilemma game

	split (cooperate)	**steal** (defect)
split (cooperate)	$^1/_2$ jackpot , $^1/_2$ jackpot	0 , jackpot
steal (defect)	jackpot , 0	0 , 0

Note: Defection is a weakly dominant strategy

Before the prisoner's dilemma is played, the contestants have to pass two rounds of pre-play. In these rounds, the finalists are selected and the jackpot is accrued. Accumulating the jackpot does not require contestant's cognitive ability or effort.[6] On average, the jackpot amounts to £12 912 and ranges from a minimum of £3 to a maximum of £93 250. The two finalists are selected through two voting decisions. After each round, the contestants need to cast a vote against one of them to leave the show. The contestant who receives the majority of votes is eliminated. Throughout the show the contestants face each other and are allowed to freely communicate with each other (see supplementary material for a detailed description of the game show).

Table A.1 in the appendix presents summary statistics of the outcomes of the prisoner's dilemma and of the contestants' personal characteristics. The unilateral cooperation rate is 54%, and contestants mutually cooperate (defect) in 32% (25%) of cases. 46% of finalists are male and the mean age is 37 years.

3.2 Survey on facial appearance

We evaluate the contestants' facial appearance using a panel of independent raters. The raters are recruited at the Euro-Airport Basel, at the University of Zurich, and at the

[6]Each pre-play round, the contestants are randomly assigned a certain cash value. For the detailed description of the show, see online appendix.

University of Zurich for the elderly.[7] All 844 contestants are judged by 728 raters and, from those, 365 raters judge the 422 finalists. Each rater was asked to individually rate the facial appearance of five randomly assigned contestants, of which two or three were male (female). On average, a finalist is judged by 4.3 raters.
Table I reports summary statistics for the finalists' raters and Table A.2 in the appendix for all contestants' raters. The mean age of the 365 raters is 41 years and 50% are male.

Table I: Summary statistics of the finalists' raters

Rater's Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Male	0.5	0.5	0	1	365
Age (in years)	40.88	15.54	17	93	361
Age of Male (in years)	41.58	15.60	17	93	183
Age of Female (in years)	40.15	15.50	18	86	178
Age of Female ($\geq$ 40.15 years)	53.51	10.36	41	86	86
Age of Female ($<$ 40.15 years)	27.66	6.38	18	40	92
Age of Male ($\geq$ 41.58 years)	55.21	11.42	42	93	85
Age of Male ($<$ 41.58 years)	29.76	6.28	17	41	98

The survey is questionnaire based. For an illustration of a sample questionnaire see Figure 2. Each questionnaire contains two portrait photographs of the same contestant and is divided into three parts. To receive non-biased evaluations of the contestants and to reduce measurement error, all photographs are selected from the same two sequences of the game show such that one photograph shows a neutral facial expression with a view to the camera and the other a neutral facial expression with a view to the side of the camera.[8]
In the first part of the questionnaire, the raters are asked to judge the contestant with respect to her age by fitting the contestant into one of seven age categories, "<20", "20-30", "30-40", "40-50", "50-60", "60-70", or ">70" years. The second part includes assessments of the contestant's appearance using four opposite word pairs, i.e., "attractive - unattractive", "likable - unlikable", "trustworthy - untrustworthy", and "honest - dishonest". These items are rated on a 1-to-7 point Likert scale, where 1 equals very unattractive, 4 comprises a neutral position and 7 equals very attractive.[9] In the last part of the questionnaire we asked the raters to give a binary response (yes/no) to two statements: "this person's appearance helps him/her in life" and "this person strikes me as calculating". These statements are included to capture a rater's overall impression of the respective contestant's

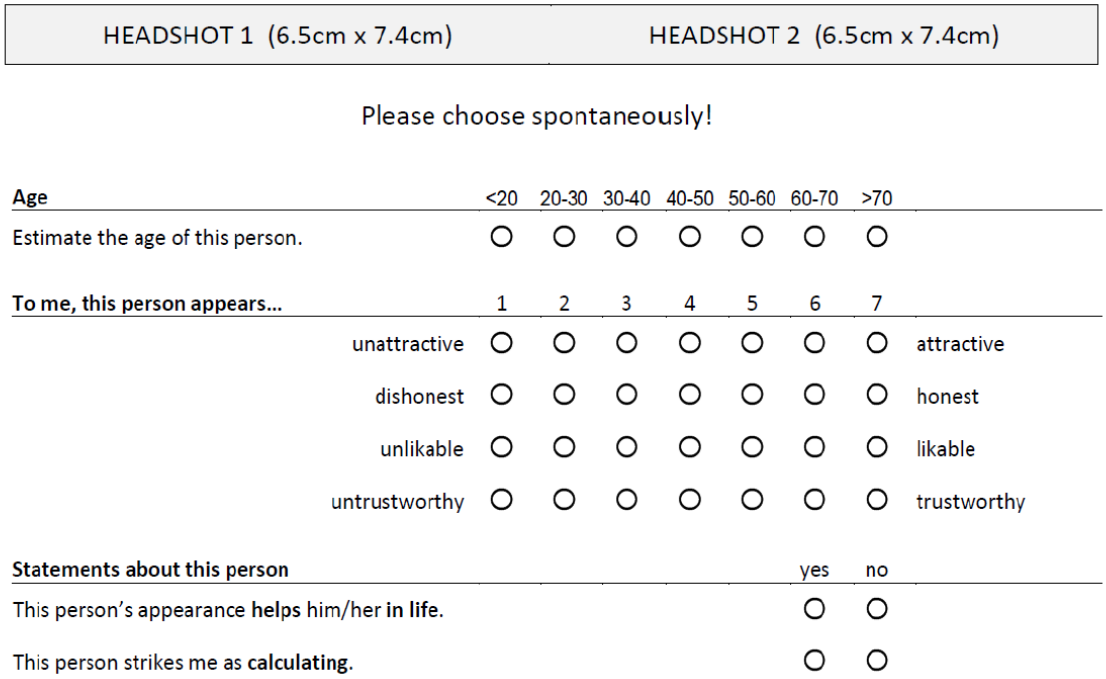
[7]The University of Zurich provides lectures for a senior audience, that are mainly attended by retired people. We only recruited raters who either had a very high proficiency or were native speakers of German, French, or English and provided the questionnaires in the respective language.

[8]If possible we chose a neutral facial expression of the contestant, otherwise a positive one was chosen, but never a negative or disadvantageous one.

[9]The use of a 7-point Likert scale, which includes a neutral element, allows for sufficient diversification and is standard in the relevant literature, see e.g., Alreck and Settle (1995), pp.113-114.

appearance, i.e., whether a contestants is either attributed a beauty premium or a strategic intention per se.

Figure 2: Sample Questionnaire



3.3 Evaluation of facial appearance ratings

The raters' evaluations are used to construct a facial appearance measure for each of the four items: attractiveness, honesty, likability, and trustworthiness. We account for rater specific variation in the perception of the respective appearance item in the construction of each measure.[10] For each rater j we calculate the average across all contestants rated by j, $\bar{x}_j$. Then we mean-center the ratings by subtracting the respective rater's mean rating, $\bar{x}_j$, from her individual rating of contestant i, x_{ij}. These demeaned ratings are

[10]Raters may differ in the interpretation of the 7-point scale, and may vary in anchoring their average rating above, below or close to 4. Student t-tests of differences between the means of the facial appearance measure by the raters' gender are: $t = 2.799$, $p = 0.005$ for attractiveness, $t = 4.746$, $p = 0.000$ for likability, $t = 4.913$, $p = 0.000$ for trustworthiness, and $t = 4.591$, $p = 0.000$ for honesty. This shows that female raters tend to give higher ratings than male raters. This phenomenon is also persistent with respect to the raters' age. Further, it is known that images of females tend to be rated higher than images of males, see e.g., Jackson, Hunter, and Hodge (1995) and Wilson and Eckel (2006). Since we control directly for gender-related effects in the regressions, the attractiveness measure can be used uncorrected for the gender of the contestants. As a robustness check we construct a measure by making use of the rater-specific fixed effects, rather than correcting for them. In particular, we attach higher weight to ratings made by raters that are similar to the contestant with respect to age and gender, and we attach lower weight to ratings made by raters that are unsimilar to the contestant. All results remain unchanged.

now anchored at 0 for each rater and are therefore corrected for rater-specific fixed effects (see Figure A.1 in the appendix). Finally we take the mean of all mean-centered ratings of contestant i, resulting in the particular facial-item measure for each contestant.

$$[\text{Facial-item}]_i = E_j[x_{ij} - \bar{x}_j] \quad \text{with} \quad \begin{cases} i = & \text{i-th contestant} \\ j = & \text{j-th rater} \end{cases}$$

Additionally, we construct two variables from the binary statements "appearance helps in life" and "appearance strikes as calculating" by taking the mean rating for each contestant.[11] In the following we refer to these variables as "statement variables". Table II provides summary statistics of all appearance measures.

Table II: Summary statistics facial appearance

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Attractiveness (mean-centered, cont.)	0	0.79	-1.8	2.65	422
Honesty (mean-centered, cont.)	0	0.6	-1.8	1.45	422
Likability (mean-centered, cont.)	0	0.66	-1.95	1.65	422
Trustworthiness (mean-centered, cont.)	0	0.6	-1.75	1.65	422
Appearance Helps In Life (cont.)	0.58	0.30	0	1	422
Appearance Strikes As Calculating (cont.)	0.35	0.25	0	1	422

All four facial appearance variables are highly positively correlated (see Table III). This is also reflected in a sufficiently high Cronbach coefficient alpha ($\alpha = 0.82$).[12]

Table III: Correlation matrix for facial appearance variables, N=422

	Attractiveness	Honesty	Likability	Trustworthiness	Helps in Life
Attractiveness[a]	1				
Honesty	0.286***	1			
Likability	0.461***	0.651***	1		
Trustworthiness	0.340***	0.760***	0.680***	1	
Helps in Life[b]	0.643***	0.229***	0.389***	0.299***	1
Strikes as Calculating	0.007	−0.313***	−0.288***	−0.315***	0.034

[a] Mean-centered, continuous variables (for attractiveness, honesty, likability, trustworthiness).
[b] Averaged, continuous variables (for "appearance helps in life", "appearance strikes as calculating").

[11] Again, there may be differences between the means of the statement variables' ratings by the raters' gender. Student t-tests w.r.t. gender are $t = 0.851$, $p = 0.395$ for "appearance helps in life" and $t = -5.437$, $p = 0.000$ for "appearance strikes as calculating", i.e., male raters are more likely to rate a contestant as calculating than female raters.

[12] We use Cronbach's alpha for standardized variables to measure the inter-item reliability for facial appearance. The measure adjusts for item specific mean and variance. Also a nonparametric test for testing whether samples originate from the same distribution cannot be rejected (Kruskal-Wallis $K = 0.9868$).

The statement variable "appearance helps in life" is also positively correlated with all four facial appearance variables. However, the statement variable "appearance strikes as calculating" is not correlated with attractiveness, and even negatively correlated with the remaining facial appearance variables.

In the following, we focus on facial attractiveness, since it is a crucial part of the first impression of a person and a stable characteristic which is almost impossible to mimic (see e.g., Grammer, Fink, Møller, and Thornhill (2003)). The two statement variables are used as additional controls.
There are many ways to define an attractive person. We use the following classifications: First, we classify a contestant as *facially attractive* if her facial attractiveness rating lies above or is equal to the mean over all facial attractiveness ratings, and as *facially unattractive* if her facial attractiveness rating lies below this mean. The average mean-centered rating of facially attractive contestants is 0.599, and the one of facially unattractive contestants is -0.650. We also used the median over all facial attractiveness ratings as a classification device. The average mean-centered rating of a facially attractive contestant, who is rated above or equal to (below) the median is 0.613 (-0.635). Second, we define extreme measures of facial attractiveness. A contestant is classified as *most attractive* if her facial attractiveness rating lies within the top 10% (25%) percentile of the distribution of facial attractiveness and as *least attractive* if her facial attractiveness rating lies within the bottom 10% (25%) percentile. Contestants who are rated as most attractive receive, on average, a mean-centered rating of 1.389 (0.993), and those who are rated as least attractive receive a mean-centered rating of -1.386 (-1.003). For an illustration of the distribution of facial attractiveness see Figure A.1 in the appendix.

Table IV: Summary statistics facial attractiveness

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Attractiveness (mean, d)	0.52	0.50	0	1	422
Attractiveness (median, d)	0.51	0.50	0	1	422
Most attractive (90% percentile, d)	0.11	0.31	0	1	422
Most attractive (75% percentile, d)	0.25	0.44	0	1	422
Least attractive (10% percentile, d)	0.10	0.30	0	1	422
Least attractive (25% percentile, d)	0.25	0.43	0	1	422
Helps in Life (d)	0.50	0.50	0	1	422
Strikes as Calculating (d)	0.47	0.50	0	1	422

(d) for dummy variable.

Third, we define a contestant's *appearance to be helpful in life* (*to strike as calculating*) if her "helps in life"-rating ("strikes as calculating"-rating) lies above or is equal to the mean over all "helps in life"-ratings ("strikes as calculating"-ratings), and as *not to be helpful in life* (*not to strike as calculating*) if her "helps in life"-rating ("strikes as calculating"-

rating) lies below this mean. Table IV summarizes the binary attractiveness and statement variables. A detailed description of the distribution of attractive and unattractive finalists with respect to gender and age is provided by Table A.3 in the appendix.

4 Results

4.1 Facial attractiveness

In order to investigate the relationship between cooperative behavior and facial attractiveness we use several binary probit models with the decision to cooperate as the dependent variable (with $y_i = 1$ equal cooperate; $y_i = 0$ equal defect). Throughout the analysis, we control for effects and interactions related to the contestant's gender and age, other demographic characteristics, as well as variables of stake size, communication, and variables describing the course of events of the game show previous to the prisoner's dilemma (pre-play), see Table A.1 in the appendix.

Own attractiveness The results depicted in Table V, model (1) to (5), show that facially attractive contestants do not behave differently with respect to cooperativeness than facially unattractive contestants, independent of the specification of the attractiveness measure.[13] This finding is in line with our conjecture from the literature (see e.g., Eagly, Ashmore, Makhijani, and Longo (1991)): physically attractive people are not more pro-social and therefore not more cooperative than physically unattractive people.
Whereas, overall, we find no difference between the cooperative behavior of attractive and unattractive people, there are some qualifications with respect to gender and age.[14] Both gender and age seem to mediate the effect of a contestant's own attractiveness on

[13]Table V reports the regression results including a dummy variable for the attractive contestant. For robustness of all our results, we also estimate the regressions including (i) the continuous measure for attractiveness, (ii) the mean over the four appearance variables, (iii) the predicted factors obtained in a confirmatory factor analysis of the four appearance variables, (iv) a normalized attractiveness measure in line with Mobius and Rosenblat (2006), in which our constructed attractiveness measure is normalized across all contestants, (v) a normalized attractiveness measure in line with Biddle and Hamermesh (1998), in which the normalization is across all raters, and (vi) the mean and median attractiveness ratings of the raw data, where the mean (median) is 0.443 (0.728) with an average rating of facially attractive contestants above the mean (median) of 5.552 (4.914), and below the mean (median) of 3.196 (2.357). All measures produce qualitatively the same results.

[14]Irrespective of a contestant's own attractiveness, we find a very strong and significant correlation between age and cooperative behavior (see Table V): older contestants (≥ 37 years) are much more likely to cooperate than younger contestants (< 37 years), regardless of the age of the opponent. This result is in line with List (2006) who finds that contestants ≥ 31 years are significantly more likely to cooperate than younger contestants. Concerning gender, we find no direct effect, which is contrary to the studies of e.g., Kahn, Hottes, and Davis (1971) and Ortmann and Tichy (1999). But we can show that younger males are significantly more likely to defect; and, as age increases, males are more likely to cooperate than females (see Table A.4 in the appendix).

cooperation, see Table A.5 in the appendix. Attractive females (model (1)) and attractive younger contestants (model (3)) show more cooperative behavior, whereas attractive males (model (1)) and attractive older contestants (model (3)) cooperate less.

Opponent's attractiveness We now turn to the impact of the opponent's facial attractiveness on a contestant's willingness to cooperate. The regression results in Table V show that contestants are $10-16$ percentage points more likely to cooperate when facing an attractive opponent than when facing an unattractive opponent. Hence, attractive contestants are rewarded with greater cooperativeness, and this provides attractive contestants a beauty premium. The premium is independent of the opponent's gender and age, see Table A.5, model (2) and (4) in the appendix. Furthermore, our results show that least attractive contestants suffer a beauty penalty due to lower cooperativeness towards them. As model (5) reports, a contestant is less likely to cooperate if the opponent is rated to be least attractive than if the opponent is neither rated to be most nor least attractive. There is no significant effect on cooperative behavior if the opponent is rated to be most attractive. The result holds independently of the definition of the extreme measure of attractiveness. These findings suggest that contestants rather focus on the opponent's "negative" than "positive" appearance in the decision to cooperate.
The results remain unchanged when adding the two binary statement variables (see model (4) of Table V). There is additional evidence that a contestants is less likely to cooperate if the opponent's appearance is rated as to "help her in life" than if it is not. But we find no interaction effect of the binary statement variables and our attractiveness measures, which indicates that the statement variables have not much additional explanatory power.[15]

Similarity Since contestants behave more cooperatively towards the attractive counterpart, the question arises whether pairs of attractive contestants behave differently in the prisoner's dilemma than pairs of unattractive contestants or pairs who are mixed in terms of facial attractiveness. We denote pairs of contestants as *similar team* if either both contestants are facially attractive or both are facially unattractive, and we identify pairs of contestants as an *attractive team* if both are facially attractive, and as an *unattractive team* if both are facially unattractive. We find no evidence that contestants who are similar with respect to attractiveness are more or less likely to cooperate or defect, see model (1) of Table A.6 in the appendix. However, we find that an unattractive contestant is less likely to cooperate with her unattractive counterpart compared to teams of contestants

[15]Note, the effect of the variable "the contestant's appearance helps her in life" without the inclusion of the facially attractiveness variables is positive, but not significant. We also tested for interaction effects between the two statement variables and gender or age (table unreported). We find that a female opponent, whose appearance is rated as helping her in life, lowers the contestant's propensity to cooperate; but this effect is not robust with respect to different specifications of the model.

Table V: Results from binary probit regressions of the decision to "cooperate" ($y_i = 1$) or "defect" ($y_i = 0$) in the prisoner's dilemma

	Marginal Effects									
	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)	
Attractiveness (mean-centered, d) (d)	0.067	(0.049)	0.067	(0.052)	0.048	(0.053)	0.026	(0.060)		
Opponent Attractiveness (mean-centered, d) (d)	0.112**	(0.046)	0.125**	(0.049)	0.101**	(0.050)	0.162***	(0.057)		
Appearance Helps In Life (d)							0.062	(0.066)	0.075	(0.061)
Opponent Appearance Helps In Life (d)							-0.110*	(0.064)	-0.050	(0.060)
Appearance Strikes As Calculating (d)							0.017	(0.056)	0.012	(0.055)
Opponent Appearance Strikes As Calculating (d)							-0.001	(0.056)	0.000	(0.056)
Most attractive (90% percentile) (d)									-0.134	(0.086)
Opp. most attractive (90% percentile) (d)									-0.096	(0.085)
Least attractive (10% percentile) (d)									-0.077	(0.090)
Opp. least attractive (10% percentile) (d)									-0.230***	(0.081)
Male (d)	-0.036	(0.048)	-0.046	(0.050)	-0.058	(0.051)	-0.052	(0.056)	-0.060	(0.051)
Age (cont.)	0.081***	(0.023)	0.082***	(0.025)	0.088***	(0.027)	0.093***	(0.030)	0.085***	(0.028)
Opp. Male (d)							-0.008	(0.056)		
Opp. Age (cont.)							0.016	(0.028)		
Promise or Vow (d)					0.240***	(0.051)	0.254***	(0.052)	0.258***	(0.052)
Opp. Promise or Vow (d)					0.006	(0.054)	-0.002	(0.054)	-0.013	(0.054)
Handshake (d)					-0.102*	(0.059)	-0.098	(0.060)	-0.117**	(0.059)
Demographics	–		X		X		X		X	
Stake Size	–		X		X		X		X	
Pre-Play	–		–		X		X		X	
Wald χ^2	18.59***		43.01***		63.96***		69.21***		68.21***	
Log-Likelihood	-282.75		-264.39		-250.74		-248.53		-246.79	
Adjusted R^2	0.013		0.039		0.055		0.042		0.048	
N	422		419		419		419		419	
Number of clusters	211		211		211		211		211	

Standard errors in parentheses are corrected for episode clusters; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
(d) for discrete change of dummy variable from 0 to 1

who are both attractive or mixed, see model (2) and (3) of Table A.6 in the appendix. This improves our attractiveness-results: contestants not only behave more cooperatively towards an attractive partner, but also more deceitfully towards an unattractive partner, and this unattractive-penalty is likely to dominate.[16]

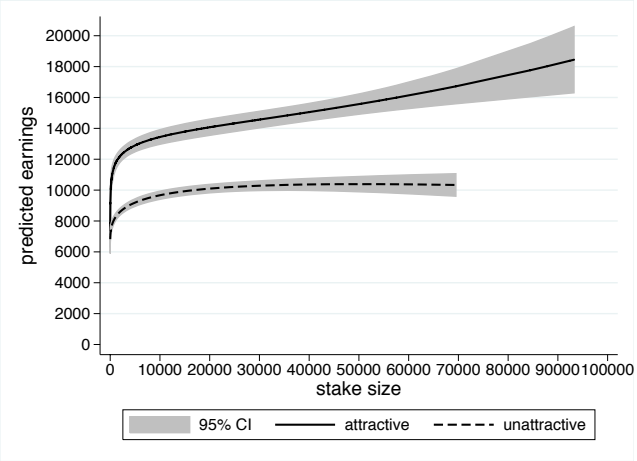
Thus, our results provide evidence for a causal relationship between the opponent's attractiveness and cooperative behavior. Facially attractive contestants are able to provoke cooperation from their counterpart, independent of their gender or age. But we do not find a significant difference in behavior between facially attractive and facially unattractive contestants.

4.2 Beauty premium

The results of the previous section should also translate into a monetary beauty premium, i.e., into higher earnings for the attractive than for the unattractive. In order to quantify the marginal beauty premium we use a standard censored tobit model. The outcome "taking no money home" from the prisoner's dilemma is interpreted as a corner solution outcome, where the response variable y_i describes the observable outcome of a contestant, which takes on the value zero with positive probability (if the opponent defects), and which is a continuous variable over strictly positive values (if the opponent cooperates).
We estimate marginal effects for contestant i's expectation of receiving positive earnings $y_i > 0$, conditional on i's own and her opponent's facial attractiveness (standardized measures), demographic characteristics, the log value of the stake size, as well as opponent characteristics, which are summarized in $\mathbf{X_i}$, $E(y_i|\mathbf{X_i}, y_i > 0)$.
We find that a one standard deviation increase in facial attractiveness significantly increases a contestant's expected positive earnings by £617 to £741 depending on the controls included, see Table A.7 in the appendix. Figure 3 provides an illustration of the beauty premium based on model (2) of Table A.7 in the appendix, and depicts the predicted positive earnings for attractive and unattractive contestants as a function of stake size. The figure shows that the predicted earnings of facially attractive contestants are always clearly above the earnings of unattractive contestants, given the size of the jackpot.[17]
Decomposing the beauty premium with respect to gender and age, we can show that there are no differences in receiving positive earnings between attractive male and female,

[16]For robustness of our results, we repeat the analysis using ordered probits on the contestants' likelihood to mutually cooperate. Again, similar teams, as well as teams of attractive contestants are no more likely to reach a certain outcome. Only pairs of unattractive contestants are significantly less likely to reach mutual cooperation (tables unreported). Additionally, we estimate regressions including the relative difference between both final contestants' attractiveness, i.e., the distance in attractiveness between both contestants, and including an index of the contestants' similarity with respect to facial attractiveness, age, and gender, weighting each component by one-third. All measures do not matter for the contestant's

Figure 3: Predicted positive earnings by attractive and unattractive contestants, given stake size



as well as attractive older and younger contestants. However, a contestant's expected earnings increase the older the opponent, and a contestants expected earnings decrease when the opponent is male (see model (2) of Table A.7 in the appendix).

5 Transmission channels

Our results show that attractive people are able to provoke cooperative behavior from their opponent and, since they are not more or less cooperative than unattractive people, they obtain a beauty premium in the prisoner's dilemma. In this section we will address the potential transmission channels suggested by the literature (see Section 2) and evaluate their explanatory content for our observed effects of attractiveness.

5.1 Beauty-is-good stereotyping and taste-based discrimination

The most prominent theory is that people hold stereotype beliefs about attractive people. People believe that attractive ones behave more pro-socially than less attractive ones. In the presence of interdependent social preferences, people may belief that attractive ones are more cooperative and behave more cooperatively towards them with the intention to reciprocate cooperation. Hence, these stereotype beliefs can mediate people's behavior. Recall, however, that we do not find any difference between attractive and unattractive contestants regarding their likelihood to cooperate. If "beauty-is-good" stereotyping is

decision to cooperate.

[17]We obtain a similar result for the overall conditional expectation that a contestant receives the outcome y_i, given $\mathbf{X_i}$, $E(y_i|\mathbf{X_i})$. A one standard deviation increase in facial attractiveness significantly increases a contestant's expected earnings by £1 784 to £2 153. Further, we estimate that the probability of obtaining positive earnings, given X_i, $Pr(y_i > 0)|\mathbf{X_i})$, increases by up to 5.9 percentage points (tables unreported).

driving our result, then we should observe a higher likelihood of cooperation towards attractive contestants in the prisoner's dilemma and the effect should be independent of the selected sample, e.g., the effect should be present in interactions between contestants of the same sex as well as between contestants of the opposite sex.
However, theories from evolutionary psychology argue that effects of physical attractiveness on behavior originate in primeval partner selection and therefore predict that effects of attractiveness are more prevalent in mixed-gender interactions (see e.g., Cosmides and Tooby (1987)). To scrutinize this argument, we parse the data in mixed-gender and same-gender interactions (63% and 37% of all interactions) and run several probits to evaluate the influence of facial attractiveness on a contestant's propensity to cooperate, controlling for demographic characteristics and stake size. The two subsamples are not different regarding the observed cooperation rate (54% in mixed- vs. 53% in same-gender interactions, see Table VI).

Table VI: Results from binary probit regressions on outcomes in the prisoner's dilemma (subsamples)

	Marginal Effects			
	(1) Mixed-gender		(2) Same-gender	
Attractiveness (mean-centered, d)	0.047	(0.067)	0.101	(0.094)
Opp. Attractiveness (mean-centered, d)	0.151**	(0.062)	0.013	(0.092)
Male (d)	-0.049	(0.059)	-0.104	(0.107)
Age (cont.)	0.076**	(0.033)	0.111***	(0.042)
Demographics	X		X	
Stake Size	X		X	
Wald χ^2	47.27***		20.01*	
Log-Likelihood	-161.07		-96.46	
Adjusted R^2				
N	265.00		154.00	
Number of clusters	133.00		78.00	

Note: Binary probit regressions of the decision to "cooperate" ($y_i = 1$) or "defect" ($y_i = 0$) in the prisoner's dilemma, restricting the sample to (1) mixed-gender and (2) same-gender interactions. The marginal effect of the respective explanatory variable determines the effective change of this variable on player i's predicted probability to "split". Standard errors are reported in parentheses and are corrected for episode clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As the regression results of Table VI show, attractiveness only matters in mixed-gender interactions, but not in same-gender ones. Males and females are about 15% percentage points more likely to cooperate if facing a facially attractive opponent of the other sex than if they face an unattractive opponent.[18] The finding applies across sexes, both females and males are biased towards attractiveness if facing someone of the other sex. This

[18]This result is also displayed in the fact that facially attractive contestants enjoy a beauty premium in mixed-gender interactions, but not in same-gender ones (tobit regression tables unreported). A one standard deviation increase in attractiveness results in a significantly higher gains for the attractive in mixed-gender (p=0.003), but not in same-gender interactions (p=0.840).

extends the finding of Ashmore and Longo (1995), who note that only attractive females tend to have the ability to make males more likely to do them a favor. We show that males have the same ability.
According to the theory of taste-based discrimination (Becker (1957)), contestants might favor to be in the final round with a facially attractive fellow player. Hence, taste-based discrimination could be the underlying mechanism if we identify a beauty premium in the pre-play, but not in all interactions in the prisoner's dilemma. Using binary and ordered probit models, we estimate the effect of attractiveness on the likelihood to be voted off the game and on the likelihood to receive a certain number of votes after the first and second round. We find no significant effect of attractiveness on the voting outcome (tables unreported).[19] This is also reflected in the almost equal share of attractive contestants in the final and the one in the initial round of the game show (52.4% vs. 51.8%).

The absence of a beauty premium or plainness penalty in the pre-play and in same-gender interactions in the final suggests that neither beauty-is-good stereotyping nor taste-based discrimination can explain our results consistently. It rather seems that the ability of attractiveness to mediate behavior is entangled to the opponent's sex and that people have a preference to cooperate with someone to whom they are personally attracted.

5.2 Vocal attractiveness and social skills

In addition to the theories discussed above, the literature reports that physical and vocal attractiveness are highly correlated and thus suggests the existence of a vocal rather than a visual attractiveness stereotype (see, e.g., Zuckerman and Driver (1989)). Furthermore, physically attractive people are also assigned stronger verbal and social skills (e.g., Goldman and Lewis (1977) and Erwin and Calev (1984)). In the presence of other regarding preferences, vocal attractiveness and potentially strong verbal and social skills might enable attractive people to trigger cooperative behavior off their opponents. Although our data does not allow us to directly test for the impact of (perceived) vocal attractiveness on cooperative behavior, we can indirectly test for a correlation between facial attractiveness and communication which comprises verbal and social skills.
Shortly before the prisoner's dilemma is played both contestants are given some extra time to talk to each other. In these short conversations which are, on average, 38 seconds long, each contestant tries to convince her opponent to cooperate. Since it has been shown that

[19] In all regressions we control for demographic characteristics such as gender, age, race and place of residence as well as for objective voting criteria such as the stake size a contestant accumulated, and whether a contestant lied in a previous round of the pre-play. We also follow Belot, Bhaskar, and van de Ven's approach and rank the contestants by their facial attractiveness to explain the likelihood to be voted off during the game, which yields no significant effects either (table unreported). Our finding is in contrast to Belot, Bhaskar, and van de Ven (2012), who find that attractive people are positively discriminated against unattractive people and hence pass the pre-play more easily.

promises effect people's behavior in experiments (e.g., Charness and Dufwenberg (2006) and Vanberg (2008)) and in the field (e.g., Belot, Bhaskar, and van de Ven (2010)), we code whether a contestant explicitly promises her opponent to cooperate.[20] Furthermore, we observe that contestants use handshakes to corroborate their mutual intention to cooperate and therefore we also code whether two contestants shake hands. As Table V shows, promises have a significantly positive impact on cooperative behavior, whereas handshakes have a negative one.[21] Comparing models (1) and (3) of Table V shows that the effect of attractiveness on cooperation remains almost unchanged when we add the variables of communication as controls.

If attractive contestants are better in terms of verbal and social skills, a promise or handshake of a facially attractive contestant might be more convincing than a promise or handshake of a facially unattractive contestant. Testing for interaction effects between facial attractiveness and the communication variables, reveals no additional effects (table unreported). Further, we could expect that facially attractive contestants are more likely to elicit a promise from their opponent and less likely to engage in a handshake. Using binary probit regressions on the contestant's propensity to promise or to shake hands, we find that facially attractive contestants are not more likely to state a promise or to shake hands than facially unattractive contestants (table unreported). However, we find limited evidence that contestants are significantly more likely to state a promise if the opponent is attractive (table unreported). Since promising to cooperate significantly increases a contestant's likelihood to cooperate, we cannot exclude that promises are the underlying transmission channel of facial attractiveness on cooperative behavior. The absence of the beauty premium in the pre-play can also not be explained by the theory of better social and verbal skills or vocal attractiveness.

6 Conclusion

This paper analyzes the relationship between attractiveness and cooperative behavior in a high-stakes field setting with two-sided communication. Two independent data sets are combined. One on cooperation, collected from decisions made in a slightly modified prisoner's dilemma played in the final round of a television game show. The other one on the physical attractiveness of the game show's contestants using a sample of indepen-

[20] We count all statements as a promise when they contain either the word "promise" or "swear" or they are a statement of intent. Examples are "I promise to split", "I promise I will not steal", "I swear I will split", "I swear I will not steal", "I will split", or "I will not steal".

[21] The effect of the communication variables is robust to various specifications of the regression model. Including only demographics, the likelihood to cooperate of a contestant who made a promise to cooperate is 25 percentage points (p=0.000) higher than the one of a contestant who did not promise to cooperate. When contestants use a handshake they are actually 13 percentage points (p=0.031) less likely to cooperate than when they do not use a handshake (table unreported).

dent third-party raters. In the prisoner's dilemma two finalists play for an accumulated jackpot by deciding either to "split" the jackpot or to "steal" it and to keep the entire amount for themselves. The results show a strong and robust effect of attractiveness: contestants are significantly more likely to cooperate with a facially attractive opponent. But, facially attractive contestants are not more or less likely to be cooperative compared to facially unattractive contestants. Hence, attractive contestants are rewarded by a beauty premium, which, at the margin, amounts to up to £2 153 for an increase in attractiveness of one standard deviation. The attractiveness effect is robust to demographic characteristics, including gender, stake size, and communication. However, the effects of attractiveness might be amplified by the attractive contestant's ability to talk their opponent into promising to cooperate, which has a significantly positive effect on a contestant's likelihood to cooperate.

However, the decomposition of the data into same- and mixed-gender interactions reveals a new picture. The ability of attractive contestants to elicit cooperative behavior from their opponent vanishes in interactions between two contestants of the same sex. That is, contestants are only biased by the facial attractiveness of their opponent when the opponent is of the other sex. This suggests that stereotype beliefs about attractive people, such as them being more pro-social, cannot explain our results. Since for this explanation to hold, the effect should prevail in all interactions. In line with our finding are rather theories from evolutionary psychology arguing that effects of physical attractiveness originate in primeval partner selection and should therefore be only or at least more present in mixed-gender interactions (see e.g., Cosmides and Tooby (1987)). Also, the absence of a beauty premium or plainness penalty in the pre-play suggests that physical appearance is particularly important as soon as people are lacking objective information and moreover it allows us to exclude taste-based discrimination as the underlying transmission channel of attractiveness on cooperative behavior.

To confirm our hypothesis of a preference based mechanism, we test it with the data from Mobius and Rosenblat's experiments on the beauty premium in the labor market. The authors argue that stereotype beliefs about attractive people's productivity and the higher confidence of attractive people in their own productivity are driving the results. Attractive people earn significantly higher wages even though they do not display a higher productivity than unattractive people. We re-estimate Mobius and Rosenblat's full model (Table 6, p.234 in Mobius and Rosenblat (2006)) parsing the data into mixed- and same-gender interactions. The results are presented in Table A.8 in the Appendix. As predicted by our hypothesis the effects of attractiveness remain only prevalent in interactions between either a female worker and a male employer or a male worker and a female employer (Model (2)). In model (3) which uses the sample of same-gender interactions, all effects of attractiveness disappear.

Our results are relevant and applicable to many, if not all, one-shot face-to-face inter-

actions. The better performance of attractive people in the labor market, for instance, may be enforced by the fact that they also benefit from greater cooperativeness towards them. And facing someone attractive of the other sex seems to bias people when making economic decisions, which may help beautiful people in negotiations when cooperation is required.

References

Alreck, P. L., and R. B. Settle (1995): *The Survey Research Handbook*. Chicago: Irwin.

Andreoni, J., and R. Petrie (2008): "Beauty, Gender and Stereotypes: Evidence from Laboratory Experiments," *Journal of Economic Psychology*, 29, 73–93.

Antonakis, J., and O. Dalgas (2009): "Predicting Elections: Child's Play!," *Science*, 323, 1183.

Ashmore, R. D., and L. C. Longo (1995): "Accuracy of Stereotypes: What Research on Physical Attractiveness can Teach us," in *Stereotype Accuracy: Toward Appreciating Group Differences*, ed. by Y. Lee, L. Jussim, and C. McCauley, chap. 3, pp. 63–86. American Psycological Association.

Becker, G. (1957): *The Economics of Discrimination*. University of Chicago Press.

Belot, M., V. Bhaskar, and J. van de Ven (2010): "Promises and Cooperation: Evidence from a TV Game Show," *Journal of Economic Behavior and Organizations*, 73, 396–405.

——— (2012): "Beauty and the Sources of Discrimination," *Journal of Human Resources*, 47(3), 851–872.

Berggren, N., H. Jordahl, and P. Poutvaara (2010): "The Looks Of A Winner: Beauty And Electoral Success," *Journal of Public Economics*, 94, 8–15.

Biddle, J. E., and D. S. Hamermesh (1998): "Beauty, Productivity, and Discrimination: Lawyers' Looks and Lucre," *Journal of Labor Economics*, 16(1), 172–201.

Charness, G., and M. Dufwenberg (2006): "Promises and Partnership," *Econometrica*, 74(6), 1579–1601.

Cosmides, L., and J. Tooby (1987): "From Evolution to Behavior: Evolutionary Psychology as the Missing Link," in *The Lates on the Best: Essays on Evolution and Optimality*, ed. by J. Dupré, chap. 13, pp. 277–306. Cambridge: MIT Press.

Dion, K., E. Berscheid, and E. Walster (1972): "What Is Beautiful Is Good," *Journal of Personality and Social Psychology*, 24, 285–290.

Eagly, A. H., R. D. Ashmore, M. G. Makhijani, and L. C. Longo (1991): "What is Beautiful is Good, But...: A Meta-Analytic Review of Research on the Physical Attractiveness Stereotype," *Psychological Bulletin*, 110(1), 109–128.

Eckel, C. C., and R. Petrie (2011): "Face Value," *American Economic Review*, 101, 1497–1513.

Erwin, P. G., and A. Calev (1984): "Beauty: More than Skin Deep?," *Journal of Social and Personal Relationships*, 1, 359–361.

Feingold, A. (1992): "Good-Looking People Are Not What We Think," *Psychological Bulletin*, 111(2), 304–341.

Goldman, W., and P. Lewis (1977): "Beautiful is Good: Evidence that the Physically Attractive are More Socially Skillful," *Journal of Experimental Social Psychology*, 13, 125–130.

Grammer, K., B. Fink, A. P. Møller, and R. Thornhill (2003): "Darwinian Aesthetics: Sexual Selection and the Biology of Beauty," *Biological Review*, 78, 385–407.

Jackson, L. A., J. H. Hunter, and C. N. Hodge (1995): "Physical Attractiveness and Intellectual Competence: A Meta-Analysis Review," *Social Psychology Quarterly*, 58(2), 108–122.

Kahn, A., J. Hottes, and W. L. Davis (1971): "Cooperation and Optimal Responding in the Prisoner's Dilemma Game: Effects of Sex and Physical Attractiveness," *Journal of Personality and Social Psychology*, 17(3), 267–279.

Langlois, J. H., L. Klakanis, A. J. Rubenstein, A. Larson, M. Hallam, and M. Smoot (2000): "Maxims or Myths of Beauty? A Meta-Analysis and Theoretical Review," *Pssychological Bulletin*, 126(3), 390–423.

List, J. A. (2006): "Friend or Foe? A Natural Experiment of the Prisoner's Dilemma," *The Review of Economics and Statistics*, 88(3), 463–471.

Mobius, M. M., and T. S. Rosenblat (2006): "Why Beauty Matters," *The American Economic Review*, 96(1), 222–235.

Mueller, U., and A. Mazur (1997): "Facial Diominance in Homo Sapiens as Hinest Signaling of Male Quality," *Behavioral Ecology*, 8, 569–579.

Mulford, M., J. Orbell, C. Shatto, and J. Stockard (1998): "Physical Attractiveness, Opportunity and Success in Everyday Exchange," *The American Journal of Sociology*, 103(6), 1565–1593.

Ortmann, A., and L. K. Tichy (1999): "Gender differences in the laboratory: evidence from prisoner's dilemma games," *Journal of Economic Behavior and Organization*, 39, 327–339.

Ottati, V. C., and M. Deiger (2002): "Visual Cues and the Candidate Evaluation Process," in *Social Psychology of Politics*, ed. by V. C. Ottati, and R. S. Tindale, pp. 75–87. New York: Kluwer Academic/Plenum.

Rosar, U., M. Klein, and T. Beckers (2008): "The Frog Pond Beauty Contest: Physical Attractiveness and Electoral Success of the Constituency Candidates at the North Rhine-Westphalia State Election of 2005," *European Journal of Political Research*, 47, 64–79.

Rosenblat, T. S. (2008): "The Beauty Premium: Physical Attractiveness and Gender in Dictator Games," *Negotiation Journal*, 24(4), 465–481.

Ruffle, B. J., and Z. Shtudiner (2010): "Are Good-Looking People More Employable?," *Working Paper*.

Solnick, S. J., and M. E. Schweitzer (1999): "The Influence of Physical Attractiveness and Gender on Ultimatum Game Decisions," *Organizational Behavior and Human Decision Process*, 79(3), 199–215.

Vanberg, C. (2008): "Why Do People Keep Their Promises? An Experimental Test of Two Explanations," *Econometrica*, 76(6), 1467–1480.

Wilson, R. K., and C. C. Eckel (2006): "Judging a Book by its Cover: Beauty and Expectations in the Trust Game," *Political Research Quarterly*, 59(189), 188–202.

Zuckerman, M., and R. E. Driver (1989): "What sounds beautiful is good: The vocal attractiveness stereotype," *Journal of Nonverbal Behavior*, 13, 67–82.

Zuckerman, M., H. Hodgins, and K. Miyake (1990): "The vocal attractiveness stereotype: Replication and elaboration," *Journal of Nonverbal Behavior*, 14, 97–112.

Appendix

A Figures and tables

Table A.1: Summary statistics prisoner's dilemma

Variable	Mean	Std. Dev.	Min.	Max.	N
Decision Variables					
Cooperate	0.54	0.5	0	1	422
Mutual Decision (0="steal-steal", 1="steal-split", 2="split-split")	1.07	0.76	0	2	422
Amount Money Taken Home	4 614.8	10 799.03	0	93 250	422
Demographics					
Male	0.46	0.5	0	1	422
Age[a](cont.)	3.21	1.1	1	6.5	422
Age of Male (cont.)	3.26	1.07	1	6.4	422
Age of Female (cont.)	3.19	1.09	1.25	6.6	422
White	0.94	0.25	0	1	422
London	0.11	0.32	0	1	422
England (1 = ENG, 0 = SCO, WAL, NIR, IRL)	0.85	0.36	0	1	420
Social Job (Reputation)[b]	0.16	0.36	0	1	421
Unexperienced (contestants of series 1)	0.19	0.39	0	1	422
Experienced (contestants of series 4)	0.19	0.39	0	1	422
Stake size					
Jackpot	12912.33	18213.95	3	93250	422
Potential Jackpot[c]	50329.69	29946.46	5000	168100	422
Communication					
Promise or Vow	0.42	0.494	0	1	422
Handshake	0.33	0.47	0	1	422
Pre-play					
Accumulated Most Money	0.5	0.5	0	1	422
Selected Higher Values in Bin/Win	0.5	0.5	0	1	422
Selected Most Killers in Bin/Win	0.33	0.47	0	1	422
"Should Have Left The Game"[d]	0.26	0.44	0	1	422
Lied during Pre-Play	0.62	0.49	0	1	422

[a] Age is judged on a 7-item scale (see questionnaire and section 3.2), where 3="30-40", and 4="40-50" implying that the scale average of 3.21 equals a mean age of 37 years.
[b] A social job is defined as a job in which people care for other people, e.g., doctors, nurses, child minders, social workers, teachers, police officers, firemen, soldiers.
[c] The potential jackpot is announced by the show host before the actual jackpot is determined. It displays the maximal amount the actual jackpot could potentially comprise of.
[d] The variable "should have left the game" points at the player who is the "weakest" in material terms in round 2.

Figure A.1: Raw [left] and demeaned [right] variables of facial appearance, N=422

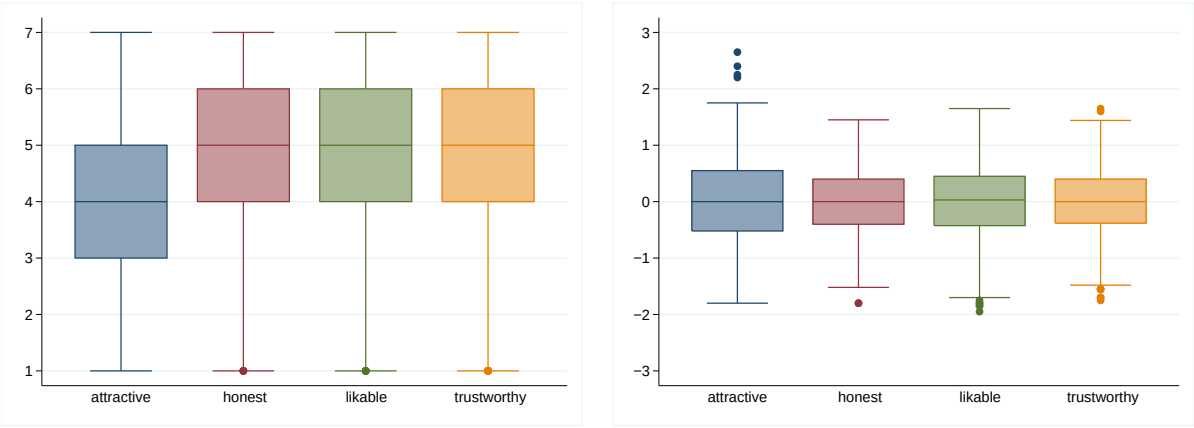


Table A.2: Summary statistics all raters

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Age (in years)	41.76	18.46	17	93	720
Age of Male (in years)	39.57	17.70	17	93	371
Age of Female (in years)	44.10	19.00	18	86	349
Male	0.51	0.5	0	1	728
Female ($\geq$ 44.1 years)	62.07	10.10	45	86	162
Female ($<$ 44.1 years)	28.52	7.84	18	44	187
Male ($\geq$ 39.6 years)	58.02	11.88	40	93	153
Male ($<$ 39.6 years)	26.62	5.17	17	39	218

Table A.3: Distribution of attractive and unattractive finalists

	facially attractive	facially unattractive
Male ($<$ 37 years)	21%	27%
Male ($\geq$ 37 years)	15%	30%
Female ($<$ 37 years)	41%	18%
Female ($\geq$ 37 years)	23%	25%
	100% (N=221)	100% (N=201)

Table A.4: Results from binary probit regressions of the decision to "cooperate" ($y_i = 1$) or "defect" ($y_i = 0$) in the prisoner's dilemma, including the interaction term "gender*age"

	Marginal Effects									
	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)	
Attractiveness (mean-centered, d)			0.062	(0.052)	0.048	(0.058)	0.025	(0.059)		
Opp. Attractiveness (mean-centered, d)			0.124**	(0.049)	0.168***	(0.056)	0.161***	(0.057)		
Appearance Helps In Life (d)					0.036	(0.062)	0.056	(0.065)	0.067	(0.061)
Opp. Appearance Helps In Life (d)					-0.090	(0.061)	-0.108*	(0.064)	-0.049	(0.060)
Appearance Strikes As Calculating (d)					0.011	(0.054)	0.018	(0.057)	0.056	(0.056)
Opp. Appearance Strikes As Calculating (d)					0.018	(0.055)	0.001	(0.056)	0.002	(0.056)
Most attractive (90% percentile, d)									-0.141	(0.087)
Opp. most attractive (90% percentile, d)									-0.090	(0.087)
Least attractive (10% percentile, d)									-0.088	(0.089)
Opp. least attractive (10% percentile, d)									-0.235*	(0.081)
Male*Age (cont.)	0.071**	(0.042)	0.067*	(0.042)	0.065*	(0.042)	0.047	(0.043)	0.059*	(0.043)
Male (d)	-0.304**	(0.147)	-0.286*	(0.149)	-0.273*	(0.149)	-0.232	(0.165)	-0.288*	(0.160)
Age (cont.)	0.037	(0.033)	0.048	(0.034)	0.050	(0.034)	0.066*	(0.039)	0.073	(0.050)
Opp. Male (d)							-0.004	(0.056)		
Opp. Age (cont.)							0.015	(0.028)		
Promise or Vow (d)							0.252***	(0.051)	0.255***	(0.052)
Opp. Promise or Vow (d)							-0.008	(0.055)	-0.020	(0.055)
Handshake (d)							-0.092	(0.060)	-0.109*	(0.059)
Demographics	X		X		X		X		X	
Stake Size	X		X		X		X		X	
Pre-Play	–		–		–		X		X	
Wald χ^2	40.35***		47.21***		48.56***		69.17***		70.26***	
Log-Likelihood	-266.43		-263.10		-261.74		-247.87		-245.75	
Adjusted R^2	0.035		0.040		0.030		0.040		0.048	
N	419		419		419		419		419	
Number of clusters	211		211		211		211		211	

Standard errors in parentheses are corrected for episode clusters; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
(d) for discrete change of dummy variable from 0 to 1

Table A.5: Results from binary probit regressions of the decision to "cooperate" ($y_i = 1$) or "defect" ($y_i = 0$) in the prisoner's dilemma, including interaction terms w.r.t. gender and age

	Marginal Effects							
	Model (1)		Model (2)		Model (3)		Model (4)	
Attractiveness*Male (d)	-0.153*	(0.104)						
Opp. Attractiveness*Male (d)			0.099	(0.103)				
Attractiveness*Age (cont.)					-0.079**	(0.044)		
Opp. Attractiveness*Age (cont.)							-0.011	(0.044)
Attractiveness (mean-centered, d)	0.133*	(0.078)			0.334**	(0.150)		
Opponent Attractiveness (mean-centered, d)			0.068	(0.073)			0.154	(0.161)
Male (d)	0.044	(0.072)			-0.036	(0.050)	-0.054	(0.050)
Opp. Male (d)			-0.055	(0.075)				
Age (cont.)	0.082***	(0.027)	0.075***	(0.024)	0.123***	(0.033)		
Opp. Age (cont.)							0.006	(0.033)
Demographics	X		X		X		X	
Stake Size	X		X		X		X	
Wald χ^2	36.01***		44.63***		43.93***		35.32***	
Log-Likelihood	-266.08		-265.29		-265.65		-270.02	
Adjusted R^2	0.033		0.035		0.034		0.019	
N	419		419		419		419	
Number of clusters	211		211		211		211	

Standard errors in parentheses are corrected for episode clusters; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; (d) for discrete change of dummy variable from 0 to 1

Table A.6: Results from binary probit regressions of the decision to "cooperate" ($y_i = 1$) or "defect" ($y_i = 0$) in the prisoner's dilemma when teams of contestants are considered

	Marginal Effects					
	Model (1)		Model (2)		Model (3)	
Similar Team (d)	-0.043	(0.054)				
Attractive Team (d)			0.045	(0.067)	0.027	(0.070)
Unattractive Team (d)			-0.154**	(0.064)	-0.113*	(0.067)
Male (d)	-0.056	(0.050)	-0.041	(0.050)	-0.064	(0.056)
Opp. Male (d)					-0.014	(0.056)
Age (d)	0.166***	(0.050)	0.187***	(0.051)	0.184***	(0.057)
Opp. Age (d)					-0.034	(0.056)
Demographics	X		X		X	
Stake Size	X		X		X	
Pre-Play & Communication	–		–		X	
Wald χ^2	40.40***		46.68***		68.32***	
Log-Likelihood	-266.77		-263.49		-250.48	
Adjusted R^2	0.034		0.042		0.049	
N	419		419		419	
Number of clusters	211		211		211	

Note: The "team variables" are indicators and equal 1 if the team is so composed and 0 otherwise, e.g., "Similar Team" equals 1 if both contestants are either attractive or unattractive, and 0 otherwise. Standard errors are reported in parentheses and are corrected for episode clusters. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

Table A.7: Censored tobit regressions results for $E[y_i|\mathbf{X_i}, y_i > 0]$

	Marginal Effects			
	Model (1)		Model (2)	
Attractiveness (standardized)	616.927*	(334.835)	741.419**	(325.057)
Opp. Attractiveness (standardized)			457.207	(285.004)
Male (d)	-209.774	(525.072)	-328.794	(540.835)
Age (cont.)	-218.725	(283.775)	-8.920	(268.037)
Opp. Male (d)			-1126.281*	(573.093)
Opp. Age (cont.)			1087.797***	(363.794)
log(Jackpot)	689.636***	(157.020)	615.929***	(146.146)
Demographics	X		X	
Opp. Demographics	–		X	
F-Statistic	3.186***		2.843***	
Log-Likelihood	-2593.35		-2578.46	
Adjusted R^2	0.003		0.006	
$\hat{\sigma}$	15059.81		14483.33	
N	419		419	
Number of clusters	211		211	

Note: Censored tobit regression for the conditional expectation, that player i receives a positive earnings $y_i > 0$ from the prisoner's dilemma game: $E[y_i|\mathbf{X_i}, y_i > 0] = \mathbf{X_i}\beta + \sigma\left[\frac{\phi(\mathbf{X_i}\beta/\sigma)}{\Phi(\mathbf{X_i}\beta/\sigma)}\right]$, where the inverse Mills ratio is evaluated at $\frac{\mathbf{X_i}\beta}{\sigma}$. The marginal effect for the jth independent (continuous) variable on $E[y_i|\mathbf{X_i}, y_i > 0]$ is computed as $ME_j = \frac{\partial E[y_i|\mathbf{X_i}, y_i>0]}{\partial x_j} = \beta_j\left[1 - \frac{\phi(\mathbf{X_i}\beta/\sigma)}{\Phi(\mathbf{X_i}\beta/\sigma)}\left(\frac{\mathbf{X_i}\beta}{\sigma} + \frac{\phi(\mathbf{X_i}\beta/\sigma)}{\Phi(\mathbf{X_i}\beta/\sigma)}\right)\right]$, and quantifies the expected increase in earnings, conditional on being positive. (d) for discrete change of dummy variable from 0 to 1. Standard errors are reported in parentheses and are corrected for episode clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Fixed effects regression of Mobius and Rosenblat (2006)

	Model (1) All interactions (Table 6, p. 234)		Model (2) Mixed gender interactions		Model (3) Same gender interactions	
LNPROJECTED	0.409***	(0.043)	0.423***	(0.066)	0.337***	(0.066)
LNPROJECTED*VISUAL	0.007	(0.059)	0.022	(0.088)	0.047	(0.089)
LNPROJECTED*AUDIO	-0.129**	(0.059)	-0.101	(0.122)	-0.091	(0.081)
LNPROJECTED*VISUAL*AUDIO	0.056	(0.084)	-0.092	(0.155)	0.039	(0.117)
LNPROJECTED*FTF	-0.069	(0.060)	0.003	(0.104)	0.038	(0.079)
LNACTUAL	-0.004	(0.027)	-0.094*	(0.048)	0.032	(0.045)
BEAUTY	-0.010	(0.031)	-0.076*	(0.046)	0.073	(0.055)
BEAUTY*VISUAL	0.094**	(0.043)	0.170**	(0.066)	-0.007	(0.080)
BEAUTY*AUDIO	0.103***	(0.035)	0.162**	(0.067)	0.030	(0.059)
BEAUTY*VISUAL*AUDIO	-0.097*	(0.050)	-0.161*	(0.093)	0.076	(0.085)
BEAUTY*FTF	0.052	(0.035)	0.105	(0.070)	0.019	(0.053)
LNESTIMATED	0.018	(0.065)	0.157	(0.096)	0.048	(0.116)
LNESTIMATED*VISUAL	0.034	(0.083)	-0.126	(0.118)	0.131	(0.149)
LNESTIMATED*AUDIO	0.265***	(0.083)	0.084	(0.142)	0.205	(0.138)
LNESTIMATED*VISUAL*AUDIO	-0.056	(0.117)	0.110	(0.200)	-0.102	(0.194)
LNESTIMATED*FTF	-0.116	(0.083)	0.065	(0.141)	-0.253**	(0.126)
F-Statistic	21.939***		8.018***		11.773***	
R^2	0.627		0.664		0.709	
N	812		384		423	

Note: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is LNWAGE, standard errors are in parentheses. The base university is UNIVERSITY1.The regression includes the following resume controls: demographic variables (sex, age and age squared, Internet at home, participation in team sports, choice of college major, hobby variables, and previous job market experience). The regression also includes SETWAGE, and SETWAGE interacted with BEAUTYand LNESTIMATED.