THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Beauty and the Sources of Discrimination

OPEN ACCESS

Beauty and the Sources of Discrimination[*]

Michèle Belot[†], V. Bhaskar[‡] & Jeroen van de Ven[§]

September 15, 2008

Abstract

We analyze behavior on a TV game show where players' earnings depend upon several factors. Attractive players fare better than less attractive ones, even though they perform no differently on every dimension. They also exhibit and engender the same degree of cooperativeness. Nevertheless, they are substantially less likely to be eliminated by their peers, even when this is costly. Our results suggest that discrimination arises due to consumption value considerations. We investigate third party perceptions of discrimination by asking experimental subjects to predict elimination decisions. Subjects' predictions implicitly assign a role for attractiveness but underestimate its magnitude.

Keywords: discrimination, beauty premium.

JEL Classification Numbers: C93, D63, J15, J16.

[†]Department of Economics and Institute for Social and Economic Research (ISER), University of Essex, mbelot@essex.ac.uk.

[‡]Department of Economics, University College London, v.bhaskar@ucl.ac.uk.

[§]Department of Economics and Amsterdam Centre for Law and Economics (ACLE), University of Amsterdam. j.vandeven@uva.nl.

1 Introduction

In a surprising and influential paper, Hamermesh and Biddle (1994) found a substantial beauty premium in the labor market, of the order of 15%. While there are several competing explanations for this premium, its source remains an open question. Attractiveness may be correlated with unobservable productive attributes such as health, education or other types of human capital. Attractive people may be more confident, thus enhancing their social skills in the workplace.[1] There may also be an element of reverse causality — individuals who fare well in the labor market may have both the ability and incentive (via greater self esteem) to invest in looking good.[2] Perhaps the simplest (and least palatable) explanation is that beauty has "consumption-value", either to the customers of the firm, fellow employees, or the boss. The beauty premium in this case is a form of taste-based discrimination, as discussed in Becker (1957). While anecdotal evidence on the importance of consumption value considerations in the hiring of air hostesses or waitresses certainly exists, the question remains whether this is a more general phenomenon. More generally, as the literature on racial/gender discrimination shows (see Altonji and Blank, 1999; Heckman, 1998), establishing discrimination and distinguishing between statistical and taste-based discrimination is difficult.

The main contribution of this paper is to disentangle the sources of advantage to attractive people. We do this in the context of a TV game show where participants are engaged in a variety of "tasks". We can ask: are attractive people more productive, and do they exhibit greater confidence? Are they more cooperative or do they engender cooperation? Are they more likely to be chosen by their peers when a selection decision has to be made? We are able to answer these questions since the game show has a rich structure, with players being involved in a number of different tasks

---

[1] See Mobius and Rosenblat (2006) for experimental evidence in favor of this hypothesis.

[2] Biddle and Hamermesh (1998) address the reverse causality problem in a study on lawyers by using a measure of beauty based on photographs taken at law school.

and decisions. It takes place over three rounds, in which players accumulate "earnings" by answering quiz questions, and their earnings depend on the accuracy of their answers, on how quickly they press the buzzer and also on their "investment decisions". Earnings therefore depend upon ability as well as a player's confidence. This allows us to study the effect of attractiveness as well as other player characteristics upon performance. At the end of each round, the lead player — the one with the highest earnings — decides which one of the remaining players to eliminate. This allows us to study the role of attractiveness and gender (in addition to performance) upon the selection decision. After the final round, when only two players remain, they play a prisoner's dilemma game, allowing us to study the relation between beauty and cooperativeness. The median stake in this prisoner's dilemma game is €1,683, so that the monetary consequences of players' decisions are substantial.

We find that attractive players fare significantly better than unattractive ones. Players can only make positive earnings by making it to the final prisoner's dilemma stage of the game show. Only 27% of the least attractive players make it to the final round, as against 49% of the most attractive ones. This difference cannot be attributed to any aspect of performance — attractive players fare no differently from unattractive ones in answering questions or in investment behavior. They are no more likely to cooperate in the final stage, and opponents also behave no differently vis-a-vis them. Nevertheless, when one player has to be eliminated by the lead player, the least attractive player of a show is significantly more likely to be chosen.

Our results support a "consumption value" basis for discrimination. Discriminating in favor of attractive players is not simply used as a tie-breaking rule between otherwise similar players. Unattractive players are more likely to be eliminated even when they have a higher score than others. This is costly to the lead player, implying an adverse selection in terms of earning potential. We estimate that the average cost to lead players from discrimi-

nation in the final round alone amounts to about €440, i.e. about 25 percent of the median stake. This is an underestimate of the overall costs of discrimination over *all the rounds*. In view of our finding that attractiveness is unrelated to cooperativeness in the prisoner's dilemma, this cost has no offsetting financial benefit.

It is noteworthy that we find discrimination against the unattractive on a TV show, where each player's performance is clear-cut and where the lead player's decision is subject to public scrutiny. Other studies using TV shows (Levitt (2004), Antonovics et al. (2005) and List (2006)) find no evidence of discrimination on the basis of race, gender or ethnic background, but some weak evidence of discrimination against older players (these papers do not examine the role of beauty). Discrimination on the basis of looks may not be so imprinted in social consciousness as racial or gender discrimination, so that people are not so aware of the possibility of discrimination, i.e. such discrimination may be *insidious*.[3] To investigate third party perceptions of discrimination, we ran an experiment where subjects watched the game show and had to predict elimination decisions. Our subjects predict that attractive players are less likely to be eliminated, although they substantially underestimate the magnitude of this effect. We also ask our subjects to list qualitative factors that influence elimination decisions, and find that very few subjects mention attractiveness. Thus third parties appear to be aware of the possibility of discrimination subconsciously rather than consciously. We also present suggestive evidence that discrimination on the basis of looks appears to entail less social opprobrium than racial discrimination, making it easier to survive and persist.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the game show and our construction of the measure of attractiveness. Section 4 analyzes behavior on the game show, in terms of performance and cooperation. Section 5 studies the selec-

---

[3] Insidious is defined as "working or spreading harmfully in a subtle or stealthy manner" (answers.com).

tion decisions by lead players and establishes that attractive players benefit in this context. We also examine gender differences and find significant differences in performance between men and women, with the result that there does not appear to be any evidence of discrimination against women. This section also presents our experimental findings on third party perceptions of discrimination. The final section concludes.

2 Related Research

Following the work of Hamermesh and Biddle (1994), who find a beauty premium in the labor market in a variety of occupations, several papers have replicated its findings, and also attempted to disentangle the components of the premium. Biddle and Hamermesh (1998) analyze a sample of lawyers, and find a premium irrespective of their area of expertise, including the self-employed. They argue that the most plausible explanation is taste-based discrimination by clients. Mocan and Tekin (2006) find that unattractive people sort into criminal activity due to the existence of a beauty premium on the legal labor market.

There are two difficulties with field studies. First, data on attractiveness is rarely available. Second, it is difficult to disentangle attractiveness from ability, since productivity is rarely observable. This makes it hard to establish whether the beauty premium is due to productivity or discrimination. Such productivity effects are sometimes present. Landry et al. (2006) find that attractive female solicitors are more productive fund-raisers. Pfann et al. (2000) study a sample of Dutch advertising firms and find that those with better looking executives have higher revenues — this is particularly pertinent for our paper, since it demonstrates that beauty plays a role in the Netherlands.

Laboratory experiments are better suited to disentangling the sources of the beauty premium, since they can be designed for this purpose. Mobius and Rosenblat (2006) take this approach, using university students in Ar-

gentina. They find that the beauty premium appears to be partly due to the fact that attractive people are more confident. Since our results differ from theirs – we find that beauty is unrelated to actual confidence, as reflected in behavior, although it is correlated with third party perceptions of confidence — we discuss their paper in more detail in section 4.1.1, while presenting these findings. Benjamin and Shapiro (2006) find that experimental subjects are able to predict the electoral fortunes of candidates on the basis of 10-second silent video clips, suggesting that attractiveness (or charisma) plays an important political role as well.

There is a large literature on discrimination on the basis of race or gender (see Altonji and Blank (1999) for a survey). One difficulty in labor market studies is in establishing discrimination — see, for example, the criticisms of Heckman (1998). Bertrand and Mullainathan (2004) conduct an field experiment which circumvents these criticisms and shows convincing evidence of employer discrimination against African-American names. However, it remains an open issue whether such discrimination is statistical or taste based, since employee performance on the job is not observable.

3   Description of the data

3.1   The game show

We use data from all 69 episodes of the game show 'Does (s)he share or not?',[4] broadcast in the Netherlands in 2002, with 345 contestants in total. In the preliminary stage of the game, six prospective players choose their initial capital, a number between one and 100. The player with the highest choice is eliminated, leaving five players for the game proper. Figure 1 shows the structure of the game.[5]

The game then proceeds with three quiz rounds, where players accumu-

---

[4] The name of the show in Dutch is 'Deelt ie 't of deelt ie 't niet?'.

[5] The format of this preliminary elimination stage was slightly different in the first few episodes.
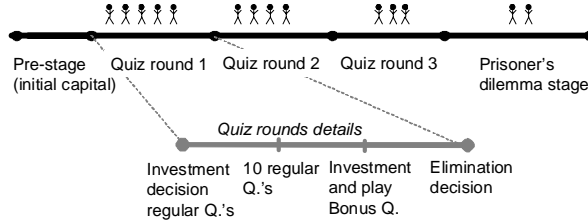
Figure 1: Timeline of the game show.

late earnings. Every round has ten regular questions and a bonus question. Players first choose how much of their capital to "invest" in answering each question (y_i). Whoever presses the buzzer first gets to answer. A correct answer yields y_i, while an incorrect answer earns $-y_i$. A player whose capital falls below his or her chosen investment may not answer any further questions. Each round ends with a bonus question, where players compete for the right to answer the question by choosing new investments. At the end of the round, the player with the highest score at that point — the *lead player* henceforth — must select one of the remaining other players for elimination. An eliminated player has no further role in the game and loses all of his or her earnings. The show then proceeds to the next round, where all players start with a capital that equals the earnings of the lead player in the previous round.

The last two remaining players play a prisoner's dilemma game. Let E denote the total prize money, which equals the sum of earnings of the two finalists. The finalists simultaneously decide whether to share or to grab. The monetary payoffs, as depicted in Table 1, correspond to a generalized prisoner's dilemma, where grab is a weakly dominant strategy.

Table 1 - *Monetary Payoffs*

	Share	Grab
Share	$\frac{1}{2}E, \frac{1}{2}E$	$0, E$
Grab	$E, 0$	$0, 0$

Table 2 presents summary statistics of the game show and the players. The total prize at stake varies between €380 and €26,600, with a median value of €1,683.[6] Players choose to share 43% of the time in the final round. The age of players varies between 18 and 64 with an average of 34.4. About a third of players is women. Of the third of the players who report their occupation, one-third is student, while the others are drawn from all the important sectors (at the two-digit classification level), including trade, information technology, education, financial services and health.

Table 2 - *Summary statistics, game show players*

	Mean	Min	Max
Mean age (years) (N = 345)	34.4	18	64
Percentage women (N = 345)	34.8	-	-
Prize (€) (N = 138)	2,976	380	26600
Percentage sharing (N = 138)	42.8	-	-

3.2 Measure of beauty

Hamermesh and Biddle (1994) argue that there are consistent standards of beauty within and even across cultures, so that subjective evaluations of attractiveness include a "common component". Our goal is to see how this common component affects performance on the game show. To this end, each of the 345 participants on the game show were rated on a scale from 1 (very unattractive) to 7 (very attractive) by approximately 10 raters, balanced by gender.

Raters were recruited in public spaces to obtain a representative sample of the adult population. Raters were on average 31.7 years old, which is close to the average age of game show participants, 33 years. We recruited 120 raters, and each rated 30 participants. This was based on watching short silent video fragments of the game show in which a player introduced him or herself. We ensured that all five players on any show were rated

---

[6]These are considerable sums given that the median disposable monthly income of a full-time employed person in the Netherlands was about €1,200 in 2000 (*Statistics Netherlands*, available at www.cbs.nl).

by the same set of raters, while varying the order in which the shows were presented. About one-third of the raters were non-Dutch and could not be familiar with the show, and only a small minority of the Dutch raters indicated that they had seen any episode of the show. Given this, and the fact that the game show was broadcast five years prior to the ratings, our results are not distorted due to familiarity with the faces.

Our measure of attractiveness is the average of the independent ratings (across raters) for each player.[7] There is a high degree of concurrence on attractiveness across raters. Across sub-samples of raters who rated the same sample of players, the Cronbach's alpha ranged from 0.70 to 0.85, showing high agreement. Table 3 reports summary statistics of the ratings. Raters were told to use the benchmark average attractiveness in the population at 4. Somewhat surprisingly, the average rating across participants is 3.51, i.e. lower than the benchmark, so that there does not appear to be any positive selection into the show according to attractiveness. Beauty is negatively correlated with age and women are, on average, rated as being more attractive than men. Average ratings are more variable across women than men, consistent with other studies (Hamermesh and Biddle, 1994). We also coded objective participant characteristics that are related to their appearance, such as eye-color, hair-length, wearing of glasses, etc. These characteristics are correlated with our beauty measure — for women, beauty is positively correlated with hair-length, while for men, beauty is negatively correlated with having a beard or moustache and wearing glasses. This suggests that our measure of beauty captures aspects of physical attractiveness.

---

[7]Raters may have different perceptions on the average beauty. To correct for such differences, some other studies use standardized measures. Each rating is adjusted for the mean rating of that rater, and is then normalized by dividing by the standard error (see e.g. Mobius and Rosenblat, 2006). We chose to have many different raters rating different subjects. The means are not comparable among raters, because the sample of episodes they rated were partly different. We therefore prefer to use the raw data. However, if we standardize ratings for the subsamples of ratings made for the same players, the results remain very similar.

Table 3 - *Summary statistics attractiveness*

	Mean	(st. dev.)	Min	Max
All ($N = 345$)	3.51	(.69)	1.7	5.75
Men ($N = 225$)	3.45	(.63)	2.0	5.20
Women ($N = 120$)	3.62	(.79)	1.7	5.75
Age $\geq$ 34 ($N = 177$)	3.30	(.60)	1.7	4.80
Age $<$ 34 ($N = 176$)	3.71	(.71)	2.2	5.75

Note: attractiveness of players is averaged across raters.

4 Beauty and Behavior

Given the many potential determinants of performance and monetary pay-offs on the show, we study each of these in turn.

4.1 Beauty and performance

We first investigate the relationship between beauty and performance in answering quiz questions. The total earnings of a player in a round is probably the most important measure of overall performance. The player with the highest earnings becomes the lead player for that round, making the elimination decision. In the first two rounds the lead player's earnings determine the initial capital for all remaining players in the next round. In the third round, a player's earnings are added to the total stake, if he is either the lead player or not eliminated by the lead player. Our focus is mainly on the first round, since this is the round with the least prior selection.

Since players compete to answer each question, their performance in the game is a relative measure. If there is a relation between beauty and performance then this will depend on the composition of players within an episode. Hence, we cannot draw conclusions from simple cross-correlations over episodes. Instead, we look at the relationship between attractiveness and relative ranking in the game. Two player positions, the first and the last, are of particular interest. The player who is ranked first in terms of earnings must choose one of the others for elimination, and cannot be eliminated himself in that round. The last ranked player is an obvious candidate for

9

elimination. We estimate a conditional logit model for the probability of being in the first or last position conditional on participating in the same episode. The conditional logit is a natural framework for modelling choices from a set of alternatives. In our context, the alternatives are the players in the round, and each player i in show j has a vector of attributes X_{ij} (gender, age, attractiveness). The conditional logit model has the form:

$$p(y_{ij} = 1) = \frac{\exp(\beta' X_{ij})}{\sum_i \exp(\beta' X_{ij})} \qquad \text{for } i = 1, ...5,$$

where y_{ij} is an indicator variable which takes value one when the player is in the first (respectively last) position.[8]

The results are reported in Table 4, columns 1 and 2. In addition to attractiveness, we also control for age and gender, and for a number of dummy variables describing the register of the introductory speech: whether the player mentions his marital status, whether he has children or not, talks about his profession or studies, or about his hobbies. Attractiveness is uncorrelated with the probability of ending first or last. Only one covariate turns out to be significant; players who mention their marital status are less likely to be ranked first. To investigate further the correlation between the players' attributes and their rank in terms of earnings, we estimate a rank-ordered logit model (a refinement of the conditional logit) that explicitly takes account of the ranking of players within a game and specifies this ranking as a function of their relative attributes. The results are reported in the third column of Table 4. We find no clear correlation between any of these characteristics and earnings ranking. In particular, there is no evidence that attractive people rank differently from unattractive ones.

---

[8]One important assumption for the validity of the conditional logit estimates is independence of irrelevant alternatives. Our tests show that this assumption is not rejected.

Table 4 - *Attractiveness and performance, 1st round*

	(1) Prob. ranked first		(2) Prob. ranked last		(3) Score ranking end of 1st round	
Attractiveness	.14	(.23)	.18	(.22)	-.06	(.12)
Age	.00	(.02)	.01	(.02)	-.01	(.01)
Female	-.32	(.29)	.02	(.27)	-.09	(.15)
Register of introductory speech						
- is married / has a partner	-2.17	(.75)***	-.30	(.54)	-.16	(.28)
- children	.67	(.45)	.04	(.44)	.08	(.24)
- study / profession	.24	(.31)	.09	(.31)	-.09	(.17)
- hobby	-1.70	(1.45)	-.31	(.93)	.09	(.49)
N. obs.	345		345		345	
Pseudo R-squared	.06		.01		-	
Method	Cond. Logit		Cond. Logit		Rank-ordered. Logit	

Notes: *** 1% significance level, in col (3), 1 is lowest score and 5 highest.

4.1.1 Decomposing performance

Performance in this game depends on a number factors: the choice of initial capital and investment, the decision to press the buzzer or not, and, finally, the answer itself. These decisions depend upon the player's ability, confidence, and risk aversion. Mobius and Rosenblat (2006) argue that confidence explains a substantial part of the beauty premium. They find that "employers" have higher estimates of the productivity of more attractive individuals, even though they are given independent evidence on productivity. An intriguing finding is that attractive subjects are estimated to have higher productivity even when their interaction with the employer is only oral, not visual. Mobius and Rosenblat attribute this to the higher self-confidence of attractive workers, and this explains about 20% of the beauty premium.

In the context of this game, we would expect highly confident players to be more active in the game in terms of investment and answering questions.

Recall that by being the first to press the buzzer, player i faces a lottery where he gets y_i (his chosen investment level) if he is correct, and gets $-y_i$ if incorrect. Let p_i be the subjective probability assigned by i to his answer being correct.[9] Let x_i be his current score, and let $V_i(.)$ denote his expected continuation value in the game after this question. It is optimal for player i to answer to the question if his expected continuation value from answering is weakly greater than his value at the current score. This implies that player i will answer the question if p_i exceeds a critical threshold:

$$p_i \geq \frac{V_i(x_i) - V_i(x_i - y_i)}{V_i(x_i + y_i) - V_i(x_i - y_i)} \equiv \bar{p}_i,$$

where $\bar{p}_i \in (0, 1)$ as long as V_i is strictly increasing in the player's earnings. Notice that the threshold value, $\bar{p}_i$, is larger if the player is more risk averse, i.e. if $V_i(.)$ is more concave. Since we observe the frequency of correct answers for every player, we have an estimate of q_i, the objective probability that a player is correct conditional on answering. That is, we have an estimate of $\mathbf{E}(q_i | p_i \geq \bar{p}_i)$. A player who is more risk averse will have a higher threshold value $\bar{p}_i$, and will therefore answer fewer questions but be observed to answer a greater proportion of questions correctly. A player who is more confident — i.e. has a greater value of p_i for a given q_i — will answer more questions and will make more mistakes. In other words, risk aversion and lower confidence act in very much the same way, in reducing both the number of answers and also the proportion of incorrect answers. On the other hand, if a player is less knowledgeable, and objectively has a lower value of q_i, this will *ceteris paribus* reduce the number of answers but not raise the proportion of correct answers.

We concentrate our analysis on the decisions made in the first round, since the second and third rounds could be subject to selection biases due

---

[9] As the player thinks about the question, his subjective probability will evolve over time. The analysis that follows pertains to any instant of time, so the continuation value from not pressing the button ($V_i(x_i)$) includes the option value of waiting, and possibly pressing the buzzer in the future, if no one else presses in interim.

to the elimination decision.[10] Table 5 presents the results of OLS estimates of the determinants of initial capital, share invested, number of answers and percentage of good answers (columns 1 to 4 respectively).[11] We do not find that attractive players are more confident. If they were, they should invest more, be more likely to answer, and, conditional on answering, perform worse. We do find a systematic difference according to gender: Women are much less likely to answer a question. This is what you would expect if women are more risk averse or less confident. However, conditional on answering, they actually do not perform better than men, what you would expect with risk aversion or lack of confidence. In the second round, we even find that they are significantly less likely to answer correctly. We also did not find any evidence that the topics of the questions could explain gender differences in performance. Overall, these results suggest that the reason why women are less likely to answer is not due to a wrong perception of their ability or a higher degree of risk aversion, but rather because they are less able to answer the type of questions on the show.

Our measure of confidence is based on actual behavior. It could be the case that attractive players are *perceived* more confident by other players. To investigate this, we constructed a measure of perceived confidence, by having independent raters assess the confidence of a player on a 7 point scale. Each player was rated by 5 to 10 raters, and each rater saw 35 players. Our player specific measure of confidence is the mean across raters. Column (5) of Table 5 shows the results of an OLS regression of perceived confidence on the characteristics of the players. We find that perceived confidence is significantly positively correlated with attractiveness. Thus, attractive players *appear* more confident but do not behave more confidently. The reason for

---

[10]We have analyzed separately the decisions made for the bonus question. Again, we found that attractiveness is uncorrelated with the share invested or the probability of giving a correct answer. The results are not reported for the sake of brevity.

[11]Alternative econometric specifications (conditional logit estimates for the probability of answering and logit estimates for the probability of answering correctly) give similar results to the ones presented here.

Table 5 - *Attractiveness, performance and confidence, 1st round*

	(1) Initial capital	(2) Share invested	(3) N. of answers	(4) Share correct answers	(5) Perceived confidence
Attractiveness	1.29 (1.05)	-.03 (.02)	.00 (.01)	-.01 (.03)	.18 (.06)***
Age	.28 (.09)***	-.004 (.002)**	.00 (.12)	-.002 (.002)	.01 (.01)
Female	-.02 (1.16)	.00 (.02)	-.50 (.17)***	.03 (.05)	-.30 (.08)***
Register introductory speech					
- is married/partner	.02 (2.27)	.04 (.05)	-.50 (.22)**	-.04 (.08)	.03 (.15)
- has children	-.71 (1.58)	-.04 (.04)	.17 (.21)	.07 (.07)	.09 (.14)
- study / profession	1.64 (1.45)	.05 (.03)*	-.02 (.18)	.00 (.04)	.17 (.08)**
- hobby	-8.54 (2.92)	-.04 (.06)	.08 (.24)	-.06 (.09)	-.18 (.15)
Constant	33.98 (5.81)***	.90 (.10)***	1.77 (.71)**	.90 (.15)	3.36 (.31)***
N. obs.	345	345	345	271	3235

Notes: OLS estimates, * significant at 10% level, ** 5% level and *** 1% level, col (4) include players who gave at least 1 answer. Standard errors are clustered by episode in col (1) to (4). Col. (5) includes predictor fixed effects and standard errors are clustered at the player level.

our finding may be that players receive continuous feedback on their performance. This contrasts with the set-up of Mobius and Rosenblat, where subjects did not receive systematic feedback on their performance through the experiment. We should also mention that in their study, there is no interaction between subjects who are assigned the role of "employers" and those assigned the role of "workers", so that consumption value considerations may be naturally less important.

4.2 Beauty and cooperativeness

Players can only materialize their gains in this game in the final stage, where the two remaining players play a prisoner's dilemma game and decide simultaneously to share the accumulated money or not. A companion paper (Belot et al., 2006) conducts a comprehensive analysis of the determinants of sharing behavior. The key findings are that own characteristics

matter — specifically, women are more likely to share than men. However, the characteristics of the opponent turn out to be irrelevant to the sharing decision.

We augment this analysis by including the player's own beauty rating and the opponent's beauty rating as explanatory variables. The results are reported in Table 6. We find no correlation between beauty and cooperative behavior. Attractive players are no more (or less) likely to share — indeed, the coefficient is very close to 0.[12] Attractive opponents are also no more (or less) likely to induce sharing behavior from their opponents. This is interesting — although our overall results suggest that players obtain consumption value from having attractive co-players, they are no more likely to share with them.[13]

It is possible that attractive people are *perceived* as being more cooperative even though they are not really so, so that selection decisions are driven by incorrect beliefs. While we do not observe the beliefs of participants on the show, a companion paper (Belot et al. (2008)) experimentally investigates third party perceptions of trustworthiness. Experimental subjects watched a random sample of shows and were asked to report a probability that a player would share. Column 3 of table 6 reports our findings on the relation between beauty and perceived cooperativeness. Our subjects predict that attractive people tend to be slightly *less* cooperative, although the effect is small. It therefore appears that attractive people are not only not more cooperative, they are also not perceived to be more cooperative by third parties.

---

[12] The coefficients are potentially biased if lead players use private signals to select players, since the sample of unattractive players may be unusually cooperative. Belot et al. (2006) addresses this issue and find no evidence of bias.

[13] Other experimental evidence on the relation between attractiveness and cooperation is mixed — see Mulford et al. (1998), Solnick and Schweitzer (1999), Eckel and Wilson (2004) and Andreoni and Petrie (2005).

Table 6 - *Attractiveness, sharing & perceptions of sharing*

	(1) Pr. sharing		(2) Pr. sharing		(3) Perceived prob. sharing	
Own attractiveness	-.03	(.07)	-.04	(.06)	-.031	(.012)**
Opponent's attractiveness			-.01	(.11)		
Age	.00	(.01)	.00	(.01)	.001	(.001)
Female	.19	(.09)**	.21	(.10)**	.091	(.015)***
Contribution to prize money (%)	-.72	(.31)**	-.82	(.35)**	-.066	(.032)**
Total prize (x €1,000)	.03	(.01)**	.03	(.01)***	.008	(.002)***
Constant					.481	(.100)***
N. obs.	138		138		1672	
Method	Bivariate probit		Bivariate probit		OLS	

Notes: * significant at 10% level, ** 5% level and *** 1% level; (1) and (2) are
bivariate probit estimates, standard errors clustered by episode;
the marginal effects in (1) and (2) are computed at the means of the lead
player's characteristics. (3) includes subject fixed effects.

5   Beauty and selection

We now study the elimination / selection decision, having established that
there is no objective reason to discriminate in favor of attractive players
either on the grounds of performance or because they are (thought to be)
more cooperative. Thus any bias towards attractive players in lead player
selection decisions can plausibly be attributed to the lead players obtaining
consumption value from having attractive co-players.

An important advantage of the rules of our game show is that in mak-
ing the elimination decision, the lead player in any round is faced with a
relatively simple *decision* problem, rather than a game. If the lead player
chooses to eliminate player i then the lead player is decisive and i will play
no further part in the game. In contrast, elimination decisions in other game
shows (such as *The Weakest Link*, analyzed by Levitt (2004) and Antonovics

et al. (2005)) are often made by majority voting, involving all the participants remaining at that stage. If a player j votes to eliminate i, then i may not be eliminated, and may in turn vote against j at a later stage. This implies that players have a strong incentive to vote to eliminate whoever they think others are going to vote against. In other words, there are multiple voting equilibria, and this is coupled with a strategic dynamic motive to vote with the majority. This may induce a significant role for irrelevant characteristics as possible focal points, even when players do not have any preference for discriminating on the basis of such a characteristic. In the context of our game, these strategic considerations do not apply, since only the lead player votes and his vote is decisive. Thus evidence of discrimination can be attributed to lead player preferences.

Figure 2 shows the average attractiveness of all players in each round, and that of the players who are eliminated in that round. In each round, eliminated players are less attractive than average, and in consequence, average attractiveness increases steadily over the rounds. Other summary statistics confirm this picture. If a player is average-looking (i.e. within one standard deviation of the mean), he or she has 0.4 probability of reaching the final round. An attractive player has a substantially higher probability of 0.51, while an unattractive player's probability is only 0.31 (see Table 7).

<div align="center">Table 7 - Attractiveness and survival, by round</div>

	% reaching round 2	% reaching round 3	% reaching final
Attractive	83	62	51
Average-looking	81	62	40
Unattractive	72	52	31

Note: Attractive (unattractive) is more than one standard deviation above (below) the mean.

We investigate in more detail the role of physical attractiveness in the selection decision by the lead player by estimating a conditional logit model, where the dependent variable indicates whether the player was eliminated
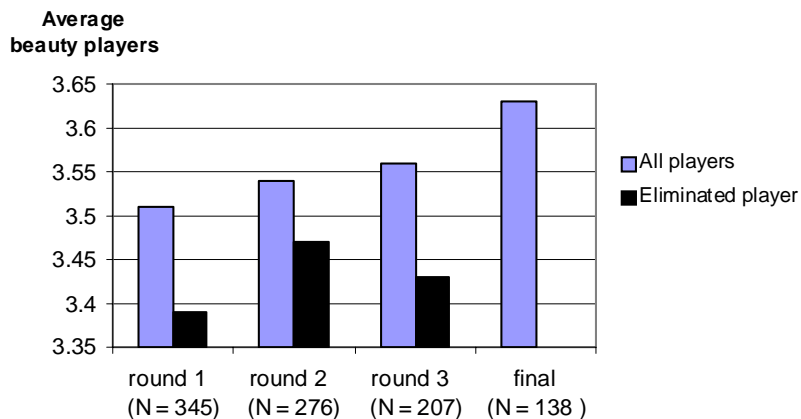
Average beauty players

Figure 2: Average beauty of all players and eliminated players.

(1) or not (0). We start by focusing on the first round, where the sample is largest and no selection has taken place. Later on we discuss results for other rounds.

The results are shown in Table 8. The first regression includes no other covariates, the second controls for performance and individual characteristics and the third controls for measures of confidence. The results are reported in terms of odds ratios with a reference value of 1. These have an easy interpretation. An odds ratio above (below) 1 signifies that the attribution increases (decreases) the probability of being eliminated.

We find that unattractive players are significantly more likely to be eliminated than average-looking players. When we control for performance and measures of confidence, we predict that unattractive players should be three times as likely to be eliminated as average-looking players. In contrast we do not find that more attractive players are better treated. Hence, our results point in the direction of an ugliness penalty, in line with the results of Hamermesh and Biddle (1994).

18

Table 8 - *Probability of being eliminated, 1st round*						
Conditional logit estimates (odds ratios)						
		(1)		(2)		(3)
Attractiveness						
Unattractive	2.37	(1.05)*	2.91	(1.43)**	3.11	(1.57)**
Attractive	1.29	(.39)	1.21	(.44)	1.15	(.43)
Performance						
Fourth highest			.47	(.15)**	.42	(.14)***
Third highest			.26	(.10)***	.23	(.09)***
Second highest			.18	(.08)***	.16	(.08)***
Confidence						
% capital invested					.43	(.33)
n. answers					.98	(.13)
perceived confid.					1.22	(.38)
Female			.67	(.21)	.70	(.24)
Age			1.00	(.02)	1.00	(.02)
Controls for topic						
registry introduction	no		yes		yes	
N. obs.	276		276		276	
Pseudo R-squared	.02		.20		.20	

Notes: Significance levels: *: 10 percent, **: 5 percent, ***: 1 percent.

Note that age and gender are irrelevant in the selection decision. Also, the score ranking is a very good predictor of elimination: the player with the lowest score (the reference category) is more than twice as likely to be eliminated as the one ranked fourth, and more than five times as likely to be eliminated as the one with the second highest score. Finally, controls for behavior during the game do not change the results and do not matter as such in the selection decision (column (3)). Less attractive players are discriminated against, for reasons that are uncorrelated with their performance or behavior during the game.

5.1 Discrimination over the rounds

One explanation for the discrimination taking place in the first round is that players have very little information about each other. They had relatively little time to get to know each other and to learn about each other's ability.

With so little information, perhaps they retreat to attractiveness to select one player over another.

If this is the reason, we expect discrimination to disappear over the rounds when more information becomes available. However, since discrimination takes place in the first round, the population of players becomes more attractive on average and less heterogeneous, which should reduce discriminatory practices. The fact that there are fewer unattractive people around also makes it harder to identify an effect. So we also propose a specification based on a relative measure of beauty, where we identify the least attractive player among the remaining contestants with a dummy variable.

The first two columns of table 9 show estimation results comparable to column (3) of Table 8, for the second and third round respectively. Columns (3) to (5) show the results for the specification including the dummy for the least attractive player for each round, including the first one.

We find no evidence of discrimination based on attractiveness in the second round. Gender is the best predictor of the elimination decision in the second round, with women being substantially more likely to stay in the game. We will come back on that point in the next section. In the third round, we find evidence of discrimination against the least attractive player, and the coefficient is comparable to the first round. So discrimination does not disappear with learning about player ability.

Table 9 - *Probability of being eliminated in 2nd and 3rd round*

Conditional logit estimates (odds ratios)									
	(1)		(2)		(3)		(4)		(5)
Round	2		3		1		2		3
Unattractive	.63	(.35)	1.02	(.84)					
Attractive	.69	(.31)	.43	(.26)					
Least attractive					2.16	(.67)**	1.00	(.33)	2.34	(.92)**
N. obs.	207		138		276		207		138

All regressions control for the same variables as Table 8, col. (3), ** indicates significance level at 5 percent

5.2 Gender differences in discriminatory behavior

We now explore the nature of discrimination in greater detail. We investigate the relation between discrimination and gender. Do men care more about looks than women do? Do people care more about the looks of the opposite sex? Table 10 reports separate regressions for male and female lead players. We find that women do discriminate more against the least attractive player than men do, i.e. women seem to care more about looks than men do. The difference in coefficients between male and female lead players is not significant though. Columns (3) and (4) investigate the elimination decision when the lead and least attractive player are of opposite sex on the one hand, and of the same sex on the other hand. We find that discrimination is indeed greater when the candidates for elimination are of the opposite sex. The difference in coefficients is significant at the 10% level. This could explain why women discriminate more, since the least attractive player in the first round is more likely to be a man (men being on average less attractive than women). We have found almost identical results for the last selection round, where the discrimination against the least attractive player took place in games where the lead player and the least attractive player were of opposite sex (results not reported for the sake of brevity). Since in the last round, most lead players are men, the discrimination plays mostly against women.

Overall, these results reinforce the idea that beauty has a consumption value, and that this is the main reason why people discriminate against less attractive players.

Table 10: *Discrimination and gender of the lead player (1st round)*

Conditional logit estimates (odds ratios)

	Lead player female		Lead player male		Lead player and least attractive - opposite sex		Lead player and least attractive - same sex	
	(1)		(2)		(3)		(4)	
Fourth highest score	.24	(.18)*	.50	(.18)*	.36	(.19)**	.53	(.23)
Third highest score	.24	(.19)*	.39	(.16)**	.51	(.25)	.18	(.11)***
Second highest score	.21	(.15)**	.13	(.08)***	.31	(.16)**	.05	(.05)***
Least attractive	4.55	(3.11)**	1.39	(.48)	2.80	(1.19)**	1.10	(.53)
Age	.95	(.04)	1.01	(.02)	.98	(.02)	1.03	(.03)
Gender	.93	(.67)	.70	(.24)	.68	(.31)	.56	(.27)
N. obs	84		192		140		136	
Test equality coefficients for the least attractive								
(1) = (2) P-value	.10							
(3) = (4) P value					.16			

Notes: Equality of coefficients is tested with a generalized Hausman test

Significance levels: *: 10 percent, **: 5 percent, ***: 1 percent.

5.3 Costs of discrimination and the benefits of beauty

We now provide an estimate of the benefits of beauty and the costs of discrimination. As in many other game shows, the rules are such that most contestants go home empty handed. In fact, only 59 out of 345 contestants (i.e. 18%) take any money home. Moreover, even among those with positive earnings, the variance is very large (the standard deviation is €2,758 compared to the mean of €2,570). These factors imply that there is a high degree of noisiness in earnings, making it a priori difficult to get statistically significant results. Furthermore, given the large number of zeroes, our estimates are not directly comparable to estimates of the beauty premium from the labour market such as those of Hamermesh and Biddle (1994).

Column 1 in table 11 presents OLS estimates of the determinants of log earnings. An increase in attractiveness of one standard deviation increases

earnings by 41 percentage points. Despite the large standard error, this is significantly different from zero at the 10% level. An alternative estimate of the beauty premium is based on estimating the probability of reaching the final round. Columns 2 and 3 provide estimates of the probability of reaching the final round with attractiveness as a continuous variable (column 2) or as a discrete variable, based on the ranking of players within each show according to attractiveness (column 3). The results confirm our previous findings: the two most attractive players are about twice as likely to reach the final round than the least attractive player (50% against 27% with a standard error of 8 percentage points). This means that, on average, their expected earnings are twice as large. Note that if we take a very conservative approach and consider the lower bound of the 95% confidence interval for the probability of reaching the final round, we still have a difference of 7 percentage points in terms of probability of reaching the final round between the most and least attractive players. In this case, the estimated lower bound on the beauty premium equals 9%, which remains substantial.

Table 11 : *The beauty premium*

	Log earnings		Probability of reaching final round			
	(1)		(2)		(3)	
Gender	-.46	(.31)	1.07	(.23)	1.12	(.24)
Age	-.005	(.018)	.98	(.01)	.98	(.01)
Attractiveness	.41	(.24)*	1.42	(.26)*		
Beauty rank 5 (least attractive)					-	-
Beauty rank 4					1.57	(.52)
Beauty rank 3					1.29	(.45)
Beauty rank 2					2.14	(.73)**
Beauty rank 1 (most attractive)					1.89	(.64)**
Constant	.15	(1.13)				
N. Obs.	345		345		345	
Pseudo R-squared	.01		.03		.04	

Notes: Col. (1): OLS estimates allowing for a show-specific random effect; Cols (2) and (3): conditional logit estimates (odds ratios); **: 5 percent, *: 10 percent.

Turning to the costs of discrimination, stakes are substantial on the

game show. By eliminating the least attractive players instead of players who would maximize their monetary payoff, players implicitly pay a price for keeping more attractive players in the game. While we cannot directly calculate the price for keeping more attractive players in the game in the first or second round (since we do not observe the earnings of those who are eliminated), we can do a back-of-the-envelope calculation of the price they pay by eliminating the least attractive player in the third round, where we do observe perfectly the earnings of the contestants. However, the sample of observations is relatively small and the results should be taken with caution.

We can identify 10 of the 69 episodes where the lead player eliminates the least attractive player, even though he does not have the lowest score. In these episodes, the less attractive players have earned €440 more on average than the player who is chosen to play the final. Hence, by eliminating the least attractive players in these cases, the lead player diminishes the prize money E by €440 on average.[14] Notice that discrimination is directly financially costly in the final round — there are only two instances where the two low ranked players had the same score. This estimate is a lower bound since we are only looking at costs associated with discrimination in the third round, and not earlier rounds, since a precise imputation of financial costs in earlier rounds is more difficult.

5.4 Gender differences

It is worthwhile contrasting our results on beauty with differences between players based on gender. Overall, women earn less than men, earning €269 on average, as compared to €428. As with attractiveness, we can decompose the gender gap in earnings into different components. In terms of performance, we found that women were less likely to answer a question, and conditional on answering were no more likely to answer correctly (in round

---

[14]It might be that lead players expect those with a lower score to be more cooperative, for which there is evidence (in our companion paper). But this doesn't explain why the player with a lower score is rarely chosen to play the final if he is is the least attractive player.

3, they are actually less likely to answer correctly). The lower performance of women could be due to the competitive character of the game show — Gneezy et al. (2003) present experimental evidence showing that women perform worse in competitive environments, especially when they compete against men.

Despite the fact that women perform worse, they are as likely to reach the final round as men — 35% of the participants on the show are women, while 37% of the players in the final round of the game are female. One explanation for discrimination in favour of women in the selection decision is their greater cooperativeness, since women are more likely to share in the final round — 55% of women share against only 36% of men. However, this does not seem to be the explanation since women are no more likely to be selected than men in the *last* round (holding earnings constant), when the lead player selects the player with whom he will play the final round and where one may expect cooperativeness considerations to be dominant.[15] It is more plausible that women are chosen despite their lower performance due to gender balance considerations (since women are relatively scarce in the show)[16] or positive discrimination (since the selection decision is public).

Women's' greater cooperativeness widens the gender gap in earnings. Indeed, since women are much more likely to share in the final round, and since the opponent does not share more often when facing a woman, the difference in take-home earnings across gender is larger than between individual scores.

Overall, we find no evidence of negative discrimination against women. The gender earnings gap is mainly driven by differences in performance and cooperativeness. Perhaps the fact that the selection decision is public prevents gender discrimination — since gender is an objective and obvious

---

[15] Our companion paper provides a comprehensive analysis of the determinants of the sharing decision in the final stage, and its implications for how selection decisions are made just before this.

[16] Some participants explicitly mentioned "gender balance" as a consideration while making the selection decision in the second round.

characteristic in contrast to attractiveness, which is more subtle.

5.5   Insidious discrimination?

Are third parties aware of the possibility of discriminatory behavior? And
if they are aware, is this knowledge conscious or sub-conscious? These questions are relevant — if discrimination is not perceived by third parties, it
may persist even under public scrutiny.[17] To investigate these questions,
we adopt a novel experimental procedure, by asking subjects to predict the
elimination decision at the end of the first round of quiz questions. We focus
on the first round since we do not want subject predictions to be influenced
by learning from observing selection decisions in previous rounds. Our subjects (71 students from the University of Amsterdam) were shown a trailer
and given a handout, setting out the overall structure of the game show,
and then shown a random sample of seven episodes. They were informed
in advance of the identity of the lead player in the first round, so that they
could focus on predicting the lead player's decision. The subject was asked
to assign a probability to the event that each of the other players would be
chosen for elimination by the lead player. Subjects were rewarded by using
a quadratic scoring rule, giving them strict incentives to report their true
beliefs. Subjects were not informed of the actual elimination decision in
any episode, so as to prevent any learning. At the end, we asked subjects a
qualitative question: what, in their opinion, are the most important factors
determining the elimination decision? Subjects were permitted to list up to
four such factors.[18]

Table 12 reports ordinary least squares estimates, where the dependent
variable is the prediction of subject i regarding the elimination probability

---

[17]This is related but not identical to the notion of implicit discrimination (see Greeenwald et al. (1998) and Bertrand and Mullainathan (2005)), which uses psychological
tests to measure implicit bias, say against African-Americans. It is worth noting that
researchers have found an implicit bias against older people and the overweight.

[18]The subjects also filled out a questionnaire on their background characteristics. They
earned on average €18 (for 90 minutes), including a €4 show up fee. The full set of
instructions is in the appendix.

of player j. We have 71 subjects, making predictions for four players in each
of seven episodes, giving us 1988 observations.

Table 12 - *Predictions of elimination in 1st round*

OLS estimates						
	(1)	(2)	(3)	(4)	(5)	(6)
Attractiveness	-2.58	-2.43	-2.45			
	(1.03)**	(.84)***	(.84)***			
Dummy least attractive				2.24	2.38	2.42
				(1.55)	(1.29)*	(1.29)*
Perceived confidence			.20			.20
			(.18)			(.17)
Age	-.09	-.03	-.03	-.04	.02	.01
	(.07)	(.06)	(.06)	(.07)	(.06)	(.06)
Gender	-3.58	-3.58	-3.52	-3.96	-3.91	-3.85
	(1.39)***	(1.11)***	(1.11)***	(1.38)***	(1.10)***	(1.10)***
Score rank 4th		-6.76	-6.74		-6.83	-6.81
		(1.37)***	(1.37)***		(1.36)***	(1.36)***
Score rank 3d		-5.22	-5.21		-4.87	-4.86
		(1.37)***	(1.37)***		(1.36)***	(1.36)***
Score rank 2d		8.41	8.42		8.61	8.62
		(1.73)***	(1.72)***		(1.74)***	(1.74)***
Constant	38.29	36.71	35.95	26.98	25.93	25.08
	(5.19)***	(4.46)***	(4.50)***	(2.63)***	(2.27)***	(2.45)***
N. Obs.	1988	1988	1988	1988	1988	1988

Notes: Standard errors are clustered by player. Significance levels: *: 10 percent,
: 5 percent, *: 1 percent.

Our main finding is that our subjects do perceive that less attractive
players are more likely to be eliminated.[19] That is, our subjects are at
least implicitly aware that discrimination on the grounds of attractiveness
is likely to occur. They predict that an increase in one standard deviation
in attractiveness should decrease the probability of being chosen by 1.7 per-
centage points. This is substantially less than what we have estimated in

---

[19]Subjects do not perceive any relation between a player's confidence (as assessed by the
subject) and the probability of elimination. Women are predicted to be more likely to be
eliminated, as are second ranked players – neither of these effects are present in the actual
selection decision. The last finding appears to be due to subjects incorrectly thinking that
the first ranked player may have an incentive to eliminate rivals for leadership.

the actual data (the corresponding number is 9 percentage points). Interestingly, attractiveness is not mentioned by anyone among the two most important criteria influencing the selection decision, and only 4 subjects list attractiveness at all. This provides suggestive evidence that discrimination is insidious, since subjects appear to be implicitly aware of its existence, but not consciously so. This is particularly noteworthy since a subject is detailing the motivations of players on the game show (rather than himself), and would therefore have little psychological reason to lie or practice self-deception.[20]

Discrimination on the basis of looks appears to meet with less social disapproval than other forms of discrimination. Following a report in *New Scientist* summarizing the findings of the present paper, it was covered by newspapers in several countries including the UK and the Netherlands. Press coverage in the Dutch newspaper "De Telegraaf" triggered a large public response — within a day, 156 readers had posted a response on the newspaper's website. We coded these responses and found that 36 of these suggested that beautiful people "deserve" to be treated better than the "ugly" — some comments were quite offensive about unattractive people. 21 people simply acknowledged the results and gave further examples of discriminatory behavior against ugly people. Only six people explicitly condemned this kind of discriminatory behavior. Furthermore, many people simply found the findings somewhat amusing (as did some of the articles in the popular press), rather than a matter for social concern. While its readers are unlikely to be representative, the newspaper is the largest with over 700,000 circulated copies daily. It is noteworthy that we find this response in the Netherlands, a country that is well known for its progressive and liberal attitudes towards some minority groups (e.g. gay people), and one with a strong commitment to equal opportunities.

Our experimental findings (and the public response) raise important is-

---

[20]Fershtman and Gneezy (2001) find that experimental subjects who discriminate seem to be unaware that they do so.

sues of social concern regarding the persistence of discrimination. Following Becker (1957), economists have focused on the extent to which the forces of competition and profit maximization eliminate discrimination. It is arguable that the social stigma associated with racial or gender discrimination is no less important a force in its elimination. To the extent that discrimination on the basis of looks or age occurs, but is either insidious or is "acceptable" even when perceived, its persistence is enhanced.

5.6   External validity

One possible concern with our study is external validity, especially if the participants on the show are not representative of the population. While genuine field data on attractiveness and individual productivity might be ideal, this is hard to get, and laboratory experiments or our type of study seem to offer the best hope for disentangling the beauty premium.[21] The producer of the show told us that no explicit criteria were used in recruiting participants — in particular, applicants were not required to submit a photograph, so that the producers do not seem to have been concerned with having good looking people. Nor does it seem, from viewing the show, that attractive people self-select into the show. This impression is confirmed by the fact that the average rating of our participants in terms of attractiveness is 3.51, whereas raters were told to use 4 as the benchmark for average attractiveness.   Participants also come from a variety of occupations, an advantage that compares favorably with laboratory experiments. Given the lack of emphasis by the producers on attractiveness, it appears unlikely that participants' behavior on the show reflects an internalization of such concerns.

In view of our findings, it is also relevant that Dutch society seems less concerned with beauty than other similar societies. Spending on cosmetics and toiletries as a proportion of GDP is lower in the Netherlands than most

---

[21] Harrison and List (2004) have an extensive discussion on the external validity of game show data.

other West European countries — the Netherlands is (joint) fourth from the bottom in the list of the EU fifteen.[22] The fact that we find discrimination in such a context seems particularly compelling.

6   Concluding comments

To summarize, our main finding is that beauty is "only skin-deep", and has no implications for a person's performance or their cooperativeness. Nevertheless, it is an attribute well worth having. Attractive players earn a premium, that arises from the reluctance of other players to eliminate them. This seems to reflect consumption value considerations on the part of the other players in the game. The preference for the beautiful is therefore a form of taste based discrimination. Our finding is also noteworthy since participants on a TV show would be reluctant to discriminate, since their behavior is subject to public scrutiny. While discrimination on the basis of gender or race are rightly frowned upon, discrimination based on a person's physical appearance is less remarked upon. Indeed, it is likely that discriminators, the discriminated against and third parties are less aware of the phenomenon, so that it is, to some extent, insidious. Our experimental evidence on third-party perceptions of discrimination is suggestive in this regard. We believe that discrimination on the basis of less obvious criteria such as attractiveness is likely to have qualitatively different characteristics from discrimination based on recognized categories. Without overstating the external validity of our results, we believe that this raises important questions for society and social policy, and merits further research.

---

[22] The role of beauty is less suprising in Argentina, where Mobius and Rosenblat's (2006) experiment was conducted. Buenos Aires is known as the "plastic surgery capital of the world".

References

[1] Altonji, J.G., and R.M. Blank (1999), Race and Gender in the Labor Market, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Amsterdam: North Holland.

[2] Andreoni, J. and R. Petrie (2004), Beauty, Gender and Stereotypes: Evidence from Laboratory Experiments, University of Wisconsin, Madison, Department of Economics Working Paper 2004-06.

[3] Antonovics, K., P. Arcidiacono and R. Walsh (2005), Games and Discrimination: Lessons From The Weakest Link, *Journal of Human Resources* 40(4), pp. 918-947.

[4] Becker, G. (1957), *The Economics of Discrimination*, Chicago University Press, Chicago.

[5] Belot, M., V. Bhaskar and J. van de Ven (2006), Social Preferences in the Public Arena: Evidence from a Prisoner's Dilemma Game on a TV show, http://else.econ.ucl.ac.uk/papers/uploaded/221.pdf.

[6] Belot, M., V. Bhaskar and J. van de Ven (2008), Can Observers Predict Trustworthiness?, http://else.econ.ucl.ac.uk/papers/uploaded/298.pdf.

[7] Benjamin, D.J., and J.M. Shapiro (2006), The Rational Voter, Thinly Sliced: Personal Appeal as an Election Forecaster, mimeo, Harvard University.

[8] Bertrand, M., D. Chugh and S. Mullainathan (2005), Implicit Discrimination, *American Economic Review Papers and Proceedings,* 95(2), 94-98.

[9] Bertrand, M., and S. Mullainathan (2004), Are Emily and Greg more Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination, *American Economic Review* 94(4), 991-1013.

[10] Biddle, J. and D. Hamermesh (1998), Beauty, Productivity, and Discrimination: Lawyers' Looks and Lucre, *Journal of Labor Economics* 16(1), 172-201.

[11] Eckel, C. and R. Wilson (2004), Detecting Trustworthiness: Does Beauty Confound Intuition?, mimeo.

[12] Fershtman, C. and U. Gneezy (2001), Discrimination in a Segmented Society: An Experimental Approach, *Quarterly Journal of Economics* 116, 351-377.

[13] Gneezy, U., M. Niederle and A. Rustichini (2003), Performance in competitive environments: Gender differences, *Quarterly Journal of Economics* 118(3), 1049-1074.

[14] Greenwald., A. G., D. McGhee and J.L.K. Schwartz (1998), Measuring Individual Differences in Implicit Cognition: The Implicit Association Test, *Journal of Personality and Social Psychology*, 74 (6), 1464-80.

[15] Hamermesh, D. and J. Biddle (1994), Beauty and the Labor Market, *American Economic Review* 84(5), 1174-1194.

[16] Harrison, G. and J. List (2004), Field Experiments, *Journal of Economic Literature* 27, 1009-1055.

[17] Heckman, J. (1998), Detecting Discrimination, *Journal of Economic Perspectives,* 12(2), 101-116.

[18] Levitt, S. (2004), Testing Theories of Discrimination: Evidence from Weakest Link, *Journal of Law and Economics* 47, 431-452

[19] Landry, C., A. Lange, J. List, M. Price and N. Rupp (2006), Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment, *Quarterly Journal of Economics* 121(2), 747-782.

[20] List, J. (2006), Friend or Foe? A Natural Experiment of the Prisoner's Dilemma, *NBER Working Paper* 12097.

[21] Mocan, N. and E. Tekin (2006), Ugly Criminals, *NBER Working Paper* 12019.

[22] Mobius, M. and T. Rosenblat (2006), Why Beauty Matters, *American Economic Review* 96(1), 222-235.

[23] Mulford, M., J. Orbell, C. Shatto and J. Stockard (1998), Physical Attractiveness, Opportunity, and Success in Everyday Exchange, *American Journal of Sociology* 103(6), 1565-1593.

[24] Pfann, G. J. Biddle, D. Hamermesh and C. Bosman (2000), Business Success and Businesses' Beauty Capital, *Economics Letters* 67(2), 201-207.

[25] Solnick, S. and M. Schweitzer (1999), The Influence of Physical Attractiveness and Gender on Ultimatum Game Decisions, *Oganizational Behavior and Human Decision Process* 79(3), 199-215.

7   Appendix: Instructions

[Translation from Dutch.] Welcome! The experiment lasts for about 90 minutes. During the experiment you will earn points that are worth money. The exact amount you earn depends on your score and can earn up to about €20. None of the other participants will know what you earn and all your answers will be confidential.

How you earn money You will see fragments of a television game show. You will be asked to predict the choices of contestants. The more accurate your predictions are, the higher your score and the more money you earn. Only your own predictions determine your score and not the predictions made by other participants.

The TV show The game show starts with 5 candidates. In the first round, the candidates have to answer quiz questions. Their score depends on the number of questions they answer correctly. At the end of the round, one candidate is eliminated by the highest-scoring candidate. You will only see the first round, and your main task is to predict who will be eliminated. You will see 7 episodes, based on a random selection from all 69 episodes.

33

For background information, we also explain the rest of the game show. You will not see any of this part of the show and do not have to make any predictions about it. There at two more rounds after the first round, where again the highest scoring person has to eliminate one of the others. After these 3 rounds, there are two candidates left. The scores of both candidates are added. This is the amount of money they will be playing for. Both players simultaneously decide whether to share or grab. There are three possible situations. (1) They both share. In this case, they both get half of the amount of money. (2) One candidate shares and the other does not share. In this case, the one who does not share gets the whole amount. The candidate who shares gets nothing. (3) They both do not share. In this case, nobody wins any money. Before they make their share/grab decisions, candidates have the opportunity to say something to each other.

Instructions We start by showing the candidates introducing themselves. You will see every candidate twice. The first time, there will be no sound and we ask you to predict the age of the candidates. The second time, you will see the candidates introducing themselves again, but with the sound. We ask you to rate the self-confidence of the candidate. You can do this on a scale from 1 (very low self-confidence) to 7 (very high self-confidence), taking 4 as the average of all Dutch people. After the introduction of the candidates, you will see the first round of quiz questions. After the end of the first round, the person with the highest score chooses which other candidate will be eliminated. At that point, we pause the show, and we will ask you to predict how likely it is that a certain candidate will be eliminated.

The way we ask you to do this is as follows. The candidate with the highest score can choose one out of the four candidates. We ask you to divide 100 points between these four candidates, in proportion to the likelihood that they will be eliminated. You can divide the 100 points in any way you like, as long as the total equals 100. You should give more points to a candidate the more likely you think that this candidate will be eliminated. If you give 100 points to one candidate, this means you are certain that this candidate will be eliminated. If you give zero points to a candidate, you are certain that this candidate will not be eliminated. [Several examples were given.]

To summarize, you give more points to a candidate if you think that candidate is more likely to be eliminated. In total, you allocate 100 points over the four candidates. For any candidate, you can use any whole number from 0 to 100 (thus including numbers such as 14, 61, 78 etc.).

Notice that the total number of points should add up to 100. However, you will not be penalized if the numbers you choose do not add up. If you allocate less, say only 90 points, we will reallocate the remaining ten points in the same ratio as your original allocation. If you allocate more points (say 150 points in total), then we will scale back the points proportionally, in this case by multiplying your allocation by 2/3.

After filling in your answer sheet we ask you to put it in the envelope on your table. After you do this, you are not permitted to take it out of the envelope. Hence, you can not go back to an earlier question.

Your earnings At the end of the experiment, we compare your predictions to the actual outcomes. You score is higher if your predictions are

better. The most you can earn per prediction is 4 points and minimum is 0 points (You will see 7 episodes). Every point is worth €0.75. For every age prediction within 2 years of the true age, you receive €0.25. Your confidence ratings have no impact on your score. The amount you earn for the selection predictions is calculated by the formula below. This formula is chosen in such a way that it is in your interest to report your true beliefs. By reporting any other number than what you truly believe, your expected earnings are decreased. A proof of this can be requested at the end of the experiment.

Questions? Please ensure yourself that you have understood the instructions. If you have any questions, please raise your hand and wait until somebody comes to you.

Formula of your score Your score is calculated in such a way that it is best for you to report your best prediction. You do not need to understand how your score is calculated in order to do well in this experiment. However, for those who want to understand the details, we explain this below.

Suppose you allocated p_i points to candidate i, $i = 1, 2, 3, 4$. If candidate i is indeed selected to be eliminated, the score for your prediction is:

$$4 - 2(1 - p_i/100)^2 - 2\sum_{j \neq i}(p_j/100)^2.$$

Your expected score is maximized by reporting what you really believe. A proof of this can be obtained at the end of the experiment.