(Work in progress, please do not circulate)

Only in the Eye of the Beholder? Attractiveness and Trust

Guido Heineck*

University of Bamberg, Germany

Abstract

This paper complements prior experimental and survey based research on beauty by analyzing whether individuals' physical attractiveness is related to their trust propensity. Results from a representative survey from Germany suggest such a relationship mainly for females, but also point to non-linear patterns, with a stronger association for less attractive women, who trust less, even when accounting for anthropometry and personality traits. Further analyses show that less attractive individuals have smaller social networks and that, as a rough approximation of observable behavior, attractiveness slightly improves the ease of persuading individuals to participate in the interview.

JEL classification: Z13

Keywords: Attractiveness; social trust

* Contact: University of Bamberg, Feldkirchenstr. 21, 96052 Bamberg, Germany, Phone: 0049-951-863-2600, guido.heineck@uni-bamberg.de

I am grateful to Helmut Rainer, Florian Herold and seminar participants at the Helmut-Schmidt-University Hamburg for comments on a prior draft. All remaining deficiencies and errors are mine.

All the beauty of the world, 'tis but skin deep.

(Ralph Venning, The Triumph of Assurance, 1647.)

I. Introduction

In whom do we trust? And does physical appearance matter in that? Whereas there is quite some research on similar questions in the psychological literature, in fact showing that we treat others differently depending on whether they are attractive or unattractive (Langlois et al., 2000), there is less evidence on this in economic research and on the follow-up question, what the trust signals we send do to those we interact with. Do good looking individuals internalize their experiences so that they in the end differ in their underlying trust propensity from less attractive ones? The latter, by and large, is what this paper is about.

We all know that trust plays a critical role in interpersonal and social processes. It allows to predict the outcomes of social interactions or at least to form expectations of how others will act or react. To this end, we often use informational cues to reduce uncertainties and to get a more reliable base for our judgement of what might be the outcome. This holds particularly true in situations of repeated interactions, for which reputation then is a major determinant. But everyday social exchange often involves persons and circumstances that are nonrecurring or – in behavorial economics terms – can be described as one-shot game. That is, we interact with others we have never seen before and might possibly never see again so that reputation building is no major concern. In such circumstances, it is both informational cues and our underlying attitude towards trust in others that matter.

It has been previously shown, shortly summarized in the following section, that beauty or physical appearance is one of the informational cues we use in social interactions. Willis and Todorov (2006) even suggest, that is does not take more than 100 ms until we make up our mind for what our perception of others is. There is also research on whether the

differential behavior that attractiveness triggers is worthwhile, so that we could rely on the early work of Dion et al. (1972), who conclude that *"what is beautiful is good"* (p. 285), which is presumably one of the most widely used phrases in this research, or whether we should be careful and rather believe that beauty is only *"skin deep"* (Feingold, 1992).

Apart from the large psychological literature on the physical attractiveness stereotype and its implications, there is also recent economic research from experiments on whether the trust signals pay off for both sides, and also which role individuals' attractiveness plays in that, but the evidence from this research is less clear, pointing to beauty premia (Mobius and Rosenblat, 2006) as well as beauty penalties (Wilson and Eckel, 2006; Andreoni and Petrie, 2008).

This study complements prior research and links the experimental trust research to results based on nationally representative survey data for Germany. To the best of my knowledge, it is also the first study at all on the relationship between attractiveness and trust that is based on large scale survey data and not on small size groups that are typical for particularly lab experiments. This allows gaining insights on whether the evidence on individuals' trust and trustworthiness found in the lab extends to the real world.

The results point to a relationship between attractiveness and trust mainly for females. They also indicate non-linear patterns, inasmuch as the associations found are stronger for unattractive women. This holds accounting for a wide range of characteristics, including individuals' anthropometry and personality. As a proxy for whether attractive individuals have better social skills, further analyses show that less attractive individuals have smaller networks and that attractive males are more likely to establish useful contacts in their leisure time, documenting a 'strategic' interest. Finally, as a rough approximation of observable outcomes, additional analyses show that attractiveness slightly adds to the ease of persuading individuals to participate in the interview.

II. Background and Previous Findings

Beauty matters in life and it seems pointless to deny this. It is in fact well documented across different disciplines that looks count for a range of different outcomes. There is particularly much research in psychology, where the interest in individuals' attractiveness has a long history (for an overview, see Swami and Furnham, 2008), covering the determinants and the perception of beauty as well as its behavioral correlates.

Within economics, beauty is addressed in survey based research as well as in experimental studies. Evidence from these two strands of research however produce a somewhat contrasting picture: Evidence from experiments points to "beauty premia" (Mobius and Rosenblat, 2006; Solnick and Schweitzer, 1999) but also to "beauty penalties" if attractive individuals do not live up to the expectations others have on their performance or cooperation (Wilson and Eckel, 2006; Andreoni and Petrie, 2008).

In particular, Mobius and Rosenblat (2006) set up an experimental labor market in order to decompose the beauty wage premium as found by many studies, for example Hamermesh and Biddle (1994). Running on a sample of 165 students from three Argentinian universities they identify three transmission channels which may result in wage premia: higher confidence of attractive workers, a biased perception of the ability of attractive workers by employers, conditional on confidence, and, last, better communication and social skills of attractive workers.

More related to the question addressed here are however the studies by Wilson and Eckel (2006) and Andreoni and Petrie (2008). The latter also start with "labor markets as inspiration" (Andreoni and Petrie, 2008, p. 74), but examine group cooperation issues with the help of a repeated linear public goods game. Based on a sample of 80 students from the University of Winsconsin-Madison they first find a beauty premium, and although attractive

subjects on average do not contribute more than others. However, if information on individuals' contributions is given, the beauty premium turns into a penalty. Andreoni and Petrie (2008) argue that this might be because the other players expect attractive persons to be more cooperative, so that unmet, yet biased expectations lead to punishment of attractive players.

On the same lines, Wilson and Eckel (2006) run a series of trust game experiments involving a trust game, an assessment of risk attitudes, and responses to a survey of attitudes toward trustworthiness and altruism with a total of 206 students from three US universities. Similar to Andreoni and Petrie (2008), the results from their experiments suggest that attractive players are thought to be more trustworthy and, because of that, earn a beauty premium in the first stage. But again, as attractive individuals do not live up to the biased expectations of the trusters, more attractive players are punished in the second stage.

Whereas this evidence calls into question that attractive persons will on average be more successful, this essentially says more about the other players and their biased, i.e. overly positive expectations in the first place.

On the other hand, Solnick and Schweitzer (1999) found in one-shot ultimatum games that attractive persons were treated differently and that they got higher offers, which is consistent with the idea of a "beauty premium". It should however also be noted that while they were offered more, more was also demanded of them. Yet, since the experiments were set up as one-shot game, it is hard to say whether the attractive players would be punished when, again, not living up the expectations of the other players.

Plus, the evidence that points to punishment or penalties is at odds with findings from survey-based research that consistently indicate a range of benefits for the attractive. These benefits include, as already mentioned, wage premia (e.g. Hamermesh and Biddle, 1994, or Scholz and Sicinski, 2015), higher employment participation (Gehrsitz, 2014), sooner and

more responses to applications (Bóo et al., 2013), higher happiness (Diener et al., 1995; Hamermesh and Abrevaya, 2013), or, for politicians, electoral success (King and Leigh, 2009; Berggren et al., 2010). In addition, there is also evidence that less attractive individuals behave differently. Mocan and Tekin (2010), for example, show that unattractive individuals are more prone to criminal behaviour, possibly a reaction to the labor market disadvantages that are induced by their looks. Without going into further details on the mentioned studies, it should suffice to say that this evidence stretches across different countries, different times, and different research questions so that it in sum is hard to believe that the beauty penalties found in experiments endure in the outside world in the long run.

But why is it that more attractive persons seem to be better off, on average? Is it possibly really the case that the good-looking are also *"good"*, as implied by Dion et al. (1972), which would, for example, suggest that they might invest more in human capital which would then pay off. Or are we simply fooled by the looks of others so that there would be indeed no substantial reason for the success of more attractive ones? There have been manifold attempts and theories to answer these questions and to disentangle the underlying causal mechanisms, mostly developed in the psychological literature.

Most of the concepts that attempt to explain why we might treat attractive others differently are in some way or another related to the idea of *"social stereotyping"* (Dion et al., 1972; Eagly et al., 1991), where stereotype can be defined as a fixed, over-generalized belief about a particular group or class of people (Cardwell, 1996). In the context here, social stereotyping is also known as the *"attractiveness halo effect"* (for the halo effect in general, see Nisbett and DeCamp Wilson, 1977). This refers to the phenomenon that we attribute positive characteristics such as friendliness or trustworthiness to attractive persons more than to average or below average looking others.[1]

At a preliminary stage or even contributing to stereotyping is another cognitive bias effect: attractiveness simply catches our eye. Maner et al. (2003) suggest that we selectively process others according to their attractiveness and that we tend to focus our attention to physically attractive targets, likely because of different mating-related motives.

Irrespective of what attracts our attention to good-looking persons in the first place, social stereotyping may spur a multistep causal mechanism that, as outlined by Darley and Fazio (1980) and repeated by Langlois et al. (2000), not only leads to differential behaviors of ourselves, but eventually also of the individuals we believe to be attractive: Following Langlois et al. (2000), think of 'targets' as the individuals who are judged on attractiveness and as 'perceivers' those persons who interact with targets.

The mechanism then suggests that, first, the physical appearance triggers stereotypes, i.e. expectations for specific behavior in the perceivers. These expectations will, next, lead to differential judgements and also treatment of targets, depending on whether they are attractive or unattractive. The next step proposes that the differential judgements and treatments induce a development that results in differential traits and behavior of targets. The final step suggests that, at the end of this process, targets will internalize both differential judgements and treatments and develop differential behavior, traits, and self-views.

Applying this mechanism here results in a testable hypothesis: If we, as an outcome of the halo effect, believe that attractive individuals are more trustworthy, we treat them according to this expectation. Imagine now that if attractive individuals receive these trust signals from 'perceivers' in numerous "one-shot" situations and across different circumstances, they will eventually incorporate that they can, in general, trust other people, as this is the pattern they learn from these interactions.

To structure the mechanism in a more formal way, one could think of it as follows. Let

$$T_i = f\left(x_i, \eta_i, \sum S_{ij}\right) , \tag{1}$$

where T_i is the individual's latent propensity to trust in others, which for ease of interpretation can be assumed to be centered at 0, so that $T_i > 0$ $\left(T_i < 0\right)$ implies above (below) average trust. x_i is a vector of trust-determining characteristics, such as age, gender, or education, and η_i is capturing unobservable components affecting *i*'s trust propensity. S_{ij} is the parameter through which other individuals' cognitive bias might add to T_i. It represents the signals, individual *i* receives from individual(s) *j* according to whether *i* is perceived as being trustworthy and the differential behavior that will follow from that. It enters the trust function as an aggregate, as individual *i* will learn from the signals he or she has received from multiple social interactions with *j*.

S_{ij} can be thought of as a function that is determined by, again, x_i, but now also by x_j, the sender's characteristics, unobserved components υ_i, and B_i, the attractiveness of individual *i:*

$$S_{ij} = g\left(x_i, x_j, \upsilon_i, B_i\right) \tag{2}$$

For exposition, it is helpful to think also of B_i in categories centered around zero, so that

$$B_i = \begin{cases} > 0 \\ = 0 \\ < 0 \end{cases} \text{ if individual } i\text{'s attractiveness } \begin{cases} \text{is above average} \\ \text{is average} \\ \text{is below average} \end{cases} \tag{3}$$

The signals that *j* sends will depend on *i*'s attractiveness, but the direction is not clear cut, at least for above attractive individuals. This is, because it is as plausible to argue in favor of positive signals as it is for negative ones. Whereas positive signals might be given if we as perceivers are subject to the cognitive bias of believing that beautiful persons are also good persons. Yet, it is on the other hand as plausible to imagine that perceivers have reason to not believe into the "beautiful equals good" notion. This might be the case if, for example, in

prior encounters the perceivers' expectations with either the same individual or other attractive individuals have not been met which might trigger an on average cautious behavior when being confronted with attractive targets.

In sum, if $B_i > 0$, i.e. if individual *i* is attractive, both positive and negative signals might be sent: $\partial S_{ij} > 0$, if, say, the halo effect dominates, but $\partial S_{ij} < 0$ if individual *j* has e.g. made overly adverse experiences with attractive individuals so that he or she is reacting to attractive others with caution, if not mistrust.

For less attractive individuals, i.e. for $B_i < 0$, it can on the other hand be argued that the "dark side" of our 'tastes for beauty' might leave an unambiguous impact on individual *i*: Given all the evidence on the disadvantages as outlined above, it is more plausible to argue that the signals sent will by and large be negative: $\partial S_{ij} < 0$.

Individual *i*'s trust propensity will at the end of an internalizing process embrace whether they receive more positive or negative signals, conditional on their attractiveness. That is:

$$\frac{\partial T_i}{\partial S_{ij}(B_i)} = \begin{cases} > 0 \\ < 0 \end{cases}, \tag{4}$$

if individual *i* is attractive, so this is an open empirical question, but

$$\frac{\partial T_i}{\partial S_{ij}(B_i)} < 0, \tag{5}$$

if individual *i* is unattractive.

This concept does not necessarily contradict the results from the above mentioned experiments that attractive individuals suffer a "beauty penalty" if they do not live up the expectations of 'perceivers' as found by Wilson and Eckel (2006) or Andreoni and Petrie (2008) in repeated game situations. However, again, everyday social exchange is often characterized by one-shot moments that nurture the outlined mechanism which, in terms of

differential treatment, starts off already in the targets' childhood (Langlois et al., 2000) and might essentially lead to a higher (lower) trust propensity of more (less) attractive individuals.

III. Data and Methods

The data are derived from the German General Social Survey (GGSS or ALLBUS).[2] The GGSS is a cross-sectional, representative survey that is conducted biennially since 1980, and since 1986 also comprises the German part of the International Social Survey Programme (ISSP). While the data come with all the potential shortcomings of cross-sectional data sources, it offers a rich and wide range of individuals' attitudes, behavior, and social structure, including typical relevant socio-demographic information on education, employment, income, etc..

The major part of the analyses below is based on the 2008 wave of the GGSS because it not only provides indicators on individuals' attractiveness, but also a few survey items on individuals' propensity to trust others. In contrast to other studies which often have to rely on the "generalized trust" question only (*"most people can be trusted"*, for details see below), this allows to retrieve an underlying latent trust factor that is more robust with regard to measurement error.

Attractiveness measures

In 2008, individuals' attractiveness was assessed three times, twice by the interviewer and once by the respondent him- or herself. The interviewers had to rate the respondent's overall attractiveness – so not their facial features only – at the beginning and at the end of the interview. For the first rating, it is important to note that interviewers were instructed to make this assessment on their first impression of the respondent, and before the interview started. In particular, the questionnaire instruction reads: *"How attractive or unattractive is the*

respondent? Please decide spontaneously according to your first impression." Responses were to be given on a 11-point Likert-type-scale, from 1 *"Unattractive"* to 11 *"Attractive"*.

Table 1 outlines the distribution of individuals' attractiveness by gender of the respondent and the interviewer, age, and by the rating method, i.e. whether it was assessed by the interviewer at either the beginning or the end of the interview or by the respondent him- or herself.[3]

Table 1 about here

The patterns in the first panel show that the interviewers give more flattering ratings at the end of the interview, with a correlation coefficient of 0.84, and shifting the whole distribution slightly upwards. This may be the result of different underlying mechanisms of which the *halo effect*, as outlined above, is critical. It implies that a smooth and, of interest here, trusting interaction during the interview would re-inforce the first impression, resulting in a more favorable assessment at the end. For an 'unbiased' association between interviewers' assessment of individuals' attractiveness and their trust propensity it is therefore most important to rely on the interviewers' first ratings. To rule out further interviewer-specific effects, the attractiveness ratings are standardized at the interviewer level so that the coefficients in the empirical analyses below can be interpreted by standard deviation changes in individuals' attractiveness, net of the interviewers' latent and unobservable rating scheme.

Also towards the end of the interview, respondents were asked to rate their own looks. Compared to the interviewers, respondents are much harder on themselves. The correlations between this self-assessment and the interviewers' ratings are remarkably lower, with 0.39 for the first rating and 0.51 for the second. Yet, as the underlying assumption suggests for an effect of impressions of others and how they react to individuals' beauty, this is not of further concern.

Table 1, panel II, also shows that, in their first ratings, female interviewers are somewhat more generous, resulting in more respondents having above average looks as compared to the assessments of their male interviewer counterparts. The pattern repeats inasmuch as males are rated less attractive, as shown by Table I, panel III and the histograms in Figure 1. On average, females are more favorably rated than males, with means of 7.3 for males and 7.5 for females (statistically different at the 1%-level, with a *p*-value of 0.0067).

Figure 1 about here

Panel IV of Table 1 finally shows that attractiveness is the realm of the young: Whereas about a quarter of the respondents of age 40 or older are deemed to have above average looks, this share is 1.7 times higher in the population that is younger than 40. Even more obvious, the share of the plain looking persons is almost three times higher in the elderly part of the sample.

Trust measures

Using the 2008 GGSS wave allows to retrieve individuals' latent trust propensity from three survey items. Two of these are variants of the well-established "generalized trust" question which is included in quite many surveys, such as the US General Social Survey, the European or the World Values Surveys and many more.[4] Whereas the original item, that asks for whether most people can be trusted or whether one cannot be careful enough, allows only for either yes or no as a response, the GGSS items try to accommodate the potential ambiguity of the item.[5]

In particular, the first item reads: *"Some people think that most people can be trusted. Others think that one can't be careful enough when dealing with other people. What do you think?"*, for which respondents' responses are categorized as follows: *"Most people can be trusted"*, *"One can't be careful enough"*, and *"It depends"* and thus go beyond the original binary response options.

The second variant asks *"Generally speaking, would you say that people can be trusted or that you can't be too careful in dealing with people?"*. The responses to this time allow for four, rather than three options: *"People can almost always be trusted"*; *"People can usually be trusted"*; *"You usually can't be too careful in dealing with people"*; *"You almost always can't be too careful in dealing with people."*.

The third item then specifically directs respondents to people they do not know, i.e. asking them *"In general, how much trust do you have in strangers?"*. The reply to this item is also more fine-grained than the two outlined variants of the generalized trust item, inasmuch as respondents can pick from a scale that runs from *"1 = Absolutely no trust at all"* to *"7 = A great deal of trust"*.

As mentioned, the original "generalized trust" item is ambiguous and it is therefore not surprising that responses to this question do not perform well as predictors in trust experiments (see Ermisch et al., 2009 or Naef or Schupp, 2009). It is instead the "trust into stranger" instrument or variants thereof that seem to capture individuals' trustworthiness much better (Naef and Schupp, 2009). It would therefore be rather careless to rely on only one of these trust items, and even worse if this would be the one of the generalized trust measures.

With no experimental data at hand, the approach chosen here therefore is to retrieve an underlying latent factor that will better reflect individuals' trust propensity. As shown in Table A1, and given that there are only three items, they load satisfactorily on the common factor. The interrelatedness of the items is further confirmed by a value of Cronbach's alpha of 0.74. The following multivariate analyses therefore include a scale that is constructed by using standardized values of the three trust items.[6]

Giving a first impression of a possible relation, Figure 2 then suggests a moderate positive association between the standardized measures of individuals' attractiveness and their trust propensity, for both males and females.

Figure 2 about here

The following regressions account for a wide range of socio-demographic controls, including e.g. gender, age, (respondent's and parental) education, self-rated health, marital status, household size and income, and occupation for the employed. To capture potential interviewer effects other than his or her unobservable propensity to assess individuals' attractiveness, which is accounted for by using the standardized beauty ratings, the analyses also control for interviewers' age, whether the interviewer is 10 years older or younger than the interviewee, whether the interviewer is of opposite sex, the interviewers' education, and their job experience. In order to exclude an impact of respondents' attractiveness-related overconfidence or pessimism, the regressions further include two dummies on differences in respondents' self-rating and the interviewers' initial rating that are larger than |1|.[7] Descriptive statistics of the final sample are given in the Appendix, Table A2.

IV. Results

As outlined above, it is more suitable to rely on the interviewers' rating at the beginning of the interview as this first impression rules out any influences that might emerge over the course of the interview. Yet, for completeness, the results from the first set of regressions shown in Table 2 also provide evidence from otherwise identical specifications that substitute the standardized attractiveness measure with the 11-point scales at that beginning and at the end of the interview as well as the respondent's self-rated attractiveness. Except for the latter, the findings for the other three models are roughly the same: They indicate that a one standard deviation change in individuals' attractiveness is associated with an increase in

individuals' trust by about 0.03 standard deviation points, statistically significant at the 5%-level or better.

Table 2 about here

In terms of magnitude, this is not an overly large association, but it is roughly twice as large as the coefficient of the linear age parameter. There anyway is only a linear relation between age and trust, as the coefficient for the quadratic of age is not statistically different from zero. The results for the other control variables – which are shortly outlined here, but omitted in the following analyses[8] – indicate that females are much less likely to trust than males.[9] Education, on the other hand, is positively related to trust. In fact, the coefficients imply that education is the most important component of trust. Respondents, in particular, who have completed upper secondary school, show a trust level that is some 0.4 standard deviation changes higher than their counterparts from the reference category, i.e. individuals who have completed only lower secondary school. Whereas being married is not related to trust, living in East Germany indicates less trust. The magnitude of 0.2 implies that living in East Germany may on average offset half of the effect of upper secondary school. This is remarkably large, but is in general in line with the findings of Heineck and Süssmuth (2013) who found that East Germans, despite twenty years of unification, are still characterized by a persistent level of social distrust.

Household income is moderately associated to trust, i.e. its coefficient indicates that a one percent increase in income relates to a 0.05 standard deviation point change in trust. In contrast to this, neither current (full- or part-time) employment nor having been unemployed in the past is statistically significant. Whereas migration background is also not associated to trust, not having adapted the German citizenship is, and its coefficient, 0.14, implies a differential in trust comparable to that found for females. The coefficient for self-rated health, which is a 5-point scale ranging from 1 "bad" to 5 "very good" points to a positive link

between individuals' attractiveness and their trust. It furthermore shows that education and its effect on trust might even be transmitted over generations: Having a father who completed upper secondary school – and who might therefore possibly be a role model for trusting behavior or trust in general – is associated to an increase in the respondent's trust of about 0.1 standard deviation points. Mothers' education, however, does not play a role in that.

Again, as outlined above, the attractiveness measure is standardized at the interviewer level to account for differences in their beauty perception. To capture further heterogeneity, the regressions include interviewers' gender, age, education, and their job experience. Of these, only age and job experience are statistically related to individuals' latent trust, and the coefficients indicate a rather small, but negative relation to the respondents' trust.[10]

Attractiveness and trust, by gender, and non-linear patterns

As attractiveness stereotypes in their consequences might be more relevant for females, Table 3 presents results from regressions for the whole sample, i.e. simply repeating the findings already given in Table 2, panel I, column 1, but also separately by gender.

Table 3 about here

Compared to the whole sample and the linear association shown there, the results now point to different associations between attractiveness and trust by gender. For females, the linear link is again visible and even stronger, with a coefficient of about 0.07, i.e. of almost double size. For males, however, there is no evidence for a linear relationship. One might argue that this has to do with the smaller sample size, but additional full sample regressions (not shown) that contain an interaction between attractiveness and gender also point towards the finding, that attractiveness has a stronger impact on the trust propensity of females.

It might in addition be that the relationship between individuals' attractiveness and their trust is not following a simple linear gradient. To account for this, Table 3, panel II, further provides results from regressions that each include two dummy variables indicating

whether the respondent's attractiveness rating is at either the lower or upper tail of the rating distribution: One set of specifications captures potential non-linearity by dummies for ratings that are below (above) the mean attractiveness minus (plus) one standard deviation, whereas the second set includes dummies that refer to the bottom or top 25% of the attractiveness distribution.

The results again indicate that beauty is mostly related to females' trust. Whereas the full sample results still indicate that respondents, whose attractiveness ratings are in the bottom 25% of the distribution, might be less prone to trust (Table 3, panel II, column 2), the link holds, again, essentially only for females if broken down by gender. Depending on the partition of the distribution, the results suggest for differentials in females' latent trust of about 0.12 or 0.16 standard deviation points (Table 3, panel II, columns 5 and 6).

Attractiveness and trust, accounting for anthropometrics

Further on, Oreffice and Quintana-Domeque (2014), who also use the GGSS 2008 data, document that respondents' anthropometry plays a role for the attractiveness assessment made by opposite-sex interviewers. They in addition find that thinner females are rated more favourably by female interviewers whereas height, weight and body-mass-index (BMI) do not contribute to male interviewers' assessment of males' attractiveness. Picking up on this evidence, additional estimations that use the ISSP-subsample of the GGSS[11] therefore control for individuals' body-mass-index and their height.[12] First, descriptive patterns (Table 4) are in line with Oreffice and Quintana-Domeque (2014) inasmuch as there is an overall inverse relation between individuals' BMI and their attractiveness ratings which is particularly pronounced for females, and ignoring the result for the three underweight males in the sample who are less favourably rated than their normal weight counterparts.

Table 4 about here

The patterns shown in Table 4 also indicate that there may be a link between BMI and trust as there is a by and large inverse relation between BMI and trust for males, and an inverse u-shaped profile for females, with the highest trust level found for females with normal weight, i.e. a BMI-value of 18.5-25 and much less trust for underweight and obese females.

Yet, accounting for individuals' BMI and height in the trust regressions by including interaction terms between BMI and attractiveness seems not to contribute any further to understanding individuals' trust propensity. This is not too surprising given that BMI is a relevant factor for the interviewers' overall attractiveness rating. Further tests for joint significance of the attractiveness measures and its interactions do not change this result.

Table 5 about here

No evidence for such an additional impact of BMI and height is also found when running regressions that include terms which interact the standardized attractiveness measure with the respondent's BMI value.[13]

Attractiveness and trust, accounting for personality traits

If the above outlined mechanism applies, it can also be expected that more attractive persons differ from attractive less ones also in terms of personality traits. To account for this, further regressions include information on individuals' personality as measured by the Five Factor Model (FFM) (McCrae and John, 1992).[14] The results are in line with the findings shown in Table 3 inasmuch as there is evidence for a relation between attractiveness and trust only for females. The specification using the linear attractiveness measure indicates a differential in females' trust propensity by about 0.05 standard deviation points (Table 6, column 5), which is only slightly lower compared to 0.07 standard deviation points found above. The coefficients for beauty ratings either in the bottom or top 25% of the distribution are also not too different from those shown above, again indicating that there is less trust among females

in the lower part of the distribution, with a difference of -0.13 standard deviation points (Table 6, column 6).

Table 6 about here

The results for the FFM traits themselves also point to a statistically significant relationship with individuals' trust. While this is an interesting research question of its own (but one which goes beyond the focus of this paper), the results imply that more extraverted and agreeable persons, both males and females, are also more trusting. Agreeableness in fact is the most relevant of all FFM traits, indicating that individuals' trust propensity is some 0.2 standard deviation points higher for a one standard deviation increase in agreeableness. Conscientiousness, on the other hand, is negatively associated to trust. Gender differences are found for openness and neuroticism: Whereas these two personality facets do not explain males' trust, open females are more trusting than neurotic ones, who are less trusting, with coefficient magnitudes of about the same, yet moderate size.[15]

Extensions: Generalized trust in an extended sample (2008-14)

All previous results were based on individuals' latent trust propensity as retrieved from the three GGSS trust items provided by the 2008 wave, as outlined above. The following waves of the GGSS in 2010, 2012, and 2014 however repeat only one of these instruments, namely the "most people can be trusted" item, i.e. one of the two variants of the "generalized trust" measures. As an extension, and if only to be able to use a larger, pooled sample, Table 8 provides results from ordered probit regressions, in particular average marginal effects for the predicted probability of individuals' response that "most people can be trusted".

Table 7 about here

First, and for comparison with the previous findings, results for the 2008 wave only indicate again that attractiveness plays a role for mainly females' trust. The average marginal effect points to an almost two percentage point increase in the predicted trust probability. Yet, the

results for the larger sample show that, as could be expected, sample size matters: In the roughly four-times larger sample attractiveness is again statistically related to females' trust, but now also relevant for males'. And compared to the cross-sectional results, the average marginal effect only slightly decreases, suggesting a one percent increase in the predicted probability that "most people can be trusted".

Extensions: Leisure time preferences and networks

Yet another extension of the previous analyses examines whether attractive individuals have different time use preferences, particularly with regard to gregariousness as this may possibly be one mechanism that shapes their social skills and eventually yields the "beauty premia" as outlined in the beginning of the paper. To this end, survey items from the "Leisure and Sports" ISSP-subsample of the 2008 GGSS wave are used that attempt to uncover individuals' preferences on being with other people in general, being with friends, doing sports to meet other people or, with a somewhat different angle, using their free time to establish useful contacts.

Table 8 about here

The results for being with other people or friends (Table 8, panels I-II) are mostly trivial. It is only for less attractive females that the estimates indicate some lower preference for being with others (panel I, column 6), and similarly for less attractive respondents who are less likely to be with friends in their free time (panel II, column 2). Yet, both coefficients are statistically significant only at the 10% level. The results in panel III then indicate that attractive individuals are more likely to do sports in order to meet other people (columns 1 and 2). While this finding holds for both the linear and the non-linear specifications, the association for the latter case is only weakly statistical significant. In the separate analyses, the picture is even worse in terms of statistical significance, which however is quite likely driven by sample size.

Interestingly, the estimates furthermore point to a possibly strategic behavior of more attractive males: The results in Table 8, panel IV show increases in males' propensity to use their free time to establish useful contacts by on average 0.05 percentage points with a one standard deviation increase in the handsomeness ratings. Looking at the non-linear patterns, this result is confirmed with a 0.07 percentage point differential for males in the 25% bottom part of the attractiveness distribution and a 0.09 percentage point increase for their counterparts in the 25% top part of it.[16]

"Revealed" trust: Network size and willingness to participate in the interview

Another sociability aspect – and possibly one, one could think of as revealed trusting behavior – is addressed in Tables 9 and 10 that provide descriptive patterns and regression results for the number of individuals in respondents' networks. The data for this additional analysis are drawn from the 2010 GGSS. The descriptive statistics first indicate that the size of individuals' social networks seems not large, ranging between 2.1 and 2.9 persons outside of one's own household. This is however due to the way, the networks were surveyed.[17] Irrespective of the network size itself, there are monotonic links between attractiveness and the size of the individuals' network, with a somewhat larger gap between persons with below average looks and those with average looks, as compared to the differences in network size between average and above average looking respondents. Yet, this might be possibly confounded with age, so that multiple regressions are required.

Tables 9 and 10 about here

Accounting for the whole range of individuals' and interviewers' characteristics, the results in Table 10 then show, that the respondents' network size increases by about 0.11 to 0.13 persons with a standard deviation change in beauty. This represents a five percent increase, given the average network size of 2.5. The relation gets even stronger when looking at possible non-linear patterns. In line with the findings on self-reported trust above, there are

again, differentials for respondents with below average looks, but this time for both females and males. The coefficients translate into network sizes that are eight percent smaller for females and 12 percent smaller for males.

As yet another "revealed" trust indicator, some final analyses examine the willingness of the respondent to take part in the interview. This indicator is not part of the questionnaire itself, but information given by the interviewer. In particular, the interviewer should indicate how difficult it was to persuade the respondent to take part in the interview, with the response options *"very difficult"*, *"tended to be difficult"*, *"tended to be easy"*, and *"very easy"*. It is important to note that this does not refer to the cooperation over the course the interview, which could be influenced by the interaction between interviewer and respondent. Again, one might to some extent think of this as an observed outcome of the respondents' trust propensity. Another advantage is, that this information is given in all of the GGSS waves of interest here, so that similar to the "generalized trust" analysis above (Table 7), a substantially larger sample can be used for the regressions.

Table 11 about here

The results from the additional analyses then show that, conditional on all other explaining covariates, there is no non-linear relation between individuals' attractiveness and their willingness to be interviewed. However, the linear attractiveness measures indicate that a one standard deviation increase in attractiveness is associated with less difficulty and more ease, although this pattern is not the most important determinant in terms of coefficient size.[18]

V. Concluding Remarks

Beauty as a cultural concept is all around us and many devote lots of monetary and non-monetary resources in their looks. Research documents that this might also pay off in terms of higher earnings (Hamermesh and Biddle, 1994; Scholtz and Sicinski, 2015) or higher

happiness (Hamermesh and Abrevaya, 2013). There are several possible mechanisms that might lead to these relative benefits. Discrimination either against unattractive or in favor of attractive individuals might for example add to these outcomes. There is, furthermore, a discussion in psychological research on whether attractive individuals differ in behavior and traits from others with average or below average looks (Langlois et al., 2000).

Related to this, this paper analyzes the link between attractiveness individuals' trust propensity. On the hand, experimental research suggests that attractive persons might suffer a "beauty penalty" if they do not live up to the expectations of others (Wilson and Eckel, 2006; Andreoni and Petrie, 2008). Yet, it is as plausible to argue that individuals learn from the accumulation of signals they receive from others in one-shot social interactions. Depending on the individuals' attractiveness, the signals might be positive or negative so that they eventually internalize that trust in others is justified or not.

The results here support these notions. Based on a representative survey from Germany, individuals' attractiveness, as approximated by interviewers' first impressions, is associated to individuals' trust propensity, as derived from three survey items on generalized trust and trust into strangers. The relationship is stronger for females and particularly so for less attractive ones, who have a lower trust propensity. This holds for a range of checks and also accounting for individuals' anthropometry and personality traits.

What exactly is at the heart of these findings and whether they are driven by individuals' experiences is however difficult to say. Yet, less attractive individuals tend to have smaller social networks, and attractiveness is also related to the difficulty or ease of persuading respondents to participate in the interview. The latter – observed – outcome indicates that the relation between attractiveness and trust propensity translates into behavior in social interactions.

There are no straightforward policy implications from this study and deriving any was not the aim of the paper in the first place anyway. However, the findings presented are in line with the multistep causal mechanism as outlined above. It should therefore be allowed to say that we should – in everyday social interactions – try to be more aware for the cognitive bias of our own behavior when responding to the attractiveness of others as this might in the end change their potential of trusting into others.

References

Alesina, Alberto, and Eliana La Ferrara, E. (2002). Who trusts others? Journal of Public Economics 85(2), 207-234.

Andreoni, James and Ragan Petrie. 2008. Beauty, gender and stereotypes: Evidence from laboratory experiments. Journal of Economic Psychology 29, 73-93.

Arrow, K.J. (1973). The Theory of Discrimination. In A. Ashenfelter and A. Rees (eds.), Discrimination in Labor Markets, Princeton, NJ: Princeton University Press.

Berggren, N., Jordahl, H., and P. Poutvaara (2010). The looks of a winner: Beauty and electoral success. Journal of Public Economics, 94(1), 8-15.

Bjørnskov, C. (2007). Determinants of generalized trust: A cross-country comparison. Public Choice, 130(1-2), 1-21.

Bóo, F. L., Rossi, M. A., & Urzua, S. S. (2013). The labor market return to an attractive face: Evidence from a field experiment. Economics Letters, 118(1), 170-172.

Cardwell, M. (1996). Dictionary of Psychology. Chicago IL: Fitzroy Dearborn.

Darley, J.M., and R.H. Fazio (1980). Expectancy-confirmation processes arising in the social interaction sequence. American Psychologist 35, 867-881.

Diener, Ed, Wolsic, Brian, & Fujita, Frank. (1995). Physical attractiveness and subjective well-being. Journal of Personality and Social Psychology, 69(1), 120.

Dion, K., Berscheid, E., and E. Walster (1972). What is beautiful is good. Journal of Personality and Social Psychology, 24, 285-290.

Ermisch, J., Gambetta, D., Laurie, H., Siedler, T., and S.C.N. Uhrig (2009). Measuring People's Trust. Journal of the Royal Statistical Society, Series A 172(4), 749-769.

Feingold, A. (1992). Good-looking people are not what we think. Psychological Bulletin, 111(2), 304-341.

Gehrsitz, M. (2014). Looks and labor: Do attractive people work more? Labour: Review of Labor Economics and Industrial Relations 28(3), 269-287.

Hamermesh, D., and J. Abrevaya (2013). Beauty is the Promise of Happiness? European Economic Review 64, 351-368.

Hamermesh, D., and J. Biddle (1994). Beauty and the labor market. American Economic Review 84(5), 1174-1194.

Heineck, G., and B. Süssmuth (2013). A different look at Lenin's legacy: Social capital and risk taking in the two Germanies. Journal of Comparative Economics 41(3), 789-803.

King, A., and A. Leigh, (2009). Beautiful Politicians. Kyklos 62(4), 579-593.

Koch, Achim, & Wasmer, Martina (2004) Die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS). Rückblick und Ausblick in die neunziger Jahre. ZUMA-Nachrichten 29, 7-28.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. Psychological Bulletin 126, 390-423.

Maner, J. K., Kenrick, D. T., Becker, D. V., Delton, A. W., Hofer, B., Wilbur, C. J., & Neuberg, S. L. (2003). Sexually selective cognition: beauty captures the mind of the beholder. Journal of Personality and Social Psychology 85(6), 1107.

McCrae, R.R., and John, O.P. (1992). An introduction to the five-factor model and its applications. Journal of Personality 60(2), 175-215.

Mobius, Markus M., & Rosenblat, Tanya S. (2006). Why beauty matters. American Economic Review 96(1), 222-235.

Mocan, N., and E. Tekin (2010). Ugly criminals. The Review of Economics and Statistics 92(1), 15-30.

Mulford, M., Orbell, J., Shatto, C., & Stockard, J. (1998). Physical attractiveness, opportunity, and success in everyday exchange. American Journal of Sociology 103, 1565-1592.

Naef, M., and J. Schupp (2009). Measuring trust: Experiments and surveys in contrast and combination. SOEP Papers No. 167.

Nisbett, R.E., and T. DeCamp Wilson (1977). The Halo-Effect: Evidence for Unconscious Alteration of Judgements. Journal of Personality and Social Psychology 35(4), 250-256.

Oreffice, S., and Quintana-Domeque, C. (2014). Attractiveness, Anthropometry or Both? Their Relationship and Role in Economic Research. IZA Discussion Paper No. 8527.

Phelps, E.S. (1972). The Statistical Theory of Racism and Sexism. American Economic Review 62: 659-661.

Rainer, H., and T. Siedler (2009). Does democracy foster trust? Journal of Comparative Economics 37(2), 251-269.

Rammstedt, B., and O.P. John (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality 41(1), 203-212.

Scholz, J. K., and K. Sicinski (2015). Facial attractiveness and lifetime earnings: Evidence from a cohort study. Review of Economics and Statistics 97(1), 14-28.

Solnick, S.J., and M.E. Schweitzer (1999). The influence of physical attractiveness and gender on ultimatum game decisions. Organizational Behavior and Human Decisions Processes 79(3), 199-215.

Swami, V., and A. Furnham (2008). The psychology of physical attraction. Routledge: New York.

Terwey, M. (2000). ALLBUS: A German General Social Survey. Journal of Applied Social Science Studies 120, 151-158.

Willis, J., and A. Todorov (2006). First Impressions. Making Up Your Mind After a 100-Ms Exposure to a Face. Psychological Science 17(7), 592-598.

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, 189-202.

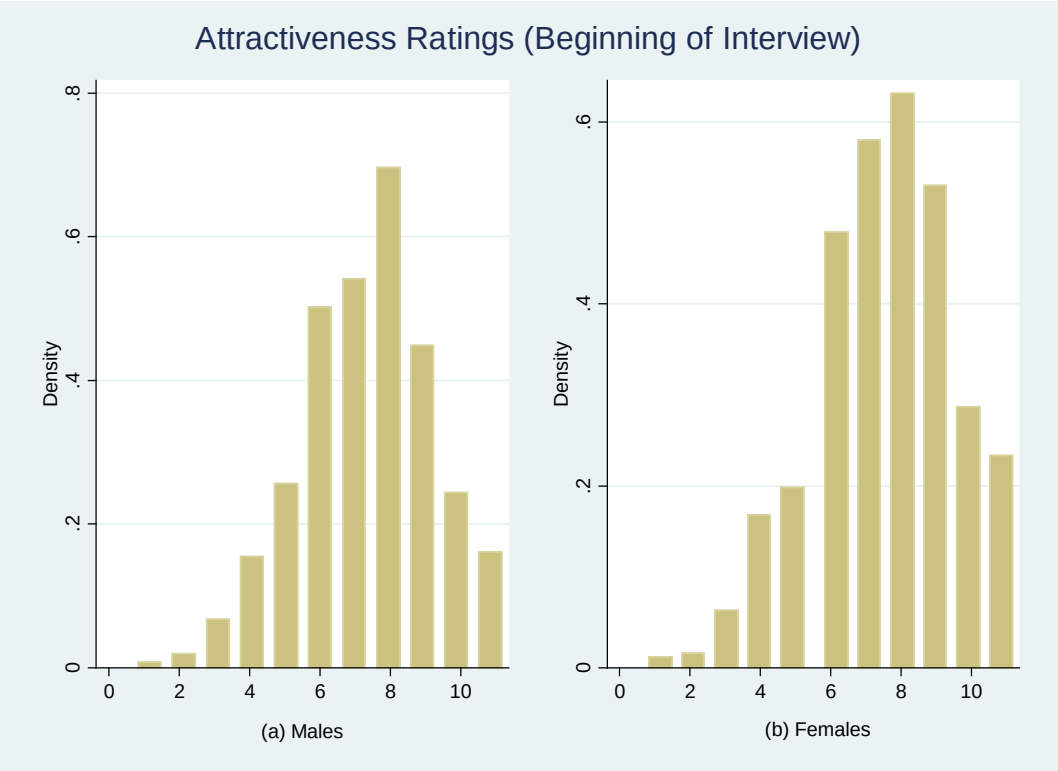

Figure 1: Attractiveness Ratings at the Beginning of the Interview, by Gender

Source: GGSS/ALLBUS, 2008.

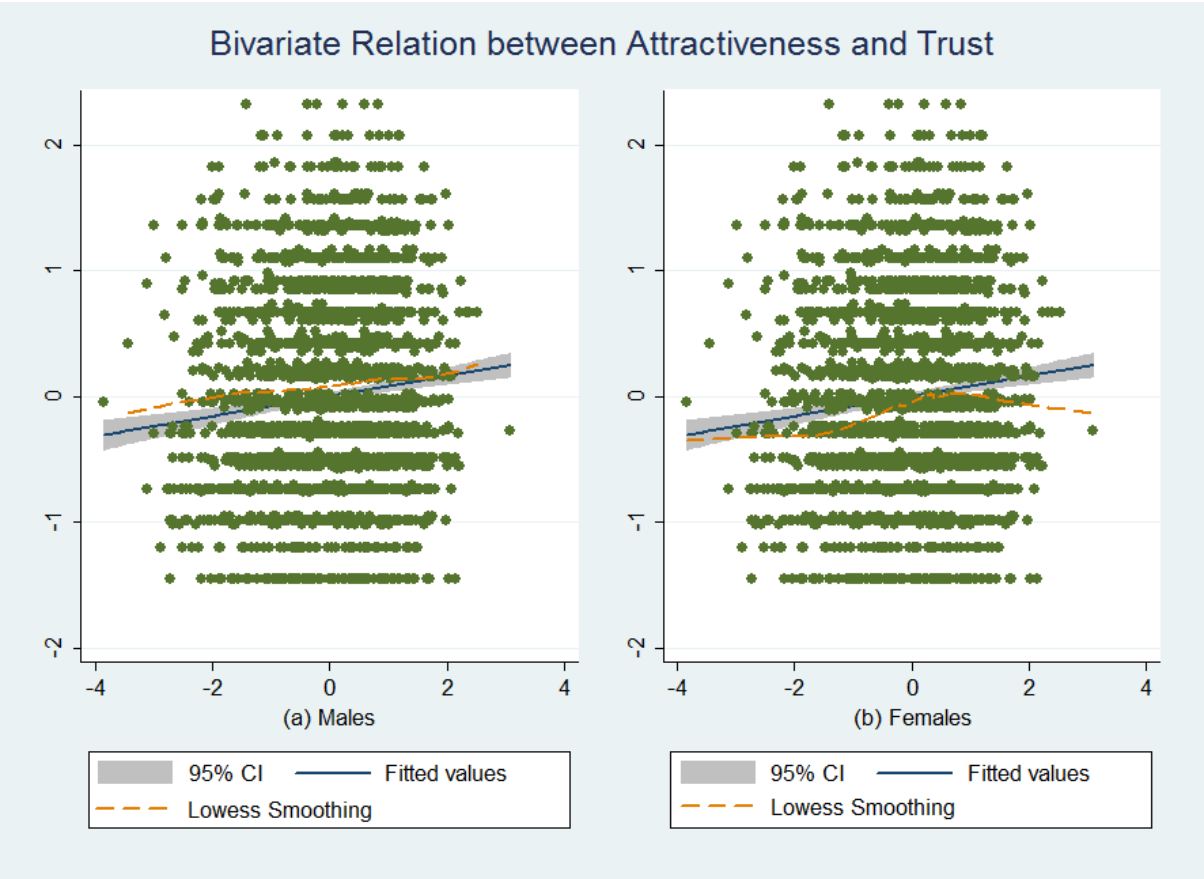

Figure 2: Bivariate Relation between Attractiveness and Trust, by Gender

Source: GGSS/ALLBUS, 2008.

Table 1: Attractiveness distributions, by rating method, gender of interviewer and respondent, and by age

(I)	Rating Method					
	Beginning of Interview		End of Interview		Self-assessed	
Below average looks	487	15.31%	423	13.30%	629	19.78%
Average looks	1732	54.47%	1722	54.15%	1907	59.97%
Above average looks	961	30.22%	1035	32.55%	644	20.25%
(II)	By Sex of Interviewer					
	Male Interviewer		Female Interviewer			
Below average looks	282	15.05%	205	15.70%		
Average looks	1062	56.67%	670	51.30%		
Above average looks	530	28.28%	431	33.00%		
(III)	By Sex of Respondent					
	Male Respondent		Female Respondent			
Below average looks	256	16.37%	231	14.29%		
Average looks	878	56.14%	854	52.85%		
Above average looks	430	27.49%	531	32.86%		
(IV)	By Age of Respondent					
	Respondent < 40 years		Respondent >=40 years			
Below average looks	62	6.47%	425	19.13%		
Average looks	479	50.00%	1253	56.39%		
Above average looks	417	43.53%	544	24.48%		

Source: GGSS/ALLBUS, 2008.

Notes: The attractiveness categories follow Gehrsitz (2014), where “below average” refers to ratings of 5 or lower, “average looks” includes all ratings larger than 5 and smaller than 9, and “above average” refers to ratings of 9 or higher.

Table 2: Attractiveness and Trust – Interviewers' (I) and Respondents' (R) Ratings

Dep. Var.: Trust (standardized)	(1) I: Standardized measure	(2) I: Rating, at beginning	(3) I: Rating, at end	(4) R: Rating, self-reported
Attractiveness	0.035**	0.036***	0.033***	0.022**
	(0.018)	(0.009)	(0.008)	(0.009)
Rating difference (R – I) < –1	-0.016	-0.046	-0.021	0.035
	(0.033)	(0.034)	(0.032)	(0.034)
Rating difference (R – I) > 1	-0.022	0.015	-0.021	-0.073*
	(0.042)	(0.043)	(0.041)	(0.042)
Female	-0.151***	-0.155***	-0.157***	-0.151***
	(0.030)	(0.030)	(0.030)	(0.030)
Age	0.013**	0.014**	0.014***	0.013**
	(0.005)	(0.005)	(0.005)	(0.005)
Age (squared/100)	-0.006	-0.007	-0.007	-0.006
	(0.006)	(0.006)	(0.006)	(0.006)
Education: mid secondary	0.160***	0.157***	0.152***	0.161***
	(0.036)	(0.036)	(0.036)	(0.036)
Education: upper secondary	0.361***	0.356***	0.352***	0.363***
	(0.043)	(0.043)	(0.043)	(0.043)
Is married	0.012	0.009	0.006	0.011
	(0.035)	(0.035)	(0.035)	(0.035)
Living in East Germany	-0.204***	-0.211***	-0.205***	-0.207***
	(0.033)	(0.032)	(0.032)	(0.033)
Log of net household-income	0.047***	0.048***	0.048***	0.050***
	(0.018)	(0.017)	(0.017)	(0.018)
Is not a German citizen	-0.147**	-0.148**	-0.146**	-0.150**
	(0.069)	(0.069)	(0.069)	(0.069)
Has migration background	-0.036	-0.035	-0.034	-0.035
	(0.047)	(0.047)	(0.046)	(0.047)
Has unemployment experience	-0.052	-0.049	-0.049	-0.052
	(0.035)	(0.035)	(0.035)	(0.035)
Is currently employed	0.045	0.041	0.040	0.044
	(0.043)	(0.042)	(0.042)	(0.042)
Self-rated health	0.084***	0.077***	0.078***	0.083***
	(0.015)	(0.015)	(0.015)	(0.015)
Father's educ.: upper secondary	0.087*	0.084*	0.084*	0.085*
	(0.049)	(0.049)	(0.049)	(0.049)
Mother‘s educ.: upper secondary	0.004	0.006	0.007	0.006
	(0.056)	(0.056)	(0.056)	(0.056)
Interviewer: has opposite sex	-0.007	-0.007	-0.008	-0.005
	(0.028)	(0.028)	(0.028)	(0.028)
Interviewer: age	-0.004*	-0.005**	-0.005**	-0.004**
	(0.002)	(0.002)	(0.002)	(0.002)
Interv. educ.: lower secondary	0.063	0.066	0.062	0.063
	(0.041)	(0.041)	(0.041)	(0.041)
Interv. educ.: upper secondary	-0.015	-0.001	-0.001	-0.009
	(0.032)	(0.032)	(0.032)	(0.032)
Interviewer: job experience	-0.006***	-0.006***	-0.006***	-0.006***
	(0.002)	(0.002)	(0.002)	(0.002)
Adj. R2	0.121	0.124	0.125	0.122

Source: GGSS/ALLBUS, 2008.

Notes: Models include a constant and further account for missing values in household size and income, household size itself, missing or no education, living in areas with less than 5,000 or more than 100,000 residents, occupation for the employed, and whether the interviewer is 10 years older or younger than the respondent; N=3,179; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Attractiveness and Trust, by Gender and Accounting for Non-Linear Patterns

(I)	(1) All	(2) Males	(3) Females
Attractiveness (std.)	0.035**	0.006	0.069***
	(0.018)	(0.027)	(0.024)
Female	-0.151***	--	--
	(0.030)		
Adj. R2	0.121	0.095	0.139
N	3,179	1,563	1,616

(II)	(1) All B: Mean – Std.dev A: Mean + Std.dev.	(2) All B: Bottom 25% A: Top 25%	(3) Males B: Mean – Std.dev A: Mean + Std.dev.	(4) Males B: Bottom 25% A: Top 25%	(5) Females B: Mean – Std.dev A: Mean + Std.dev.	(6) Females B: Bottom 25% A: Top 25%
Below: …	-0.054	-0.072*	0.001	0.005	-0.114**	-0.155***
	(0.040)	(0.037)	(0.059)	(0.053)	(0.056)	(0.052)
Above: …	-0.027	-0.011	-0.065	0.025	0.017	-0.033
	(0.041)	(0.035)	(0.064)	(0.051)	(0.054)	(0.049)
Adj. R2	0.120	0.121	0.095	0.095	0.136	0.139
N	3,179	3,179	1,563	1,563	1,616	1,616

Source: GGSS/ALLBUS, 2008.

Notes: All estimates include the set of controls as given in Table 2; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Attractiveness by Gender and Body-Mass-Index

BMI	Males			Females		
	Mean attractiveness	Mean Trust, standardized	N	Mean attractiveness	Mean Trust, standardized	N
<18,5	6.00	0.34	3	8.52	-0.34	17
	(1.00)	(0.79)		(1.90)	(0.64)	
18,5 – 25	7.67	0.12	294	7.97	-0.00	435
	(1.81)	(0.79)		(1.86)	(0.78)	
25 – 30	7.31	0.12	338	7.18	-0.09	214
	(1.88)	(0.81)		(1.96)	(0.81)	
30+	6.59	-0.02	116	6.69	-0.15	104
	(2.08)	(0.84)		(1.93)	(0.71)	
Total	7.33	0.10	751	7.59	-.057	770
	(1.92)	(0.81)		(1.96)	(0.78)	

Source: ISSP-subsample of GGSS/ALLBUS, 2008.

Table 5: Attractiveness and Trust – Models including Individuals' Body-Mass-Index

	(1) All Std. measure	(2) Males Std. measure	(3) Females Std. measure
Attractiveness	0.031	-0.018	0.043
	(0.035)	(0.063)	(0.046)
BMI: 25–30	-0.020	0.006	-0.031
	(0.047)	(0.068)	(0.067)
BMI: 30 +	-0.008	0.005	-0.015
	(0.069)	(0.098)	(0.098)
(BMI: 25–30) x (Attractiveness)	-0.009	-0.031	0.072
	(0.045)	(0.071)	(0.062)
(BMI: 30 +) x (Attractiveness)	0.021	0.066	-0.002
	(0.069)	(0.099)	(0.106)
Height	-0.027	-0.045	0.049
	(0.332)	(0.489)	(0.491)
N	1,466	730	736
Adj. R2	0.139	0.115	0.151
F-test	0.47	0.47	1.73
p-value	0.70	0.70	0.16

Source: ISSP-Subsample of GGSS/ALLBUS, 2008.

Notes: The F-test tests for joint significance of the attractiveness measures and its interactions with the two BMI-categories. All estimates include the set of controls as given in Table 2; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Attractiveness and Trust – Models including FFM-traits

	(1) All	(2) All	(3) Males	(4) Males	(5) Females	(6) Females
Attractiveness (std.)	0.021	--	-0.004	--	0.050**	--
	(0.018)		(0.028)		(0.024)	
Attr.: Bottom 25%	--	-0.028	--	0.052	--	-0.116**
		(0.037)		(0.053)		(0.051)
Attr.: Top 25%	--	-0.029	--	0.026	--	-0.072
		(0.034)		(0.049)		(0.048)
FFM: Openness	0.028*	0.029**	0.002	0.003	0.051**	0.052**
	(0.015)	(0.015)	(0.022)	(0.022)	(0.020)	(0.021)
FFM: Conscientiousness	-0.070***	-0.069***	-0.049**	-0.049**	-0.082***	-0.082***
	(0.015)	(0.015)	(0.021)	(0.021)	(0.022)	(0.022)
FFM: Extraversion	0.070***	0.071***	0.046**	0.047**	0.090***	0.093***
	(0.015)	(0.015)	(0.022)	(0.022)	(0.020)	(0.020)
FFM: Agreeableness	0.231***	0.231***	0.238***	0.239***	0.216***	0.215***
	(0.014)	(0.014)	(0.019)	(0.019)	(0.019)	(0.019)
FFM: Neuroticism	-0.018	-0.018	0.011	0.011	-0.042**	-0.042**
	(0.015)	(0.015)	(0.022)	(0.022)	(0.021)	(0.021)
N	2,965	2,965	1,467	1,467	1,498	1,498
Pseudo-R2	0.208	0.207	0.184	0.184	0.226	0.227

Source: GGSS/ALLBUS, 2008.

Notes: All estimates include the set of controls as given in Table 2; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Attractiveness and “Generalized Trust”, cross-section 2008 and extended sample 2008-14

	(1)	(2)	(3)	(4)	(5)	(6)
(I)	08: All	08: M	08: F	08-14: All	08-14: M	08-14: F
Attractiveness	0.009	0.003	0.015*	0.010***	0.010**	0.010**
	(0.006)	(0.010)	(0.008)	(0.003)	(0.005)	(0.004)
Female	-0.051***	--	--	-0.035***	--	--
	(0.011)			(0.006)		
N	3,291	1,620	1,671	12,865	6,379	6,486
Pseudo-R2	0.049	0.042	0.067	0.054	0.049	0.062
(II)	08: All	08: M	08: F	08-14: All	08-14: M	08-14: F
Attr.: Bottom 25%	-0.034**	-0.013	-0.051***	-0.025***	-0.014	-0.033***
	(0.014)	(0.022)	(0.018)	(0.007)	(0.010)	(0.010)
Attr.: Top 25%	-0.013	0.004	-0.022	-0.002	0.006	-0.007
	(0.013)	(0.021)	(0.016)	(0.007)	(0.011)	(0.009)
Female	-0.050***	--	--	-0.034***	--	--
	(0.011)			(0.006)		
N	3,291	1,620	1,671	12,865	6,379	6,486
Pseudo-R2	0.050	0.043	0.068	0.054	0.049	0.062

Source: GGSS/ALLBUS, 2008-14.

Notes: Ordered probit regressions, average marginal effects for P(y=“most people can be trusted”|X); all estimates include the set of controls as given in Table 2, plus year fixed effects for 2010, 2012, and 2014 in the models (4)-(6); robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Attractiveness and Preferences for Leisure Time Use

	All		Males		Females	
(I)	In free time: prefers to be with other people					
	(1)	(2)	(3)	(4)	(5)	(6)
Attractiveness	-0.001	--	-0.009	--	0.005	--
	(0.015)		(0.023)		(0.020)	
Attractiveness: Bottom 25%	--	-0.049	--	-0.026	--	-0.075*
		(0.031)		(0.045)		(0.045)
Attractiveness: Top 25%	--	-0.032	--	-0.018	--	-0.050
		(0.029)		(0.045)		(0.040)
N	1,550	1,550	750	750	800	800
Adj. R2	0.053	0.054	0.093	0.092	0.014	0.017
(II)	Use free time: Get together with friends (Several times a week or daily)					
	(1)	(2)	(3)	(4)	(5)	(6)
Attractiveness	0.016	--	0.014	--	0.018	--
	(0.015)		(0.022)		(0.020)	
Attractiveness: Bottom 25%	--	-0.051*	--	-0.068	--	-0.025
		(0.030)		(0.044)		(0.043)
Attractiveness: Top 25%	--	0.012	--	-0.013	--	0.050
		(0.029)		(0.042)		(0.041)
N	1,605	1,605	773	773	832	832
Adj. R2	0.162	0.163	0.207	0.208	0.126	0.126
(III)	Take part in sports: To meet other people (Important or very important)					
	(1)	(2)	(3)	(4)	(5)	(6)
Attractiveness	0.028**	--	0.028	--	0.025	--
	(0.013)		(0.018)		(0.018)	
Attractiveness: Bottom 25%	--	-0.049*	--	-0.065*	--	-0.019
		(0.027)		(0.038)		(0.040)
Attractiveness: Top 25%	--	0.044*	--	0.049	--	0.045
		(0.026)		(0.040)		(0.036)
N	1,563	1,563	756	756	807	807
Adj. R2	0.016	0.017	0.013	0.016	0.009	0.008
(IV)	Use free time: Establish useful contacts (Often or very often)					
	(1)	(2)	(3)	(4)	(5)	(6)
Attractiveness	0.039***	--	0.067***	--	0.013	--
	(0.015)		(0.022)		(0.020)	
Attractiveness: Bottom 25%	--	-0.067**	--	-0.099**	--	-0.046
		(0.029)		(0.041)		(0.042)
Attractiveness: Top 25%	--	0.031	--	0.095**	--	-0.028
		(0.031)		(0.048)		(0.041)
N	1,563	1,563	762	762	801	801
Adj. R2	0.055	0.054	0.078	0.079	0.036	0.036

Source: GGSS/ALLBUS, 2008.

Notes: Results from linear probability models. All estimates include the set of controls as given in Table 2; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Number of people in respondents' networks

	(1) All	(2) Males	(3) Females
Attractiveness: Bottom 25%	2.16	2.09	2.24
	(1.31)	(1.36)	(1.25)
Attractiveness: Mid 50%	2.58	2.54	2.62
	(1.23)	(1.23)	(1.2)
Attractiveness: Top 25%	2.76	2.59	2.88
	(1.23)	(1.15)	(1.27)

Source: GGSS/ALLBUS, 2010.

Notes: N=2,584; standard deviations in parentheses.

Table 10: Attractiveness and Individuals‘ Network Size

	(1)	(2)	(3)
	All	Males	Females
Attractiveness (std.)	0.119***	0.110**	0.129***
	(0.029)	(0.044)	(0.039)
Female	0.193***	--	--
	(0.051)		
Adj. R2	0.125	0.089	0.147
Attractiveness: Bottom 25%	-0.248***	-0.312***	-0.194**
	(0.064)	(0.093)	(0.087)
Attractiveness: Top 25%	0.017	-0.079	0.092
	(0.058)	(0.086)	(0.079)
Female	0.200***	--	--
	(0.051)		
Adj. R2	0.124	0.093	0.143
N	2,584	1,264	1,320

Source: GGSS/ALLBUS, 2010.

Notes: All estimates include the set of controls as given in Table 2; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Level of Difficulty to Have a Respondent Participate in the Interview

(I)	(1) very difficult	(2) difficult	(3) easy	(4) very easy
Attractiveness	-0.003**	-0.005**	0.0003*	0.008**
	(0.001)	(0.002)	(0.000)	(0.004)
(II)				
Attractiveness: Bottom 25%	-0.0003	-0.0005	0.00003	0.0008
	(0.003)	(0.005)	(0.0004)	(0.008)
Attractiveness: Top 25%	-0.003	-0.006	0.0004	0.008
	(0.003)	(0.005)	(0.0003)	(0.008)

Source: GGSS/ALLBUS, 2008-14.

Notes: Average marginal effects from ordered probit models; all estimates include the set of controls as given in Table 2, plus year dummies for 2010, 2012, and 2014; N=12,865; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A1: GGSS 2008 – Factor Loadings on Trust

Variable Name	Questions and response options	Factor Loading
v126	Q: Some people think that most people can be trusted. Others think that one can't be careful enough when dealing with other people. What do you think? R: Most people can be trusted ; One can't be careful enough ; It depends.	0.69
v130	Q: In general, how much trust do you have in strangers? R: 1 = Absolutely no trust at all … 7 = A great deal of trust	-0.57
v620	Q: Generally speaking, would you say that people can be trusted or that you can't be too careful in dealing with people? R: People can almost always be trusted ; People can usually be trusted ; You usually can't be too careful in dealing with people ; You almost always can't be too careful in dealing with people.	0.71

Source: GGSS/ALLBUS, 2008.

Notes: In the factor analysis, v126 has been rescaled to ascending order; Cronbach's alpha = 0.736.

Table A2: Attractiveness and Trust – Descriptive Statistics for the 2008 GGSS

Variable	Mean	Std.Dev.	Min	Max
Attractiveness, beginning of interview	7.44	(1.97)	1	11
Attractiveness, end of interview	7.55	(1.93)	1	11
Attractiveness, self-rated	6.92	(1.94)	1	11
“Most people can be trusted” (v126)	2.17	(0.73)	1	3
“How much trust in strangers” (v130)	2.86	(1.33)	1	7
“Generally, people can be trusted” (v620)	2.59	(0.72)	1	4
Is female	0.51	(0.50)	0	1
Age	49.87	(17.77)	18	97
Age (squared/100)	28.03	(18.24)	3.24	94.09
Education: no or missing	0.04	(0.19)	0	1
Education: Lower secondary (Hauptschule)	0.36	(0.48)	0	1
Education: Mid secondary (Realschule)	0.32	(0.47)	0	1
Education: Upper secondary ((Fach-)Abitur)	0.28	(0.45)	0	1
Is married	0.58	(0.49)	0	1
Living in East Germany	0.32	(0.47)	0	1
Living in rural area	0.19	(0.39)	0	1
Living in urban area	0.28	(0.45)	0	1
Household size: missing	0.00	(0.07)	0	1
Household size	2.52	(1.31)	0	12
Household-income: missing	0.17	(0.38)	0	1
Log of net household-income	6.23	(2.92)	0	10.15
Is not a German citizen	0.06	(0.24)	0	1
Has migration background	0.16	(0.36)	0	1
Has unemployment experience	0.22	(0.41)	0	1
Is currently employed	0.49	(0.50)	0	1
Self-rated health	3.56	(1.05)	1	5
Father’s education: Upper secondary	0.14	(0.35)	0	1
Mother’s education: Upper secondary	0.09	(1.05)	0	1
Interviewer: is female	0.41	(0.49)	0	1
Interviewer: has opposite sex than respondent	0.48	(0.49)	0	1
Interviewer: age	59.54	(8.76)	27	77
Interviewer: 10 yrs younger than respondent	0.18	(0.39)	0	1
Interviewer: 10 yrs older than respondent	0.50	(0.50)	0	1
Interviewer education: Lower secondary	0.19	(0.39)	0	1
Interviewer education: Upper secondary	0.35	(0.48)	0	1
Interviewer experience	12.42	(9.69)	0	40
Body-Mass-Index (N=1,521)	25.69	(4.50)	14.69	56.82

Source: GGSS/ALLBUS, 2008. N=3,179.

[1] It is interesting to note that economists also dealt with stereotyping at the beginning of the 1970s, leading to the concept of statistical discrimination (Arrow, 1973; Phelps, 1972).

[2] Allgemeine Bevölkerungsumfrage (ALLBUS); for details see Terwey (2000) or, in German language, Koch and Wasmer (2004).

[3] This breakdown follows Gehrsitz (2014) who also used the 2008 GGSS in his analysis of the labor market returns to beauty. A comparison shows that the distributions here are somewhat less favorable which can be explained by the different age composition of the sample: Whereas Gehrsitz (2014) analyses the working age population, this study comprises the full age range.

[4] A helpful overview of surveys that contain the "generalized trust" item or one of its variants can be found at: http://www.esds.ac.uk/international/resources/trust.asp [last accessed August 05, 2015].

[5] See e.g. Bjørnskov (2007) and particularly Naef and Schupp (2009) for critical reflections on the original item and how its deficiencies can be dealt with.

[6] As a sensitivity check, additional analyses employ only the "trust into strangers" measure. The results are similar to the main findings, available upon request.

[7] This definition captures some 55% percent of the sample, including 23% with difference in ratings; 30% percent of the respondents are more pessimistic and have lower self-ratings compared to the interviewer, whereas 15% are more overconfident. Experiments with different thresholds do not alter the findings as outlined in the following.

[8] The results for the control variables in all the following estimations are quite similar to the ones discussed here. Full regression results are available upon request.

[9] This is in line with the findings of Alesina and La Ferrara (2002), and, also using the GGSS, Rainer and Siedler (2009), although there seems to be no clear evidence on the direction of the gender-trust differential in the literature.

[10] It is not quite clear, why this should be the case. There, to the best of my knowledge, is however no substantial or survey-methodological evidence that examines this issue. Addressing this question in more detail is interesting, but way beyond the interest of this paper, so this is left for future research.

[11] The anthropometric characteristics were surveyed only for individuals participating in the ISSP module "Leisure and Sports". This however reduces the sample by about a half.

[12] Further estimations also accounted for weight, but this does not change the findings.

[13] These results are not shown to save space, but available upon request. Note further, that attractiveness is statistically related to trust in models without (BMI x attractiveness) interaction terms. Using the attractiveness rating as is also results in attractiveness coefficients that are statistically different from zero, in both the base models and those with interactions terms. This again highlights the importance of using the standardized measure in order to prevent false implications.

[14] Note that including the FFM indicators lowers the sample size by almost 7 percent as they have been generated from the two ISSP modules "Leisure and Sports" and "Religion" that were also implemented in the 2008 GGSS, but come along with some non-response. The FFM is implemented as a short 10-item scale, the

BFI-10, developed by Rammstedt and John (1997). While the authors claim the BFI-10 scales retain significant levels of reliability and validity, the Cronbach's alpha values in this sample (O: 0.410, C: 0.440, E: 0.603, A: 0.133, N: 0.499) point to severe measurement issues, particularly for agreeableness.

[15] The relationship between individuals' attractiveness and their personality is interesting in itself and has a long history in psychology (for a meta-analysis, see Feingold, 1992). Here, additional analyses use data from the 2010 GGSS wave that provides instruments for individuals' positive and negative reciprocity. The results point to no association between attractiveness and positive reciprocity. Yet, attractiveness and negative reciprocity are on average negatively linked which is driven by females, whose rating scores in the 25% bottom part of the beauty distribution, who seem to be more inclined to retaliate. To save space, results are not provided here, but are available upon request.

[16] As a somewhat loose complement to this, further analyses that use data from the 2014 GGSS examine whether individuals are more likely to believe in nepotism or meritocracy as a means to success. Although the results would imply that more attractive individuals are more (less) likely to believe in nepotism (meritocracy), no coefficient is statistically different from zero.

[17] There are two questions in the 2010 wave from which the network-size variable here is generated. One question limits the number of persons to three and the second question has a top-coding at "6 and more". Unfortunately, because of a sample-split, respondents had to reply to either one of the two questions, but not both.

[18] Similar to the trust equations, it is the respondents' education which matters most among the vector of individuals' characteristics, followed by whether the respondent lives in East Germany and by gender (full results available upon request).