The Role of Body Size in Economic Research Above and Beyond Beauty*

Sonia Oreffice
University of Surrey and IZA

Climent Quintana-Domeque
University of Oxford and IZA

April 2015

Abstract

We analyze how attractiveness rated at the start of the interview in a nationally representative sample is related to weight, height, and body mass index (BMI), separately by gender and accounting for interviewers' characteristics or fixed effects. We also compute the non-anthropometric residual attractiveness, and present novel estimates of how non-anthropometric attractiveness and anthropometric attributes are related to labor and marital outcomes such as hourly wage and spousal education. We show that height, weight, and BMI all strongly contribute to male and female attractiveness when attractiveness is rated by opposite-sex interviewers, and that anthropometric characteristics are irrelevant to male interviewers when assessing male attractiveness. In addition, we estimate that non-anthropometric attractiveness and height matter for both men and women in the labor market, while BMI plays a stronger role than (residual) beauty in the marriage market.

JEL Classification Codes: J01, J10.
Keywords: Attractiveness, Body Mass Index, Height, Weight, Wage, Spousal Education.

*Oreffice (corresponding author): University of Surrey, School of Economics, Guildford, GU2 7XH, United Kingdom; s.oreffice@surrey.ac.uk.

We thank Damian Clarke, Simon Quinn, Curtis Simon and seminar participants at the IFS and the University of Oxford for useful comments. The usual disclaimers apply.

1 Introduction

The salience of physical attributes to economic behavior and socioeconomic outcomes is well-established in the social sciences, where research consistently reports that beauty or anthropometric measures (height, weight, and body mass index (BMI)[1]) are significantly related to labor and marriage market outcomes.[2] In this paper, we examine how anthropometric characteristics are related to beauty, estimate the non-anthropometric residual component of attractiveness, and investigate the relationship of all these measures with socioeconomic outcomes.

We use nationally representative German data where the respondents provide information on their anthropometric attributes (height and weight) and the interviewers assess their attractiveness at the start of the interview on an 11-point Likert scale. The fact that our attractiveness measure is based on the interviewer rating respondents' overall attractiveness rather than a photograph of part of the body, that this happens at the start of the interview, that our sample is nationally representative, and that we can control for interviewers' characteristics (including fixed effects), allow us to provide a novel answer to the question: Do anthropometric characteristics contribute to attractiveness? When and to what extent?

Once attractiveness is predicted using both anthropometric and interviewer's characteristics, we decompose it netting out its anthropometric component and obtaining the residual non-anthropometric attractiveness. Thus, we work with three measures of beauty, namely attractiveness, anthropometric measures (height and weight or BMI) and the residual non-anthropometric attractiveness, to estimate how they are associated with "premia" in the labor and marriage markets, in terms of hourly wages and spousal education. As such, we reassess one of the open questions in the social sciences regarding beauty and socioeconomic outcomes: Is it non-anthropometric attractiveness, anthropometric measures or both that

[1]BMI is defined as the individual's body weight (in kg) divided by the square of his/her height (in meters).

[2]For example, Averett and Korenman (1996), Biddle and Hamermesh (1998), Caliendo and Gehrsitz (2014), Cawley (2004), Scholz and Sicinski (2014).

matter(s) for socioeconomic outcomes? When and why?

The existing research has unveiled several interesting patterns typically using either beauty or anthropometric measures, but not both types of measures in a *nationally* representative sample. This is unfortunate, because it has not been established yet whether and how anthropometric measures should be conceptually distinguished from beauty measures, or whether the observed anthropometric "premia" (or "penalties") in the labor and marriage markets simply reflect "beauty premia". For instance, Hamermesh and Biddle's (1994) seminal work on beauty and the labor market considers *during*-interview ratings of height, weight, and beauty on *non*-representative samples.[3]

Our analysis uses the German General Social Survey (ALLBUS) data for 2008 and 2012, two nationally representative cross-sections of the German population. We run least squares regressions of attractiveness rated at the start of the interview on anthropometric measures and several groups of control variables, including age, region, year, interviewer fixed effects, and interactions with gender and age group of the interviewer. We find that height, weight and BMI all strongly contribute to male and female attractiveness when attractiveness is rated by opposite-sex interviewers, whereas only female anthropometric measures are relevant when attractiveness is assessed by same-sex interviewers. To the best of our knowledge, we are the first to estimate that anthropometric characteristics are irrelevant to male interviewers in assessing male attractiveness, while they are important for both male and female interviewers in assessing female attractiveness. We unveil that this differential beauty assessment by interviewer's gender cannot be explained by the existence of a non-monotonic relationship between beauty and BMI.

This is a new and intriguing finding prompting future researchers to seriously consider and account for the gender of the interviewer in any beauty analysis. Hamermesh and Biddle (1994) write that "within a culture and at a point in time there is tremendous agreement on standards of beauty". Our analysis provides a clarification of such a statement: we show

[3]In psychology, Weeden and Sabini (2007) use measures of face photo ratings, body size, and the residual attractiveness component in a sample of undergraduate students to study their sexual behavior.

that these standards and their anthropometric determinants may differ by *gender*.

One immediate concern regarding our evidence on the role of the gender of the interviewer is that (on average) individuals interviewed by same-sex individuals may be different than those interviewed by opposite-sex interviewers. However, when testing for mean differences of respondents by sex of the interviewer (individually or simultaneously), and separately for men and women, we find that individuals interviewed by same-sex interviewers have the same average characteristics as those interviewed by opposite-sex interviewers.

In the second part of our analysis, we decompose beauty and investigate the presence of beauty "premia", specifically of anthropometric and residual non-anthropometric "premia" in the labor and marriage markets. We show that attractiveness, its non-anthropometric residual component, and height, matter in the labor market in terms of higher wages for both men and women. Male BMI is non-monotonically related to wages, consistent with Caliendo and Gehrsitz (2014). In the marriage market, women's attractiveness and its non-anthropometric component, weight or BMI, are all significantly related to their husbands' education, whereas for men neither attractiveness nor its non-anthropometric component matter, with only height and BMI exhibiting a weak link with their wives' education. We also compute the marital trade-off between the attributes of own BMI and own education, and cannot reject that these trade-offs are the same for married men and women.

Our evidence provides the missing econometric rationale for the interpretation that BMI is one of the relevant dimensions of attractiveness in the marriage market for both men and women, in line with the recent findings in Chiappori, Oreffice and Quintana-Domeque (2012) and Oreffice and Quintana-Domeque (2010, 2012), and earlier evidence that heavier women tend to have poorer husbands (Averett and Korenman, 1996). Furthermore, comparing the role of body size to the estimated non-anthropometric component of beauty indicates that anthropometric attributes matter beyond their relationship with beauty, at least in the marriage market.

Overall, our paper represents a step forward in our understanding of the role of different physical attributes in determining beauty, and in how body size and beauty explain socioeconomic outcomes. Our study contributes to bridge the gap between the literature on BMI and economic outcomes, on one hand, and beauty and economic outcomes, on the other, estimating the relevance of body size and beauty in a nationally representative sample and comparing them in two different markets. One key message is that when "translating" BMI into attractiveness, the gender of the rater/interviewer should be taken seriously. At the same time, our findings of the interplay of attractiveness, anthropometric characteristics and non-anthropometric (residual) attractiveness in the labor and marriage markets suggest that body size is not overshadowed by beauty.

The paper is organized as follows. Section 2 describes the data and discusses the approaches and challenges to measuring beauty, and our contribution, also providing a survey of the existing attractiveness measures and related economic papers. Section 3 presents a standard statistical model to understand the determinants of beauty. Section 4 analyzes the role of anthropometric attractiveness in explaining attractiveness after accounting for interviewers' characteristics. Section 5 explains how to measure non-anthropometric residual attractiveness, based on the previous estimation of attractiveness models, and how to analyze the role of non-anthropometric attractiveness and anthropometric measures on the labor and the marriage markets. Section 6 reports the results on attractiveness, non-anthropometric attractiveness, anthropometric measures, hourly wages and spousal education. Section 7 concludes the paper.

2 Data Description and Measuring Attractiveness

2.1 The data set

Estimation is carried out on the German General Social Survey (ALLBUS) data (GESIS, 2014), a biennial survey that started in 1980 on "the attitudes, behaviour, and social structure of persons resident in Germany": a nationally representative cross-section of the German population is interviewed every two years, and detailed demographic and socioeconomic information at the individual and household level is collected for thousands of respondents. In addition, the interviewer's identifier and main demographic characteristics (age and sex) are also recorded, which would prove useful in our present analysis.[4]

We use the cumulative series ALLBUS GESIS-Cumulation 1980-2012, focusing our study on the waves of 2008 and 2012, i.e., the *only* waves containing information on *both* attractiveness *and* anthropometric measures, where 2012 is the most recently released wave.[5] Our main variables of interest are height (in cm), weight (in kg), and BMI (body mass index) of the respondent as well as his/her attractiveness, which is rated by the interviewer at the *start* of the interview. The respondent's attractiveness rating is reported on an 11-point (Likert) scale from 1 to 11 (from unattractive to attractive).

To perform our analysis, we work with an additional set of variables: age, gender, a West-East region and a 2008-2012 year dummy-variable indicators, and education. In the ALLBUS data, educational attainment is measured through a series of yes/no questions on the attainment of specific types of certificates in schools and universities according to the features and dual paths of the German education system. We construct a binary vari-

[4]Interviews are performed with CAPI (computer assisted personal interviewing). In 2008 a total of 3,469 respondents participated in the survey, which was conducted between March and August of 2008 by 185 interviewers. The response rate was 51.3% (see Menold and Zuell (2010) for details), higher than the 40.2% response rate in the 2006 cross-section of the GSOEP (German Socio-Economic Panel).

[5]In the ALLBUS some questions are asked in some or alternate waves. The anthropometric measures are not available in 2010 or in the years before 2008. An additional feature of these anthropometric measures is that they are not asked in the basic questionnaire but in the rotating ISSP modules "Health" or "Leisure time and sports" to about 50% of the respondents in selected years (other respondents are asked other "split" questionnaires).

able which takes value of 1 if the respondent has a university or polytechnic degree, or a master/technician college certificate (i.e., "some college and above"), and 0 otherwise.[6]

To be able to measure socioeconomic outcomes of the respondents and their spouses, we also consider the respondent's own net monthly income and hours worked per week to generate the log of his/her hourly wage rate (own net monthly income divided by hours worked), and the spouse's education. Note that neither own net monthly income nor hours worked per week is available for the spouse.

The main analysis considers men and women who are German citizens born in Germany, between 25 and 50 years of age and with BMI in the range 18.5 to 39.99, to keep uniform reference groups with respect to attractiveness and marital and labor market outcomes, and to exclude (medically) morbid obese or underweight individuals (WHO, 2009). The restriction on place of birth and German citizenship is prompted by the fact that being foreign-born may be related differently to attractiveness. Finally, observations are weighed by the available East-West weight to adjust for the oversampling of East German respondents and make the sample nationally representative.

2.2 Measuring attractiveness

In the large body of literature on attractiveness or anthropometric measures, and related outcomes, the key challenge has always been to find a "good" measure of beauty. Even more rare is the availability of anthropometric and attractiveness measures at the same time in a representative sample. Beauty is hard to measure and difficult to have access to in a representative data set, for a variety of reasons, including the availability of only facial photographs, the potential contamination with the personality or grooming of the rated individuals, or with the raters' role into the actual measure. Another challenge is associated with the information on basic characteristics of the raters, that is often absent or unexploited in previous work. Indeed, Scholz and Sicinski (2014) state that "high quality data on beauty,

[6]This schooling variable and the related dummy are not defined for those respondents who are still in school.

augmented with household economic and demographic characteristics, are rare."

In the next subsections we survey the most popular measures of beauty and the economic studies that use them, to emphasize the challenges that this literature faces as well as to highlight the novel approach undertaken by our analysis.

2.2.1 Attractiveness at the start versus at the end of the interview

While the seminal work of Hamermesh and Biddle (1994) on beauty and the labor market uses during-interview ratings, Biddle and Hamermesh (1998) acknowledge that an attractiveness rating during or at the end of the interview would be "contaminated by other information about the subject obtained during an interview" and by other factors related to the interview process per se. For instance, it is likely that the personality of the respondent affects the rating, with polite and friendly participants receiving higher attractiveness ratings.[7] In this paper, we follow Gehrsitz (2014) and Hamermesh and Abrevaya (2013) in using the measure recorded *at the start* of the interview, and in addition we consider the numerical identifier, age and gender of the interviewer.

2.2.2 Facial attractiveness versus full-body attractiveness

Most studies actually measure attractiveness from facial photographs, rated by undergraduate students or other individuals (Biddle and Hamermesh, 1998; Hitsch, Hortaçsu and Ariely, 2010; Mobius and Rosenblat, 2006, Scholz and Sicinski, 2014). For instance, Biddle and Hamermesh (1998) consider matriculation photographs of law school students, while Scholz and Sicinski (2014) use senior year high school yearbook photographs for male high-school graduates from Wisconsin. Hitsch, Hortaçsu and Ariely (2010) estimate mate preferences and sorting patterns from an online dating service in the US, using profile photographs uploaded by users and rated by students, with photographs available *only* for

[7]Extremely few attempts have been made to measure this contamination: Doorley and Sierminska (2012) compare the rating at the start to the one at end of the interview. Robins, Romer and French (2011) consider a sample of very young workers and regress their wages on end-of-interview attractiveness rating, considering personality and grooming as additional controls.

27.5% of the sample (a self-rated attractiveness measure is also used). We see the use of our "full-body" measure of attractiveness as an advantage with respect to the "facial" measures of attractiveness (such as facial photographs). Indeed, Biddle and Hamermesh (1998) state that "a photograph captures only facial features and to some extent grooming, and captures them imperfectly". Among the few analyses not using facial beauty, Hamermesh, Meng and Zang (2002) analyze primping and beauty of working women from Shangai using *end*-of-interview ratings, while Gehrsitz (2014) considers start-of-interview measures and labor supply.

One criticism to the "full-body" measure of attractiveness is the fact that it may capture other characteristics, such as grooming and fine clothes. It is true that individuals attending a job market interview or going on a first date are likely to dress up to make an impression, and this may correlate to the success of the date or the job market interview (and other unobservable characteristics). However, we honestly do not expect interviewees of the biannual ALLBUS nationally representative survey to dress up in their own home to make an impression on the interviewers (or that dressing up would depend on their beauty/anthropometric characteristics). Finally, the same type of concern can be put forth for most facial measures of attractiveness: faces can (if anything) be "modified" with make-up, contact-lenses, earrings, hairstyle, etc.

2.2.3 Anthropometric characteristics and attractiveness

A relevant strand of literature uses anthropometric measures such as height, weight, or BMI to assess their impact on labor or marriage market outcomes. The few studies on marriage market that consider anthropometric measures as mate attributes, directly use BMI (or obesity) as a proxy for physical attractiveness (Averett and Korenman, 1996; Oreffice and Quintana-Domeque, 2010; Chiappori, Oreffice and Quintana-Domeque, 2012). The same approach is undertaken by the empirical analyses of how wages and employment status are related to BMI and obesity (Brunello and d'Hombres, 2007; Caliendo and Gehrsitz, 2014; Cawley, 2004; Garcia and Quintana-Domeque, 2007).

Articles bringing together measures of some anthropometric attributes and attractiveness include Scholtz and Sicinski (2014), Hitsch, Hortaçsu and Ariely (2010) and Hamermesh and Biddle (1994). Scholtz and Sicinski (2014) control for height in addition to facial attractiveness to estimate the effects of beauty on male life-time earnings. Hitsch, Hortaçsu and Ariely (2010) estimate mate preferences in online dating, and claim that height and weight are self-reported with "only small levels of misrepresentation", although the nature of measurement error in an online dating service is worrisome.[8] Finally, the seminal work of Hamermesh and Biddle (1994) includes in some wage regressions the height and weight ratings, in addition to the beauty rating, and concludes that anthropometric characteristics are not relevant. However, not only both types of characteristics are assessed by the interviewer, but height and weight are reported in categories.[9]

2.3 Measurement error in self-reported anthropometric measures

Self-reported anthropometric measures and measurement error. It is well-known that self-reported anthropometric variables contain measurement error. In the US, Thomas and Frankenberg (2002), using data from the National Health and Nutrition Examination Survey III (NHANES III), show that men tend to overstate their height by around 1 cm until age 50 (Figure 1), after that the overstatement increases with age (Figure 2). Women also overstate their height but the extent of overstatement is small (and not significant) until they reach age 50 (Figure 1), after that the overstatement increases with age (Figure 2). On average, men tend to overstate their weight by nearly two-thirds of a kilogram. Women tend to understate weight by nearly 1.5 kg. After age 50, the understatement decreases with age (Figure 2). It is remarkable to note that the average bias (difference) is

[8] If anything we would expect these measures to be reported with nonclassical errors, that is, people over-reporting their heights (and their other "good" traits) and under-reporting their weights (and their other "negative" attributes). In particular, this is a concern also for the self-reported annual income of the users' dating profiles, with this variable being available for only 50% of the sample.

[9] The same can be said for Hamermesh, Meng and Zhang (2002), where estimates of height and weight are not reported.

constant for the age group of analysis in our paper (25-50 years old).

[Figure 1 about here]

[Figure 2 about here]

What are the implications of measurement error? Cawley (2004) uses the same data as Thomas and Frankenberg (2002) to estimate the relationship between measured height and weight and their self-reported counterparts. He estimates regressions of the corresponding measured variable to its self-reported counterpart by age and race. Then, assuming transportability, he uses the NHANES III estimated coefficients to adjust the self-reported variables from the National Longitudinal Survey of Youth. His estimates of the effect of BMI on wages are *very similar*, accounting or not for measurement error. Hence, and following Caliendo and Gehrsitz (2014), who note that there is no such benchmark study available for Germany, we refrain from adjusting our anthropometric measures.

2.4 Descriptive statistics

Table 1 presents the descriptive statistics of the respondents and the interviewers. As to respondents, their average age is about 39, and 80% of the them lives in the former West Germany. Women report being on average less educated than men, while men exhibit a higher average BMI but a lower average attractiveness than women. Men are slightly overweight (with an avearage BMI of 26.2), and their mean attractiveness rating is 7.7, while women score 8.2 on average, although the rating standard deviations are the same. This higher mean female rating is consistent with a large body of findings across disciplines and data sets, reporting that on average women are rated more attractive than men (e.g., Doorley and Sierminska, 2012; Gehrsitz, 2014; Hamermesh and Biddle, 1994).

[Table 1 about here]

The features in Table 1 indicate that the distribution of the key variables in the ALLBUS data, such as anthropometric measures, are comparable to other well-known German data sets (e.g., GSOEP) and stylized facts (OECD, 2014). In addition, our interviewers are numerous (258), 38% of them are women, and on *average* 59 years old. Hence, they are older than the respondents, whose average age is 39, and less likely to be female (48% of respondents are women).

[Figure 3 about here]

Finally, Figure 3 displays the histograms of the attractiveness ratings for men and women. Reassuringly, almost all values (from 1 to 11) are used by the interviewers when rating both male and female respondents.

3 The Determinants of Beauty

3.1 Decomposing Beauty

Suppose that interviewer j assigns to individual i the attractiveness score A_{ij}. This score depends on the observable (to the econometrician) characteristics of the interviewee ($\mathbf{x}_i$) and of the interviewer ($\mathbf{x}_j$) and the unobservable (to the econometrician) interviewer-interviewee characteristics (ϵ_{ij}). Formally, A_{ij} can be decomposed into two components

$$A_{ij} = E(A_{ij}|\mathbf{x}_i, \mathbf{x}_j) + \epsilon_{ij} \quad (1)$$

where $E(A_{ij}|\mathbf{x}_i, \mathbf{x}_j)$ is the conditional expectation function (CEF) of attractiveness –the part of attractiveness of individual i explained by $\mathbf{x}_i$ and $\mathbf{x}_j$– and ϵ_{ij} is the part of attractiveness orthogonal to any function of $\mathbf{x}_i$ and $\mathbf{x}_j$.

3.2 A Linear Regression Model of Beauty

If the CEF is linear, then

$$E(A_{ij}|\mathbf{x}_i, \mathbf{x}_j) = \alpha_j + \mathbf{x}_i\boldsymbol{\beta} + \mathbf{x}_i\mathbf{x}_j\boldsymbol{\theta} \quad (2)$$

where α_j captures interviewer fixed effects.[10] Hence,

$$A_{ij} = \alpha_j + \mathbf{x}_i\boldsymbol{\beta} + \mathbf{x}_i\mathbf{x}_j\boldsymbol{\theta} + \epsilon_{ij} \quad (3)$$

The first part of our study is devoted to estimate this regression.

[10]Even if the CEF is not linear, (2) provides the best linear approximation to it.

3.3 Taking the Regression to the Data

We will estimate regression (3) separately for female and male respondents. In our context $\mathbf{x}_i$ is a vector of anthropometric $\mathbf{a}_i$ (weight, height, or BMI) and demographic $\mathbf{d}_i$ (age, region, year) characteristics of the respondent. Similarly, we can think of $\mathbf{x}_j$ as a vector of demographic and anthropometric characteristics of the interviewer. Unfortunately, while we have information on demographic characteristics of the interviewer (such as gender and age), we lack information on their anthropometric characteristics.[11] Hence, $\mathbf{x}_j = \mathbf{d}_j$ is a vector of demographic characteristics (gender and age-group of the interviewer). Our most complete model of attractiveness is

$$A_{ij} = \alpha_j + \mathbf{a}_i\boldsymbol{\beta}_0 + \mathbf{a}_i\mathbf{d}_j\boldsymbol{\theta}_0 + \mathbf{d}_i\boldsymbol{\beta}_1 + \mathbf{d}_i\mathbf{d}_j\boldsymbol{\theta}_1 + \epsilon_{ij} \tag{4}$$

where $\boldsymbol{\theta}_0$ captures the potential *heterogeneous* relationship between the anthropometric characteristics and the attractiveness of the interviewee depending on the interviewer's characteristics (age and gender). In particular, in our analysis

- $\mathbf{a}_i$ contains either weight and height or BMI of respondent i,
- $\mathbf{d}_i$ contains age and region of respondent i, and year of the interview,
- $\mathbf{d}_j$ contains two indicator dummy variables: female-indicator and same-age group (50 years or less) interviewer.

Furthermore, we will also allow for a quadratic term in $\mathbf{a}_i$ (BMI^2). Regressions are weighted to adjust for the over-sampling of East German respondents.

[11] To the best of our knowledge no study has interviewers' anthropometric characteristics in a representative sample.

4 Do Anthropometric Measures influence Beauty?

Table 2 displays a series of least square regressions of attractiveness on anthropometric measures for women; in Panel I the focus is on weight and height, while BMI is the focus of Panel II. In each panel, we present five types of regressions depending on the control variables used.[12] The point estimates of the regressions of attractiveness on height and weight in Panel I indicate that taller and slimmer women are ranked as more attractive by interviewers: the corresponding estimated coefficients are all significant at the 1% level and of very similar size across all the specifications, indicating that the interviewer characteristics do not affect at all how female anthropometric characteristics "translate" into beauty. When we control for both interviewer's fixed effects and interactions of their age and gender with all the respondents' characteristics, the influence of anthropometric characteristics on attractiveness is, if anything, slightly larger (column 5). Panel II conveys the same pattern of results for BMI: it reveals that, across specifications, female BMI is negatively correlated with female beauty.

[Table 2 about here]

It is worth noting that these point estimates are sizable: for example, column (4) indicates that a one standard deviation increase in weight is associated with an approximately 0.3 standard deviation decrease in attractiveness, while a one standard deviation increase in height is associated with a 0.2 standard deviation increase in attractiveness. For BMI, a one standard deviation increase would lead to a 0.3 standard deviation decrease in attractiveness. These findings clearly indicate that: first, weight and height (or BMI) of a woman are relevant in explaining her attractiveness; second, the extent to which women's height, weight, and BMI are considered attractive is independent of the gender (and age) of the interviewer.

[12] Heteroskedasticity robust standard errors clustered at the interviewer's level are used in all the empirical analysis.

Table 3 presents our empirical findings on the role of anthropometric characteristics in explaining male attractiveness. Panel I shows that male height is positively correlated with male attractiveness, in columns (1)-(3), while weight does not seem to be relevant at all. However, this zero result changes dramatically when we allow for the gender of the interviewer to be interacted with the anthropometric characteristics of the respondent. Indeed, columns (4) and (5) reveal that male weight is negatively assessed by *female* interviewers. Panel II reveals the same pattern for BMI: while male BMI is irrelevant to male interviewers, it is clearly not neutral to female interviewers.

[Table 3 about here]

The evidence in Tables 2 and 3 indicates that the rating of female attractiveness does not vary by the gender of the interviewer, while that of male attractiveness does. Male interviewers do not consider anthropometric attributes at all when assessing male attractiveness, whereas for female interviewers, looking at column (4), we can see that: a one standard deviation increase in male weight (or BMI) is associated with a 0.2 standard deviation decrease in attractiveness, and a one standard deviation increase in height is associated with a 0.2 standard deviation increase in male attractiveness. Tables A1 and A2 in the Appendix provide a comparison of ordered probit against least squares estimates: no differences are detected.

These are intriguing findings, and we are the first to document them, at least on a nationally representative sample. They may help explain the commonly observed similar correlation between own BMI and spousal socioeconomic status of men and women. In the US, Chiappori, Oreffice and Quintana-Domeque (2012) show that heavier individuals (men or women) tend to have less educated and heavier spouses. We contend that BMI (or weight controlling for height) is in general negatively valued in the marriage market because individuals in the heterosexual marriage market are of opposite-sex with respect to potential spouses.

This new interpretation is made possible by our approach unravelling the influence of the gender of the interviewer, without simply resorting to the argument that tastes for female and male attributes are different in the US and Germany, or, more generally, that *de gustibus non est disputandum.*

Before concluding this section we discuss two potential concerns: the presence of non-monotonicities, and that (on average) individuals interviewed by same-sex individuals may be different from those interviewed by opposite-sex interviewers. Table 4 shows that while there is no evidence of a non-monotonic relationship between female attractiveness and BMI, there is evidence of a quadratic relationship with male attractiveness: the relationship is positive up to a certain threshold, after which it becomes negative. This confirms that the differential beauty rating by the interviewer's gender is not due to non-monotonicities.

Regarding the characteristics of the interviewers, Table 5 shows that the gender of the interviewer and the gender of the interviewee are independent. In addition, Table 6 reports the respondents' mean characteristics by interviewer's sex. Remarkably enough, when testing for mean differences of respondents by sex of the interviewer (individually or simultaneously), and separately for men and women, we find that individuals interviewed by same-sex interviewers have the same average characteristics as those interviewed by opposite-sex interviewers. In particular, the means of weight, height and BMI are the same for those interviewed by same-sex interviewers and those interviewed by opposite-sex interviewers: there is no evidence that the sex (or age group) of the interviewer is related to the way the respondent reports his/her anthropometric measures. If anything, sex of the interviewer can be thought of as being as good as randomly assigned across respondents.[13]

[Table 4 about here]

[Table 5 about here]

[Table 6 about here]

[13] We have also tested for the presence of "halo effects" by including dummy variables for the type of building where the respondent lives (results available upon request).

5 Decomposing Beauty: Estimating its Non-anthropometric Component

In the first part of the paper we have investigated the role of anthropometric characteristics in explaining attractiveness, so that we can use our previous estimates to predict attractiveness $\widehat{A}_{ij}$ as follows

$$\widehat{A}_{ij} = \widehat{\alpha}_j + \mathbf{a}_i \widehat{\boldsymbol{\beta}}_0 + \mathbf{a}_i \mathbf{d}_j \widehat{\boldsymbol{\theta}}_0 + \mathbf{d}_i \widehat{\boldsymbol{\beta}}_1 + \mathbf{d}_i \mathbf{d}_j \widehat{\boldsymbol{\theta}}_1 \tag{5}$$

This allows us to construct a measure of non-anthropometric attractiveness as the (predicted) residual from our estimated attractiveness regression

$$\widehat{NA}_{ij} = A_{ij} - \widehat{A}_{ij} \tag{6}$$

Hence, we define non-anthropometric attractiveness $\widehat{NA}_{ij}$ for individual i assessed by interviewer j, as the difference between the attractiveness score, A_{ij}, and the predicted attractiveness $\widehat{A}_{ij}$ using anthropometric characteristics (together with the interviewee's and interviewer's demographic characteristics).

5.1 Beauty and its Components in the Labor Market

We estimate the following regressions separately for women and men

$$\ln wage_i = \delta^L + \rho^L A_{ij} + \mathbf{d}_i \boldsymbol{\phi}^L + \sigma^L E_i + u_i^L \tag{7}$$

$$\ln wage_i = \delta^{L'} + \rho_1^{L'} \widehat{NA}_{ij} + \mathbf{a}_i \boldsymbol{\rho}_0^{L'} + \mathbf{d}_i \boldsymbol{\phi}^{L'} + \sigma^{L'} E_i + u_i^{L'} \tag{8}$$

where $\ln wage_i$ is the log of hourly wage rate of individual i and E_i is education (1 if some college and above, 0 otherwise) of individual i. Equation (7) is a standard wage regression to measure the beauty (A_{ij}) "premium", whereas equation (8) contains non-anthropometric attractiveness $\widehat{NA}_{ij}$ and anthropometric measures $\mathbf{a}_i$. Since $\widehat{NA}_{ij}$ is a generated regressor, we need to account for its variability when computing the standard errors. To that end, we bootstrap our standard errors taking into account both the two stages involved in the estimation and the clustered structure of our data (interviewers).[14]

We estimate different first stage regressions, without and with interviewer fixed effects, corresponding to columns (4) and (5) of Table 2 (women) and Table 3 (men). When including both BMI and BMI^2, the first stage regressions without and with fixed effects correspond to the specifications in columns (1) and (2) of Table 4, for women, and of columns (3) and (4), for men.

5.2 Beauty and its Components in the Marriage Market

Similarly, we estimate the following regressions

$$E_s = \delta^M + \rho^M A_{ij} + \mathbf{d}_i \boldsymbol{\phi}^M + \sigma^M E_i + u_i^M \tag{9}$$

$$E_s = \delta^{M'} + \rho_1^{M'} \widehat{NA}_{ij} + \mathbf{a}_i \boldsymbol{\rho}_0^{M'} + \mathbf{d}_i \boldsymbol{\phi}^{M'} + \sigma^{M'} E_i + u_i^{M'} \tag{10}$$

where E_s is education (1 if some college and above, 0 otherwise) of the spouse s of individual i.

[14] We perform 300 replications.

6 Is it non-anthropometric attractiveness, anthropometric characteristics, or both?

We first consider the benchmark case, using attractiveness to estimate the "beauty premium" in the labor and marriage markets, as it is standard in the literature (e.g., Hamermesh and Biddle, 1994). Then, we take advantage of the previous decomposition procedure and control for weight, height and the corresponding non-anthropometric (residual) attractiveness, to investigate the source of the "beauty premium". We also replicate this decomposition using BMI, instead of weight and height, and check for non-monotonicities in BMI.

6.1 Beauty Decomposition and Wages

Table 7 shows that attractiveness is positively related to wages for women. Interestingly, both the non-anthropometric attractiveness and height are positively correlated with wages in the labor market, whereas weight (or BMI) exhibits a point estimate close to zero and not statistically significant at conventional levels. We can see in column (1) that attractiveness explains 10% of the variation in female log hourly wages, while non-anthropometric attractiveness and anthropometric characteristics (column (3)) jointly explain 12% of the same variation. Column (4) indicates that a one standard deviation increase in the non-anthropometric component is associated with a 0.13 standard deviation increase in the log hourly wage rate, and a one standard deviation in height is associated with a 0.14 standard deviation increase in the log hourly wage rate. Finally, columns (7) and (8) reveal that there is no evidence of non-monotonicities in BMI. Table A3 in the Appendix shows that our results are robust to selection into working.

[Table 7 about here]

Table 8 presents the estimates of the same type of regressions as in Table 7, but for the sample of working men; interestingly enough, they exhibit a similar pattern of results, except for the fact that now BMI and wages are non-monotonically related, as it can be seen from columns (7) and (8), consistent with Caliendo and Gehrsitz (2014). Attractiveness, its component unexplained by anthropometric characteristics and height are all positively related to wages. Column (1) shows that attractiveness explains 17% of the variation in male log hourly wages, while non-anthropometric attractiveness and anthropometric characteristics (column (3)) jointly explain 19% of the same variation. Column (4) shows that a one standard deviation increase in the non-anthropometric component is associated with a 0.21 standard deviation increase in the log hourly wage rate, while a one standard deviation in height is associated with 0.12 standard deviation increase. Table A3 in the Appendix shows that our results are robust to selection into working.

[Table 8 about here]

Our findings regarding BMI match those in Garcia and Quintana-Domeque (2007), who do not find a clear penalty for heavier women in European countries and any penalty at all for men (neither Cawley (2004) does in the US). The quadratic relationship between log wages and BMI for men is consistent with the analysis using German data from the GSOEP by Caliendo and Gehrsitz (2014). The positive correlation between height and log hourly wage rates is consistent with the height "premium" documented by Case and Paxson (2008) across genders. Moreover, the fact that overall attractiveness is positively related to wages for both women and men is in line with the seminal work by Hamermesh and Biddle (1994) and the literature using beauty ratings (instead of anthropometric measures) developed after their work. For instance, Mobius and Rosenblat (2006) find a sizable beauty "premium" in wages in a lab experimental setting, while Gehrsitz (2014) reports that good looks improve labor market outcomes for both men and women, as do Doorely and Sierminska (2012) focusing on women. Finally, Scholz and Sicinski (2014) find that both facial attractiveness and height increase male earnings using data from the Wisconsin Longitudinal Survey.

6.2 Beauty Decomposition and Spousal Education

Table 9 shows that women's attractiveness is positively related to their husbands' education, and that weight and BMI contribute to the female beauty premium in the marriage market, and play a very significant role. Thinner women tend to have better educated husbands, whereas non-anthropometric beauty is significantly related to spousal quality. According to column (4), a one standard deviation in weight is associated with a reduction in this probability of 8 percentage points. Similar results are found when using BMI instead of weight controlling for height. Columns (7) and (8) show no evidence of non-monotonicities.

[Table 9 about here]

Table 10 presents the estimates of the same type of regressions as in Table 9, but for the sample of married men. Interestingly, the evidence on the role of male attractiveness and its determinants on the marriage market reveals an important gender difference: Male attractiveness (overall or non-anthropometric) is not at all related to their wives' education. Instead, being thinner and taller plays a mild advantage in securing a better quality wife. A weakly significant relationship is also found when using BMI rather than weight controlling for height. According to column (5), a one standard deviation increase in BMI is associated with a decrease of 4 percentage points in the probability of having a higher educated wife. As before, there is no evidence of non-monotonicities (columns (7) and (8)).

[Table 10 about here]

Overall, it seems that in the marriage market body size plays a stronger role than (residual) beauty for both men and women. This is a relevant finding since the literature on beauty argues that anthropometric characteristics (in particular BMI) do not matter after controlling for beauty. This is clearly not the case here, where this type of estimation is performed for the first time using nationally representative data.

Finally, one may ask whether the marital trade-off between BMI and education is the same for men and women. Under certain assumptions, the rate at which the marriage market allows an individual to trade-off own BMI and own education can be measured by the ratio of the coefficients of BMI and education (Chiappori, Oreffice and Quintana-Domeque, 2012). That is, it is possible to test whether mate preferences for two individual attributes are different between men and women. If BMI and education both significantly contribute to a mate's suitability in the marriage market, then the marginal rate of substitution between these characteristics can be exactly identified, provided that this "suitability" function is differentiable in BMI and education, and that these characteristics are (weakly) separable and (conditionally) independent of unobservable characteristics (see Chiappori, Oreffice and Quintana-Domeque (2012) for technical details).

In Table 11 we test whether the ratio of the estimated coefficients on BMI and education is different for married women and men: the test is performed after simultaneously estimating the regressions in column (5) in Tables 9 and 10. We cannot reject the null hypothesis that the ratio of these coefficients is the same for women and men.

[Table 11 about here]

7 Conclusions

We examine how attractiveness rated at the start of the interview is related to weight (controlling for height), or BMI, separately by gender and accounting for interviewer's characteristics (or fixed effects), in a nationally representative sample (German General Social Survey, 2008, 2012). We also decompose attractiveness netting out its anthropometric component and obtaining the residual non-anthropometric attractiveness, thus working with three measures of beauty to assess their "premia" (or "penalties") in the labor and marriage markets, in terms of hourly wages and spousal education.

Our paper not only highlights the importance of accounting for the gender of the rater/interviewer when "translating" anthropometric measures into attractiveness, but that body size (in particular BMI) is a relevant dimension above and beyond beauty, at least in the marriage market.

References

Averett, Susan and Sanders Korenman. 1996. “The Economic Reality of the Beauty Myth.” *Journal of Human Resources* 31 (2):304–330.

Biddle, Jeff E. and Daniel S. Hamermesh. 1998. “Beauty, Productivity, and Discrimination: Lawyers’ Looks and Lucre.” *Journal of Labor Economics* 16 (1):172–201.

Brunello, Giorgio and Béatrice d’Hombres. 2007. “Does body weight affect wages? Evidence from Europe.” *Economics and Human Biology* 5 (1):1–19.

Caliendo, Marco and Markus Gehrsitz. 2014. “Obesity and the Labor Market: A Fresh Look at the Weight Penalty.” *IZA Discussion Paper* 7947.

Cawley, John. 2004. “The Impact of Obesity on Wages.” *Journal of Human Resources* 39 (2):451–474.

Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque. 2012. “Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market.” *Journal of Political Economy* 120 (4):659–695.

Doorley, Karina and Eva Sierminska. 2012. “Myth or Fact? The Beauty Premium across the Wage Distribution.” *IZA Discussion Paper* 6674.

Gehrsitz, Markus. 2014. “Looks and Labor: Do Attractive People Work More?” *LABOUR* 28 (3):269–287.

GESIS. 2014. “GESIS - Leibniz Institute for the Social Sciences (2014).” *ALLBUS / GGSS: German General Social Survey - Cumulation 1980-2012* .

Hamermesh, Daniel S. and Jason Abrevaya. 2013. “Beauty is the promise of happiness?” *European Economic Review* 64:351–368.

Hamermesh, Daniel S. and Jeff E. Biddle. 1994. "Beauty and the Labor Market." *American Economic Review* 84 (5):1174–1194.

Hamermesh, Daniel S., Xin Meng, and Junsen Zhang. 2002. "Dress for success – does primping pay?" *Labour Economics* 9 (3):361–373.

Menold, Natalja and Cornelia Zuell. 2010. "Reasons for refusals, their collection in surveys and interviewer impact." *GESIS Working Papers* .

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

Oreffice, Sonia and Climent Quintana-Domeque. 2010. "Anthropometry and Socioeconomics among Couples: Evidence in the United States." *Economics and Human Biology* 8 (3):373–384.

———. 2012. "Fat spouses and hours of work: are body and Pareto weights correlated?" *IZA Journal of Labor Economics* 1 (6):1–21.

Robins, Philip, Jenny Homer, and Michale French. 2011. "Beauty and the Labor Market: Accounting for the Additional Effects of Personality and Grooming." *LABOUR* 25 (2):228–251.

Scholz, John Karl and Kamil Sicinski. 2014. "Facial Attractiveness and Lifetime Earnings: Evidence from a Cohort Study." *Review of Economics and Statistics* , forthcoming.

Thomas, Duncan and Elizabeth Frankenberg. 2002. "The measurement and interpretation of health in social surveys." *Summary Measures of Population Health* :387–420.

Weeden, Jason and John Sabini. 2007. "Subjective and Objective Measures of Attractiveness and Their Relation to Sexual Behavior and Sexual Attitudes in University Students." *Archives of Sexual Behavior* 36 (1):79–88.

WHO. 2009. "Global Health Risks: mortality and burden of disease attributable to selected major risks." *World Health Organization, Geneva* .

Table 1. Summary statistics. ALLBUS: 2008, 2012

Women	N	Mean	SD	Min	Max
Age	514	39.1	7.5	25	50
Height (cm)	514	167.6	5.9	150	186
Weight (kg)	514	67.5	12.5	48	120
BMI (kg/m^2)	514	24.0	4.2	18.5	38.87
Attractiveness (1-11)	514	8.2	1.8	1	11
West	514	0.82	0.39	0	1
Education (some college and above)	494	0.26	0.44	0	1
Men	**N**	**Mean**	**SD**	**Min**	**Max**
Age	561	39.4	7.4	25	50
Height (cm)	561	180.2	6.7	158	200
Weight (kg)	561	85.0	12.5	53	135
BMI (kg/m^2)	561	26.2	3.6	18.52	39.85
Attractiveness (1-11)	561	7.7	1.8	1	11
West	561	0.81	0.40	0	1
Education (some college and above)	551	0.37	0.48	0	1
All respondents	**N**	**Mean**	**SD**	**Min**	**Max**
Age	1,075	39.3	7.5	25	50
Female	1,075	0.48	0.50	0	1
All interviewers	**N**	**Mean**	**SD**	**Min**	**Max**
Age	258	59.1	8.9	27	78
Female	258	0.38	0.49	0	1

Note: We focus our analysis on German citizens born in Germany, aged 25-50 and with BMI in the range 18.5-39.99. Attractiveness is assessed by the interviewer at the start of the interview. Observations have been weighted to adjust for the oversample of East German respondents. See ALLBUS: German General Social Survey-Cumulation 1980-2012.

Table 2. LS Regressions of female attractiveness on anthropometric measures

	(1)	(2)	(3)	(4)	(5)
Panel I					
weight	−0.042***	−0.041***	−0.048***	−0.042***	−0.053***
	(0.009)	(0.009)	(0.013)	(0.012)	(0.012)
weight × female interviewer				0.004	0.011
				(0.021)	(0.027)
weight × same-age group interviewer				−0.006	0.006
				(0.021)	(0.028)
height	0.053***	0.052***	0.073***	0.057***	0.067***
	(0.017)	(0.016)	(0.020)	(0.017)	(0.022)
height × female interviewer				−0.005	−0.003
				(0.035)	(0.044)
height × same-age group interviewer				−0.021	0.038
				(0.036)	(0.063)
Clusters	215	215	215	215	215
Observations	514	514	514	514	514
R^2	0.10	0.10	0.66	0.12	0.67
Panel II					
BMI	−0.117***	−0.116***	−0.140***	−0.119***	−0.151***
	(0.027)	(0.026)	(0.036)	(0.034)	(0.035)
BMI × female interviewer				0.012	0.027
				(0.059)	(0.075)
BMI × same-age group interviewer				−0.018	−0.005
				(0.059)	(0.063)
Clusters	215	215	215	215	215
Observations	514	514	514	514	514
R^2	0.10	0.10	0.66	0.11	0.67
Interviewee controls?	Yes	Yes	Yes	Yes	Yes
Interviewer controls?	No	Yes	No	Yes	No
Interviewer fixed effects?	No	No	Yes	No	Yes
Interviewer × Interviewee interactions?	No	No	No	Yes	Yes

Note: Robust standard errors clustered at the interviewer level in parentheses.
Interviewee controls: age, region and year of the interview.
Interviewer controls: same-age group (1 if less than 50, 0 otherwise) and female (0-1).
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table 3. LS Regressions of male attractiveness on anthropometric measures

	(1)	(2)	(3)	(4)	(5)
Panel I					
weight	−0.009	−0.009	−0.014	0.000	0.001
	(0.008)	(0.008)	(0.011)	(0.010)	(0.013)
weight × female interviewer				−0.030*	−0.040**
				(0.015)	(0.020)
weight × same-age group interviewer				0.023	0.045
				(0.023)	(0.031)
height	0.047***	0.047***	0.063***	0.018	0.041
	(0.014)	(0.015)	(0.021)	(0.017)	(0.027)
height × female interviewer				0.058*	0.057
				(0.030)	(0.041)
height × same-age group interviewer				0.010	−0.052
				(0.039)	(0.047)
Clusters	211	211	211	211	211
Observations	561	561	561	561	561
R^2	0.06	0.06	0.56	0.09	0.58
Panel II					
BMI	−0.032	−0.033	−0.043	0.000	−0.007
	(0.024)	(0.024)	(0.036)	(0.027)	(0.041)
BMI × female interviewer				−0.090*	−0.106*
				(0.048)	(0.064)
BMI × same-age group interviewer				0.054	0.141
				(0.077)	(0.100)
Clusters	211	211	211	211	211
Observations	561	561	561	561	561
R^2	0.04	0.04	0.54	0.07	0.57
Interviewee controls?	Yes	Yes	Yes	Yes	Yes
Interviewer controls?	No	Yes	No	Yes	No
Interviewer fixed effects?	No	No	Yes	No	Yes
Interviewer × Interviewee interactions?	No	No	No	Yes	Yes

Note: Robust standard errors clustered at the interviewer level in parentheses.
Interviewee controls: age, region and year of the interview.
Interviewer controls: same-age group (1 if less than 50, 0 otherwise) and female (0-1).
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table 4. LS Regressions of attractiveness: non-monotonicities in BMI

	Female Attractiveness		Male Attractiveness	
	(1)	(2)	(3)	(4)
BMI	−0.392	−0.277	0.325	0.473
	(0.277)	(0.323)	(0.256)	(0.448)
BMI^2	0.506	0.233	−0.580	−0.855
	(0.509)	(0.571)	(0.443)	(0.794)
BMI × female interviewer	−0.428	−0.092	0.981**	0.992
	(0.553)	(0.704)	(0.486)	(0.719)
BMI^2 × female interviewer	0.851	0.240	−1.986**	−2.037
	(1.075)	(1.321)	(0.870)	(1.278)
BMI × same-age group interviewer	0.875	0.815	−0.599	−0.650
	(0.767)	(1.098)	(0.733)	(0.948)
BMI^2 × same-age group interviewer	−1.684	−1.538	1.255	1.492
	(1.461)	(2.047)	(1.322)	(1.702)
Clusters	215	211	215	211
Observations	514	514	561	561
R^2	0.12	0.67	0.09	0.59
Interviewee controls?	Yes	Yes	Yes	Yes
Interviewer controls?	Yes	No	Yes	No
Interviewer fixed effects?	No	Yes	No	Yes
Interviewer × Interviewee interactions?	Yes	Yes	Yes	Yes

Note: Robust standard errors clustered at the interviewer level in parentheses.
Interviewee controls: age, region and year of the interview.
Interviewer controls: same-age group (1 if less than 50, 0 otherwise) and female (0-1).
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table 5. Interviewer-Interviewee Matching Patterns by Gender

	Male interviewee	Female interviewee	Total
Male interviewer	29.30% (315)	26.14% (281)	55.44% (596)
Female interviewer	22.89% (246)	21.67% (233)	44.56% (479)
Total	52.19% (561)	47.81% (514)	100% (1,075)
Pearson's chi-squared test		$\chi^2(1)$=0.2380 p-value=0.626	

Note. Number of observations in parentheses.

Table 6. Respondents' mean characteristics by interviewer's sex

	Women			Men		
	Interviewer's sex			Interviewer's sex		
	Same	Opposite	*Difference*	Same	Opposite	*Difference*
Panel I. One-by-one estimation						
Age	39.5	38.8	0.7 (0.7)	39.8	39.0	0.8 (0.7)
Height	167.1	167.9	0.8 (0.5)	180.2	180.3	0.1 (0.6)
Weight	68.1	67.1	1.0 (1.2)	85.0	85.0	0.0 (1.1)
BMI	24.4	23.8	0.6 (0.4)	26.2	26.1	0.1 (0.3)
Attractiveness	8.08	8.25	0.17 (0.24)	7.64	7.73	0.09 (0.21)
West	0.81	0.82	0.01 (0.05)	0.78	0.83	0.05 (0.06)
Year 2012	0.48	0.50	0.02 (0.07)	0.52	0.52	0.00 (0.08)
Education (some college and above)	0.31	0.23	0.08* (0.04)	0.35	0.41	0.06 (0.05)
Panel II. Simultaneous estimation						
Adjusted Wald Test		$F_{8,207} = 1.52$ p-value = 0.1533		$F_{8,203} = 0.76$ p-value = 0.6368		

Note: The means in Panel I are obtained from individual regressions of each of the variables in the column on an interviewer's sex indicator. *Difference* is the coefficient on the interviewer's sex indicator (the difference in means between respondents interviewed by opposite-sex interviewers and those interviewed by same-sex interviewers). Panel II contains the result of the *Adjusted Wald Test* (Ho: no mean differences in any of these characteristics) after simultaneous estimation of all the previous individual regressions and its associated p-value. Observations have been weighted to adjust for the oversample of East German respondents. Standard errors clustered at the interviewer level are reported in parentheses.
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table 7. LS Regressions of female log hourly wage rate on attractiveness and other variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attractiveness	0.052***	0.040**						
	(0.018)	(0.017)						
Non-anthropometric attractiveness			0.038**	0.056*	0.039**	0.061*	0.040**	0.060*
			(0.017)	(0.033)	(0.017)	(0.032)	(0.017)	(0.033)
Weight			−0.002	−0.002				
			(0.003)	(0.003)				
Height			0.012**	0.012**				
			(0.005)	(0.005)				
BMI					−0.006	−0.006	−0.063	−0.062
					(0.008)	(0.008)	(0.064)	(0.064)
BMI^2							0.106	0.105
							(0.117)	(0.117)
Education (some college and above)		0.328***	0.314***	0.322***	0.324***	0.332***	0.320***	0.331***
		(0.057)	(0.061)	(0.063)	(0.060)	(0.062)	(0.060)	(0.063)
Clusters	163	163	163	163	163	163	163	163
Observations	309	309	309	309	309	309	309	309
First stage with interviewer FE	–	–	NO	YES	NO	YES	NO	YES
R^2	0.10	0.20	0.22	0.21	0.20	0.20	0.21	0.20

Note: All regressions include interviewee controls: age, region and year of the interview.
Non-anthropometric attractiveness in columns (3) and (4) is the residual from regression in columns (4) and (5) of Panel I in Table 2.
Non-anthropometric attractiveness in columns (5) and (6) is the residual from regression in columns (4) and (5) of Panel II in Table 2.
Non-anthropometric attractiveness in columns (7) and (8) is the residual from regression in columns (1) and (2) of Table 4.
Since these residuals are predicted in a first-stage, our standard errors (in parentheses) are bootstrapped to take into account both the two stages involved in the estimation and the clustered structure of our data (interviewers). We perform 300 replications. Observations have been weighted to adjust for the oversample of East German respondents.
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table 8. LS Regressions of male log hourly wage rate on attractiveness and other variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attractiveness	0.074***	0.065***						
	(0.014)	(0.013)						
Non-anthropometric attractiveness			0.065***	0.057**	0.065***	0.055**	0.066***	0.061***
			(0.015)	(0.022)	(0.016)	(0.023)	(0.016)	(0.023)
Weight			−0.002	−0.002				
			(0.002)	(0.002)				
Height			0.009**	0.009**				
			(0.004)	(0.004)				
BMI					−0.009	−0.009	0.155***	0.154**
					(0.007)	(0.007)	(0.071)	(0.071)
BMI^2							−0.301***	−0.298**
							(0.127)	(0.127)
Education (some college and above)		0.272***	0.262***	0.269***	0.266***	0.275***	0.254***	0.262***
		(0.051)	(0.052)	(0.054)	(0.052)	(0.054)	(0.052)	(0.054)
Clusters	185	185	185	185	185	185	185	185
Observations	415	415	415	415	415	415	415	415
First stage with interviewer FE	–	–	NO	YES	NO	YES	NO	YES
R^2	0.17	0.23	0.25	0.22	0.24	0.21	0.25	0.22

Note: All regressions include interviewee controls: age, region and year of the interview.
Non-anthropometric attractiveness in columns (3) and (4) is the residual from regression in columns (4) and (5) of Panel I in Table 3.
Non-anthropometric attractiveness in columns (5) and (6) is the residual from regression in columns (4) and (5) of Panel II in Table 3.
Non-anthropometric attractiveness in columns (7) and (8) is the residual from regression in columns (3) and (4) of Table 4.
Since these residuals are predicted in a first-stage, our standard errors (in parentheses) are bootstrapped to take into account both the two stages involved in the estimation and the clustered structure of our data (interviewers). We perform 300 replications. Observations have been weighted to adjust for the oversample of East German respondents.
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table 9. LS Regressions of spousal education on female attractiveness and other variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attractiveness	0.058***	0.040***						
	(0.016)	(0.015)						
Non-anthropometric attractiveness			0.032*	0.012	0.033*	0.019	0.033*	0.033
			(0.017)	(0.037)	(0.017)	(0.035)	(0.017)	(0.035)
Weight			−0.007***	−0.007***				
			(0.003)	(0.003)				
Height			0.009	0.009				
			(0.006)	(0.006)				
BMI					−0.019***	−0.019***	−0.045	−0.044
					(0.007)	(0.007)	(0.065)	(0.064)
BMI^2							0.048	0.047
							(0.120)	(0.120)
Education (some college and above)		0.418***	0.398***	0.410***	0.401***	0.413***	0.401***	0.408***
		(0.062)	(0.068)	(0.067)	(0.066)	(0.065)	(0.066)	(0.065)
Clusters	156	156	156	156	156	156	156	156
Observations	291	291	291	291	291	291	291	291
First stage with interviewer FE	–	–	NO	YES	NO	YES	NO	YES
R^2	0.05	0.18	0.20	0.19	0.20	0.19	0.20	0.19

Note: All regressions include interviewee controls: age, region and year of the interview.
Non-anthropometric attractiveness in columns (3) and (4) is the residual from regression in columns (4) and (5) of Panel I in Table 2.
Non-anthropometric attractiveness in columns (5) and (6) is the residual from regression in columns (4) and (5) of Panel II in Table 2.
Non-anthropometric attractiveness in columns (7) and (8) is the residual from regression in columns (1) and (2) of Table 4.
Since these residuals are predicted in a first-stage, our standard errors (in parentheses) are bootstrapped to take into account both the two stages involved in the estimation and the clustered structure of our data (interviewers). We perform 300 replications. Observations have been weighted to adjust for the oversample of East German respondents.
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table 10. LS Regressions of spousal education on male attractiveness and other variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attractiveness	0.009	−0.000						
	(0.014)	(0.013)						
Non-anthropometric attractiveness			−0.003	−0.022	0.001	−0.016	−0.000	−0.028
			(0.013)	(0.026)	(0.013)	(0.025)	(0.014)	(0.024)
Weight			−0.004	−0.004				
			(0.003)	(0.003)				
Height			0.010*	0.010*				
			(0.006)	(0.006)				
BMI					−0.013	−0.013	−0.076	−0.078
					(0.008)	(0.008)	(0.075)	(0.075)
BMI^2							0.116	0.118
							(0.134)	(0.134)
Education (some college and above)		0.278***	0.274***	0.280***	0.275***	0.281***	0.278***	0.287***
		(0.047)	(0.045)	(0.046)	(0.046)	(0.046)	(0.045)	(0.045)
Clusters	145	145	145	145	145	145	145	145
Observations	251	251	251	251	251	251	251	251
First stage with interviewer FE	–	–	NO	YES	NO	YES	NO	YES
R^2	0.01	0.13	0.15	0.16	0.14	0.14	0.15	0.15

Note: All regressions include interviewee controls: age, region and year of the interview.
Non-anthropometric attractiveness in columns (3) and (4) is the residual from regression in columns (4) and (5) of Panel I in Table 3.
Non-anthropometric attractiveness in columns (5) and (6) is the residual from regression in columns (4) and (5) of Panel II in Table 3.
Non-anthropometric attractiveness in columns (7) and (8) is the residual from regression in columns (3) and (4) of Table 4.
Since these residuals are predicted in a first-stage, our standard errors (in parentheses) are bootstrapped to take into account both the two stages involved in the estimation and the clustered structure of our data (interviewers). We perform 300 replications. Observations have been weighted to adjust for the oversample of East German respondents.
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table 11. Trade-offs between BMI and education in the marriage market

	Husband's Education	Wife's Education
Spousal Non-Anthropometric Attractiveness	0.032**	−0.007
	(0.016)	(0.013)
Spousal BMI	−0.019***	−0.013*
	(0.007)	(0.008)
Spousal Education (some college and above)	0.400***	0.278***
	(0.063)	(0.047)
Ratio of Coefficients		
$\frac{\beta_{BMI}}{\beta_{Education}}$	−0.049**	−0.047
	(0.022)	(0.029)
Test of Equality		
Wald Test	$\chi^2(1) = 0.00$	
	p-value=0.9718	
Clusters	208	
Observations	542	

Note: All regressions include interviewee controls: age, region and year of the interview.
Spousal non-anthropometric attractiveness in husband's education column is the residual from regression in column (4) of Panel II in Table 2.
Spousal non-anthropometric attractiveness in wife's education column is the residual from regression in column (4) of Panel II in Table 3.
Simultaneous estimation (standard errors) accounting for weighting and clustering.
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

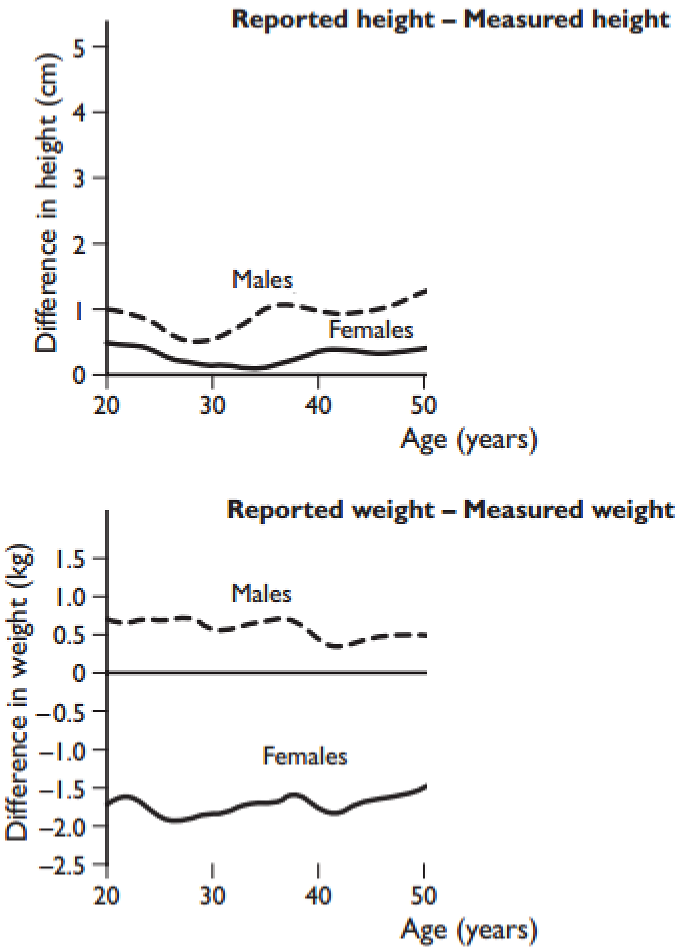

Figure 1: Height and weight of adults 20-50 years old: difference between self-reports and measurements by age (NHANES III). Source: Based on Figure I in Thomas and Frankenberg (2002).

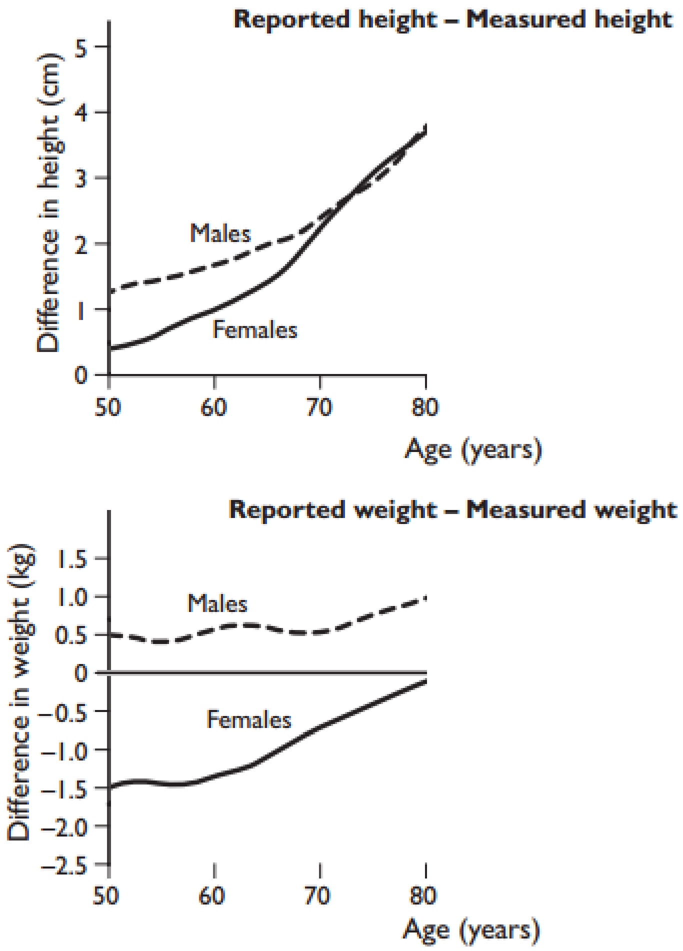

Figure 2: Height and weight of adults 50-80 years old: difference between self-reports and measurements by age (NHANES III). Source: Based on Figure I in Thomas and Frankenberg (2002).

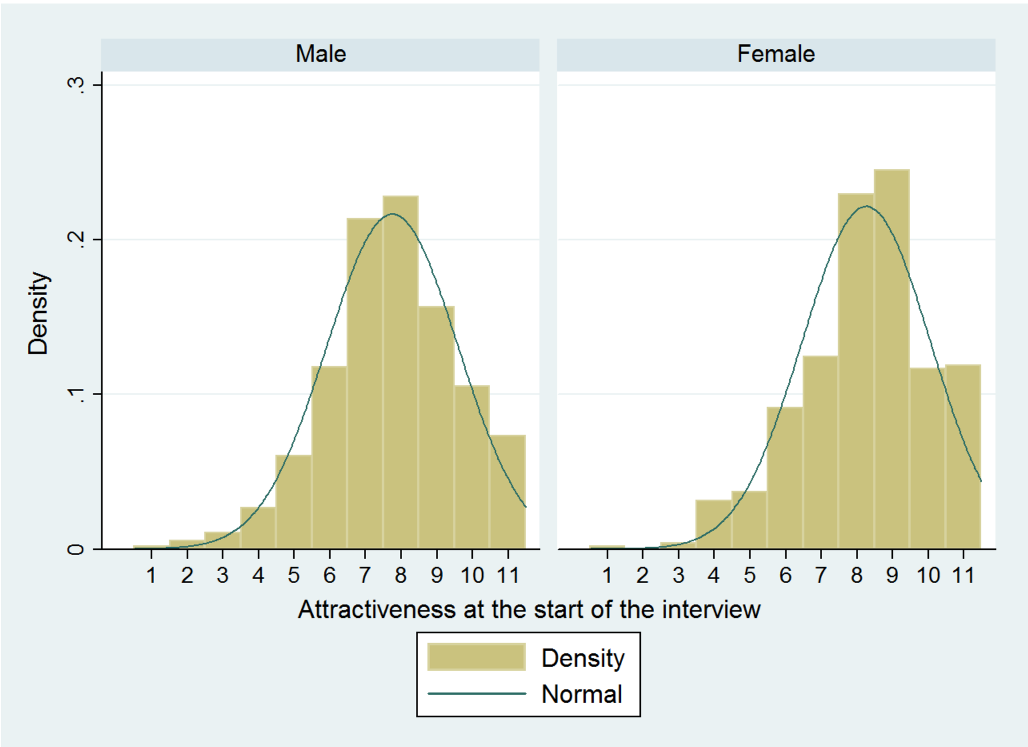

Figure 3: Histograms of attractiveness distributions for men and women.

APPENDIX

Table A1: Ordered Probit (OPROBIT) vs Least Squares (LS) regressions. Dependent variable: Female attractiveness

	OPROBIT	LS	OPROBIT	LS
BMI	−0.066***	−0.116***	−0.071***	−0.119***
	(0.017)	(0.026)	(0.021)	(0.034)
BMI × female interviewer			0.012	0.012
			(0.034)	(0.059)
BMI × same-age group interviewer			−0.008	−0.018
			(0.035)	(0.059)
Observations	514	514	514	514
Interviewee controls?	Yes	Yes	Yes	Yes
Interviewer controls?	Yes	Yes	Yes	Yes
Interviewer × Interviewee interactions?	No	No	Yes	Yes

Note: Robust standard errors clustered at the interviewer level (215) in parentheses.
Interviewee controls: age, region and year of the interview.
Interviewer controls: same-age group (1 if less than 50, 0 otherwise) and female (0-1).
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table A2: Ordered Probit (OPROBIT) vs Least Squares (LS) regressions
Dependent variable: Male attractiveness

	OPROBIT	LS	OPROBIT	LS
BMI	−0.018	−0.033	0.000	0.000
	(0.014)	(0.024)	(0.016)	(0.027)
BMI × female interviewer			−0.050*	−0.090*
			(0.034)	(0.028)
BMI × same-age group interviewer			0.024	0.054
			(0.044)	(0.077)
Observations	561	561	561	561
Interviewee controls?	Yes	Yes	Yes	Yes
Interviewer controls?	Yes	Yes	Yes	Yes
Interviewer × Interviewee interactions?	No	No	Yes	Yes

Note: Robust standard errors clustered at the interviewer level (211) in parentheses.
Interviewee controls: age, region and year of the interview.
Interviewer controls: same-age group (1 if less than 50, 0 otherwise) and female (0-1).
*** p-value<0.01, ** p-value<0.05, * p-value<0.1

Table A3: Wage Equations accounting for selection. Heckman's full MLE. Dependent variable: Log(Hourly Wage Rate)

	Female		**Male**	
Non-Anthropometric Attractiveness	0.039**	0.041**	0.070***	0.071***
	(0.017)	(0.017)	(0.015)	(0.015)
BMI	−0.007	−0.074	−0.012	0.189***
	(0.007)	(0.063)	(0.007)	(0.069)
BMI^2		0.125		−0.364***
		(0.114)		(0.125)
Education (some college and above)	0.336***	0.331***	0.292***	0.283***
	(0.056)	(0.055)	(0.054)	(0.054)
Clusters	212	212	210	210
Observations	494	494	551	551
Censored	185	185	136	136
Log Pseudo-Likelihood	−475.54	−474.17	−532.57	−529.37

Note: Robust standard errors clustered at the interviewer level in parentheses.
Controls: age, region and year of the interview.
The same explanatory variables are used in both the wage and the selection equation.
*** p-value<0.01, ** p-value<0.05, * p-value<0.1