

Beauty, Body Size and Wages: Evidence from a Unique Data Set*

Sonia Oreffice
University of Surrey and IZA

Climent Quintana-Domeque
University of Oxford and IZA

November 2015

Abstract

We analyze how attractiveness rated at the start of the interview in a nationally representative sample from Germany is related to weight, height, and body mass index (BMI), separately by gender and accounting for interviewers' characteristics or fixed effects. 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. We also estimate whether, controlling for beauty, body size measures are related to hourly wages. We find that anthropometric attributes play a significant role in wage regressions *in addition to* attractiveness, showing that body size cannot be dismissed as a simple component of beauty. Our findings are robust to controlling for health status and accounting for selection into working.

JEL Classification Codes: J01, J10.

Keywords: Attractiveness, Body Mass Index, Height, Weight, Wages.

*Quintana-Domeque (corresponding author): University of Oxford, Department of Economics, Oxford, OX1 3UQ, United Kingdom; climent.quintana-domeque@economics.ox.ac.uk. We thank Damian Clarke, Tel Enyecu, Marco Gonzalez-Navarro, David Margolis, Daniele Paserman, Simon Quinn, Curtis Simon and seminar participants at Alacant, IFS, Oxford, PSE and TSE for useful comments. The usual disclaimers apply.

1 Introduction

The salience of physical attributes to economic behavior 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 socioeconomic outcomes, from schooling attainment and wages to crime.² In this paper, we examine how anthropometric characteristics are related to beauty, and investigate the relationship of attractiveness, weight, height and BMI with wages, with a unique data set.

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?

We then 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 hourly wages. As such, we reassess one of the open questions in the social sciences regarding beauty and socioeconomic outcomes: Is it beauty, anthropometric measures or both that matter(s) for wages?

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,

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

²For example, Averett and Korenman (1996), Biddle and Hamermesh (1998), Caliendo and Gehrsitz (2014), Cawley (2004), Mocan and Tekin (2010) and Scholz and Sicinski (2015).

or whether the observed anthropometric “premium” (or “penalty”) in the labor market simply reflects a “beauty premium”, or vice versa. 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 estimate 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. Our findings are robust to controlling for health status. In addition, 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.

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 investigate the presence of both beauty and

³In anthropology or psychology studies typically work with non-representative samples. For instance, 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.

anthropometric “premia” in the labor market by means of wage regressions. 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. Moreover, male BMI is non-monotonically related to wages, consistent with Caliendo and Gehrsitz (2014). The height “premium” may reflect the fact that adult stature is positively correlated with cognitive ability (Case and Paxson, 2008), while the non-monotonic relationship of BMI with wages for men is consistent with BMI not being able to distinguish fat from muscle (Burkhauser and Cawley, 2008; Tekin and Wada, 2010). Body size –height for both men and women, BMI for men only– explains wages above and beyond beauty, even when controlling for health status and accounting for selection into working.

Our first finding should prompt 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 statement: we show that these standards and their anthropometric determinants may differ by *gender*. Moreover, our paper contributes to bridge the gap between the literature on the economics of anthropometric measures (including height and BMI), on one hand, and the economics of beauty, on the other, estimating the relevance of body size and beauty in a nationally representative sample. It seems that body size cannot be dismissed as a component of beauty since it significantly explains hourly wages beyond attractiveness.

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 our analysis of the determinants of beauty. Section 4 investigates whether body size matters above and beyond beauty in explaining wages, accounting for differences in health status and selection into working. Section 5 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. In addition, the interviewer’s identifier and main demographic characteristics (age and sex) are also recorded, which will 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 cumulation 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, health status (1 if health is self-assessed as good or very good, 0 otherwise), number of children, 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

⁴Interviews 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).

⁵In 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).

features and dual paths of the German education system. We construct a binary variable 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 Finally, we use 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).

The main analysis considers men and women who are German citizens born in Germany, between 25 and 50 years of age (i.e., prime-age workers) and whose BMI is in the range 18.5 to 39.99, to exclude (medically) morbid obese or underweight individuals (WHO, 2009).⁷ 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 (2015) state that “high quality data on beauty, augmented with household economic and demographic characteristics, are rare.”

In the next subsections we survey the most popular measures of beauty and the economic

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

⁷This last requirement excludes 1.4% of observations with $BMI \geq 40$ and 1.8% of observations with $BMI < 18.5$.

studies that use them, to emphasize the challenges that this literature faces as well as to highlight the 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. 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; Deryugina and Shurchkov, 2015; Hitsch, Hortaçsu and Ariely, 2010; Mobius and Rosenblat, 2006; Scholz and Sicinski, 2015). For instance, Biddle and Hamermesh (1998) consider matriculation photographs of law school students, Deryugina and Shurchkov (2015) retrieve those from student ID cards in a women’s college, while Scholz and Sicinski (2015) 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 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 a key feature of our analysis, which complements those using “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 priming and beauty of working women from Shanghai using *end*-of-interview ratings, while Gehrsitz (2014) considers start-of-interview measures and labor supply. Although a “full-body” measure of attractiveness may capture other characteristics, such as grooming and fine clothes, in principle this holds true of most facial measures of attractiveness: faces can (if anything) be “modified” with make-up, contact-lenses, earrings, hairstyle, etc. Finally, while the ALLBUS survey interviews individuals in their own homes, including dummy variables for the type of building respondents live in to capture “halo” effects does not affect our results.

2.2.3 Anthropometric characteristics and attractiveness

A relevant strand of literature uses anthropometric measures such as height, weight, or BMI to assess their correlation with (or sometimes “impact” on) labor or marriage market outcomes. The few studies on the marriage market that consider anthropometric measures as mate attributes, directly use BMI (or obesity) as a proxy for physical attractiveness (Averett and Korenman, 1996; Chiappori, Oreffice and Quintana-Domeque, 2012; Oreffice and Quintana-Domeque, 2010). 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; Cawley, Grabka and Lillard, 2005; Garcia and Quintana-Domeque, 2007).

Research bringing together measures of some anthropometric attributes and attractiveness include Scholz and Sicinski (2015), Hitsch, Hortaçsu and Ariely (2010) and Hamermesh and Biddle (1994). Scholz and Sicinski (2015) 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 *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

⁸If 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.

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

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 them lives in the former West Germany. Women report being on average less educated than men, while men exhibit a higher average BMI than women, with the same average good health status. Men are slightly overweight (with an average BMI of 26.2), 14% of them are obese (11% of females are), and their mean attractiveness rating is 7.7, while women score 8.2 on average, although the standard deviations of their ratings 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, 2015; 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) and 38% of them 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 male and female respondents.

3 The Determinants of Beauty

3.1 A Linear Regression Model

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}). A linear regression model can then be written as follows

$$A_{ij} = \alpha_j + \mathbf{x}_i\boldsymbol{\beta} + \mathbf{x}_i\mathbf{x}_j\boldsymbol{\theta} + \epsilon_{ij} \quad (1)$$

where α_j captures interviewer fixed effects.

We will estimate regression (1) 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.¹⁰ 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

$$A_{ij} = \alpha_j + \mathbf{a}_i\boldsymbol{\beta}_0 + \mathbf{a}_i\mathbf{d}_j\boldsymbol{\theta}_0 + \mathbf{d}_i\boldsymbol{\beta}_1 + \mathbf{d}_i\mathbf{d}_j\boldsymbol{\theta}_1 + \epsilon_{ij} \quad (2)$$

¹⁰To the best of our knowledge no study has interviewers' anthropometric characteristics in a representative sample.

where θ_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 (BMI²). Regressions are weighted to adjust for the over-sampling of East German respondents.

3.2 How do Anthropometric Measures correlate with Beauty?

Table 2 displays a series of least squares 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.¹¹ 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 the 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.

¹¹Heteroskedasticity robust standard errors clustered at the interviewer’s level are used in all the empirical analysis.

[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.

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 negatively assessed by 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 (OP) against least squares (LS) estimates: no differences are detected, meaning that treating attractiveness as a cardinal variable (LS) or as an ordinal one (OP) is immaterial for our conclusions. In addition, controlling for health status does not alter our findings either (results available upon request).

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. 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).

[Table 4 about here]

[Table 5 about here]

[Table 6 about here]

4 Does Body Size matter above and beyond Beauty?

4.1 Wage Regressions with Attractiveness and Body Size

We first consider the benchmark case, using attractiveness to estimate the “beauty premium” in the labor market, as it is standard in the literature (e.g., Hamermesh and Biddle, 1994). Then, we include the body size measures of weight and height, or BMI (and BMI squared), as additional variables, to test whether anthropometric characteristics play a significant role in explaining wages in addition to an individual’s beauty. Table 7 shows that attractiveness is positively related to wages for working men and women. We can see in columns (1) and (5) that (conditional) attractiveness explains 19% (23%) of the variation in female (male) log hourly wages. Adding weight and height, columns (2) and (6) reveal that both attractiveness and height are positively related with wages for both men and women. When replacing weight and height with BMI (or BMI and its square), we find that BMI and wages are not statistically related for women, but they follow an inverted U-shape relationship for men. One of the key messages of this table is that the inclusion of body size measures increases the explanatory power of wage regressions, above and beyond beauty: the adjusted R^2 s are higher in column (2) than (1), and in column (8) than (5).

[Table 7 about here]

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 afterwards. For instance, Mobius and Rosenblat (2006) report 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 (2015) find that both facial attractiveness and height increase male earnings using data from the Wisconsin Longitudinal Survey.

The height “premium” is consistent with adult stature being positively correlated with cognitive ability (Case and Paxson, 2008). Our findings regarding female BMI match those in Garcia and Quintana-Domeque (2007), who do not find a clear penalty for heavier women in European countries. 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).

While the estimates from Table 7 suggest that body size matters above and beyond attractiveness, these are obtained without taking into account selection into working and omitting the role of differences in health status. To overcome these limitations, Table 8 includes health status and accounts for selection into working using the standard Heckman selection model, where the selection equation (probit) includes the same explanatory variables as the wage equation plus the number of children. Reassuringly, the results displayed in Table 8 are virtually the same as those reported in Table 7. The fact that controlling for health status does not affect our findings is consistent with Gregory and Ruhm (2011), who show that the relationship between BMI and earnings is not driven by health expenditures.

[Table 8 about here]

4.2 Wage Regressions with Non-anthropometric Attractiveness and Body Size

In the same spirit as above, we now conduct an additional comparison between the role of beauty and body size measures in explaining wages, using the non-anthropometric residual attractiveness instead of the full beauty measure. We use our previous estimates in Tables 2 and 3 to predict attractiveness \hat{A}_{ij} , constructing a measure of non-anthropometric

attractiveness as the (predicted) residual from our estimated attractiveness regression

$$\widehat{NA}_{ij} = A_{ij} - \widehat{A}_{ij} \quad (3)$$

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). As such, we can now estimate the role of attractiveness, and specifically of both the anthropometric characteristics (either weight and height or BMI) and the non-anthropometric attractiveness, in explaining hourly wages. 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 (performing 300 replications) taking into account the two stages involved in the estimation and the clustered structure of our data (interviewers). 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 BMI², 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.

Table 9 shows that both the non-anthropometric attractiveness –the part of the attractiveness score uncorrelated to body size and interviewer fixed effects– and height are positively correlated with wages for both women (columns 1-6) and men (columns 7-12), and with similar magnitudes across genders. Column (1) 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. Column (8) 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. In addition, BMI is relevant for men, in a non-monotonic way, but not for women.

[Table 9 about here]

While the role of height and BMI can be assessed already from Tables 7 and 8, the key insight of Table 9 is that non-anthropometric attractiveness plays a similar quantitative role for men and women, once interviewer fixed effects are accounted for.

5 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, the German General Social Survey of 2008 and 2012. We also work with three measures of beauty to assess the sources of the associated "premia" in the labor market, in terms of hourly wages.

Our analysis highlights the importance of accounting for the gender of the rater/interviewer when "translating" anthropometric measures into attractiveness. Furthermore, it shows that body size measures cannot be dismissed as simply reflecting beauty, since they are estimated to be relevant dimensions in explaining wages above and beyond attractiveness, even after controlling for health status and accounting for selection into working.

Perhaps, the main point of our paper is to provide a first clear attempt to bridge the gap between two literatures, the economics of beauty and the economics of anthropometric measures, which so far have been developing in parallel.

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.
- Burkhauser, Richard and John Cawley. 2008. "Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research." *Journal of Health Economics* 27 (2):519–529.
- Caliendo, Marco and Markus Gehrsitz. 2014. "Obesity and the Labor Market: A Fresh Look at the Weight Penalty." *IZA Discussion Paper* 7947.
- Case, Anne and Christina Paxson. 2008. "Stature and Status: Height, Ability, and Labor Market Outcomes." *Journal of Political Economy* 116 (3):499–532.
- Cawley, John. 2004. "The Impact of Obesity on Wages." *Journal of Human Resources* 39 (2):451–474.
- Cawley, John, Markus Grabka, and Dean Lillard. 2005. "A Comparison of the Relationship Between Obesity and Earnings in the U.S. and Germany." *Journal of Applied Social Science Studies (Schmollers Jahrbuch)* 125 (1):119–129.
- 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.
- Deryugina, Tatyana and Olga Shurchkov. 2015. "Does Beauty Matter in Undergraduate Education?" *Economic Inquiry* 53 (2):940–961.

- Doorley, Karina and Eva Sierminska. 2015. "Myth or Fact? The Beauty Premium across the Wage Distribution." *Economics Letters* 129:29–34.
- 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* .
- Gregory, Christian A. and Christopher J. Ruhm. 2011. "Where Does the Wage Penalty Bite?" In *Economic Aspects of Obesity*, edited by Michael Grossman and Naci Mocan. Chicago: The University of Chicago Press.
- 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 priming pay?" *Labour Economics* 9 (3):361–373.
- Hitsch, Gunter J., Ali Hortaçsu, and Dan Ariely. 2010. "Matching and Sorting in Online Dating." *American Economic Review* 100 (1):130–63.
- 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.
- Mocan, Naci and Erdal Tekin. 2010. "Ugly Criminals." *Review of Economics and Statistics* 92 (1):15–30.

- OECD. 2014. "OECD Health Statistics 2014." *OECD, Paris* .
- 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.
- Scholz, John Karl and Kamil Sicinski. 2015. "Facial Attractiveness and Lifetime Earnings: Evidence from a Cohort Study." *Review of Economics and Statistics* 97 (1):14–28.
- Tekin, Erdal and Roy Wada. 2010. "Body Composition and Wages." *Economics and Human Biology* 8 (2):242–254.
- 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					
<hr/>					
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
Obese (BMI \geq 30)	514	0.11	0.31	0	1
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
Good Health	514	0.68	0.47	0	1
<hr/>					
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
Obese (BMI \geq 30)	561	0.14	0.35	0	1
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
Good Health	561	0.68	0.47	0	1
<hr/>					
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
<hr/>					
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 (1-11) on anthropometric measures

	(1)	(2)	(3)	(4)	(5)
Panel I					
weight	-0.042*** (0.009)	-0.041*** (0.009)	-0.048*** (0.013)	-0.042*** (0.012)	-0.053*** (0.012)
weight × female interviewer				0.004 (0.021)	0.011 (0.027)
weight × same-age group interviewer				-0.006 (0.021)	0.006 (0.028)
height	0.053*** (0.017)	0.052*** (0.016)	0.073*** (0.020)	0.057*** (0.017)	0.067*** (0.022)
height × female interviewer				-0.005 (0.035)	-0.003 (0.044)
height × same-age group interviewer				-0.021 (0.036)	0.038 (0.063)
Observations	514	514	514	514	514
R ²	0.10	0.10	0.66	0.12	0.67
Panel II					
BMI	-0.117*** (0.027)	-0.116*** (0.026)	-0.140*** (0.036)	-0.119*** (0.034)	-0.151*** (0.035)
BMI × female interviewer				0.012 (0.059)	0.027 (0.075)
BMI × same-age group interviewer				-0.018 (0.059)	-0.005 (0.063)
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 the interviewer's age is 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 (1-11) on anthropometric measures

	(1)	(2)	(3)	(4)	(5)
Panel I					
weight	-0.009 (0.008)	-0.009 (0.008)	-0.014 (0.011)	0.000 (0.010)	0.001 (0.013)
weight × female interviewer				-0.030* (0.015)	-0.040** (0.020)
weight × same-age group interviewer				0.023 (0.023)	0.045 (0.031)
height	0.047*** (0.014)	0.047*** (0.015)	0.063*** (0.021)	0.018 (0.017)	0.041 (0.027)
height × female interviewer				0.058* (0.030)	0.057 (0.041)
height × same-age group interviewer				0.010 (0.039)	-0.052 (0.047)
Observations	561	561	561	561	561
R ²	0.06	0.06	0.56	0.09	0.58
Panel II					
BMI	-0.032 (0.024)	-0.033 (0.024)	-0.043 (0.036)	0.000 (0.027)	-0.007 (0.041)
BMI × female interviewer				-0.090* (0.048)	-0.106* (0.064)
BMI × same-age group interviewer				0.054 (0.077)	0.141 (0.100)
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 the interviewer's age is 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 (1-11): non-monotonicities in BMI

	Female Attractiveness		Male Attractiveness	
	(1)	(2)	(3)	(4)
BMI	-0.392 (0.277)	-0.277 (0.323)	0.325 (0.256)	0.473 (0.448)
BMI ²	0.506 (0.509)	0.233 (0.571)	-0.580 (0.443)	-0.855 (0.794)
BMI × female interviewer	-0.428 (0.553)	-0.092 (0.704)	0.981** (0.486)	0.992 (0.719)
BMI ² × female interviewer	0.851 (1.075)	0.240 (1.321)	-1.986** (0.870)	-2.037 (1.278)
BMI × same-age group interviewer	0.875 (0.767)	0.815 (1.098)	-0.599 (0.733)	-0.650 (0.948)
BMI ² × same-age group interviewer	-1.684 (1.461)	-1.538 (2.047)	1.255 (1.322)	1.492 (1.702)
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 Same	Opposite	<i>Difference</i>	Interviewer's sex Same	Opposite	<i>Difference</i>
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* (H_0 : 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 log hourly wage rate on attractiveness (1-11) and other variables

	(1)	Females		(4)	(5)	Males		(8)
		(2)	(3)			(6)	(7)	
Attractiveness	0.040** (0.017)	0.038** (0.016)	0.039** (0.016)	0.039** (0.017)	0.065*** (0.013)	0.061*** (0.013)	0.064*** (0.014)	0.063*** (0.014)
Weight		-0.000 (0.002)				-0.002 (0.002)		
Height		0.010** (0.005)				0.006* (0.003)		
BMI			-0.002 (0.007)	-0.051 (0.062)			-0.007 (0.007)	0.136* (0.070)
BMI ²				0.091 (0.114)				-0.261** (0.128)
Education (some college and above)	0.328*** (0.057)	0.316*** (0.060)	0.326*** (0.058)	0.324*** (0.059)	0.272*** (0.051)	0.265*** (0.051)	0.267*** (0.050)	0.258*** (0.051)
Observations	309	309	309	309	415	415	415	415
Adjusted R ²	0.19	0.20	0.19	0.19	0.23	0.23	0.23	0.24

Note: All regressions include interviewee controls: age, region and year of the interview. Standard errors (in parentheses) are clustered at the interviewer level. 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. Wage equations controlling for health status and accounting for selection. Heckman's full MLE.
Dependent variable: Log(Hourly Wage Rate)

	Females				Males			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Attractiveness	0.039* (0.022)	0.037** (0.017)	0.028 (0.018)	0.038** (0.019)	0.057*** (0.014)	0.060*** (0.014)	0.057*** (0.014)	0.056*** (0.014)
Weight		-0.000 (0.002)				-0.001 (0.002)		
Height		0.011** (0.005)				0.008** (0.004)		
BMI			-0.002 (0.007)	-0.049 (0.068)			-0.002 (0.007)	0.129* (0.077)
BMI ²				0.088 (0.125)				-0.238* (0.141)
Education (some college and above)	0.319*** (0.112)	0.315*** (0.071)	0.252*** (0.070)	0.318*** (0.082)	0.217*** (0.056)	0.271*** (0.053)	0.215*** (0.055)	0.209*** (0.055)
Good Health	0.014 (0.098)	0.018 (0.068)	-0.040 (0.073)	0.012 (0.078)	0.109* (0.056)	0.182*** (0.048)	0.105*** (0.056)	0.110*** (0.055)
Observations	491	491	491	491	543	543	543	543
Censored	182	182	182	182	134	134	134	134
Log Pseudo-Likelihood	-460.49	-457.39	-458.70	-459.36	-515.82	-514.44	-514.01	-511.52

Note: Wage regressions include interviewee controls: age, region and year of the interview.

The selection equation includes the same explanatory variables as the wage equation plus number of children.

Standard errors (in parentheses) are clustered at the interviewer level. 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 log hourly wage rate on attractiveness (1-11) and other variables

	Females						Males					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Non-anthropometric attractiveness	0.038** (0.017)	0.056* (0.033)	0.039** (0.017)	0.061* (0.032)	0.040** (0.017)	0.060* (0.033)	0.065*** (0.015)	0.057** (0.022)	0.065*** (0.016)	0.055** (0.023)	0.066*** (0.016)	0.061*** (0.023)
Weight	-0.002 (0.003)	-0.002 (0.003)					-0.002 (0.002)	-0.002 (0.002)				
Height	0.012** (0.005)	0.012** (0.005)					0.009** (0.004)	0.009** (0.004)				
BMI			-0.006 (0.008)	-0.006 (0.008)	-0.063 (0.064)	-0.062 (0.064)			-0.009 (0.007)	-0.009 (0.007)	0.155*** (0.071)	0.154** (0.071)
BMI ²					0.106 (0.117)	0.105 (0.117)					-0.301*** (0.127)	-0.298** (0.127)
Education (some college and above)	0.314*** (0.061)	0.322*** (0.063)	0.324*** (0.060)	0.332*** (0.062)	0.320*** (0.060)	0.331*** (0.063)	0.262*** (0.052)	0.269*** (0.054)	0.266*** (0.052)	0.275*** (0.054)	0.254*** (0.052)	0.262*** (0.054)
Observations	309	309	309	309	309	309	415	415	415	415	415	415
First stage with interviewer FE	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES

Note: All regressions include interviewee controls: age, region and year of the interview.

Non-anthropometric attractiveness in columns (1) and (2) is the residual from regression in columns (4) and (5) of Panel I in Table 2.

Non-anthropometric attractiveness in columns (3) and (4) is the residual from regression in columns (4) and (5) of Panel II in Table 2.

Non-anthropometric attractiveness in columns (5) and (6) is the residual from regression in columns (1) and (2) of Table 4.

Non-anthropometric attractiveness in columns (7) and (8) is the residual from regression in columns (4) and (5) of Panel I in Table 3.

Non-anthropometric attractiveness in columns (9) and (10) is the residual from regression in columns (4) and (5) of Panel II in Table 3.

Non-anthropometric attractiveness in columns (11) and (12) 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

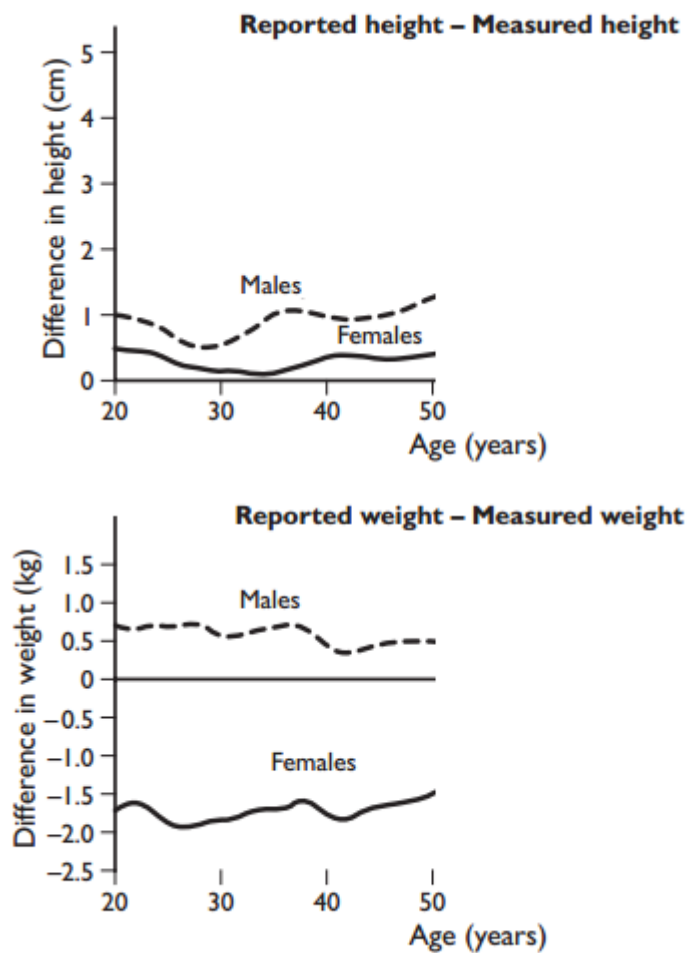


Figure 1: Height and weight of adults 20-50 years old: difference between self-reports and measurements by age (NHANES III). Source: Based on Figure I in Thomas and Frankenberg (2002).

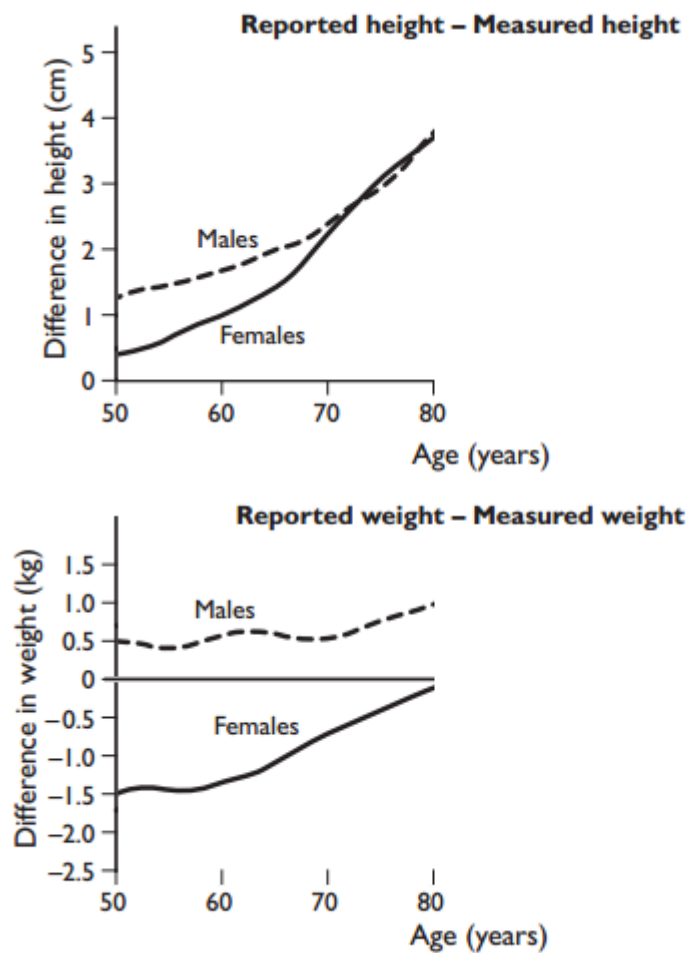


Figure 2: Height and weight of adults 50-80 years old: difference between self-reports and measurements by age (NHANES III). Source: Based on Figure I in Thomas and Frankenberg (2002).

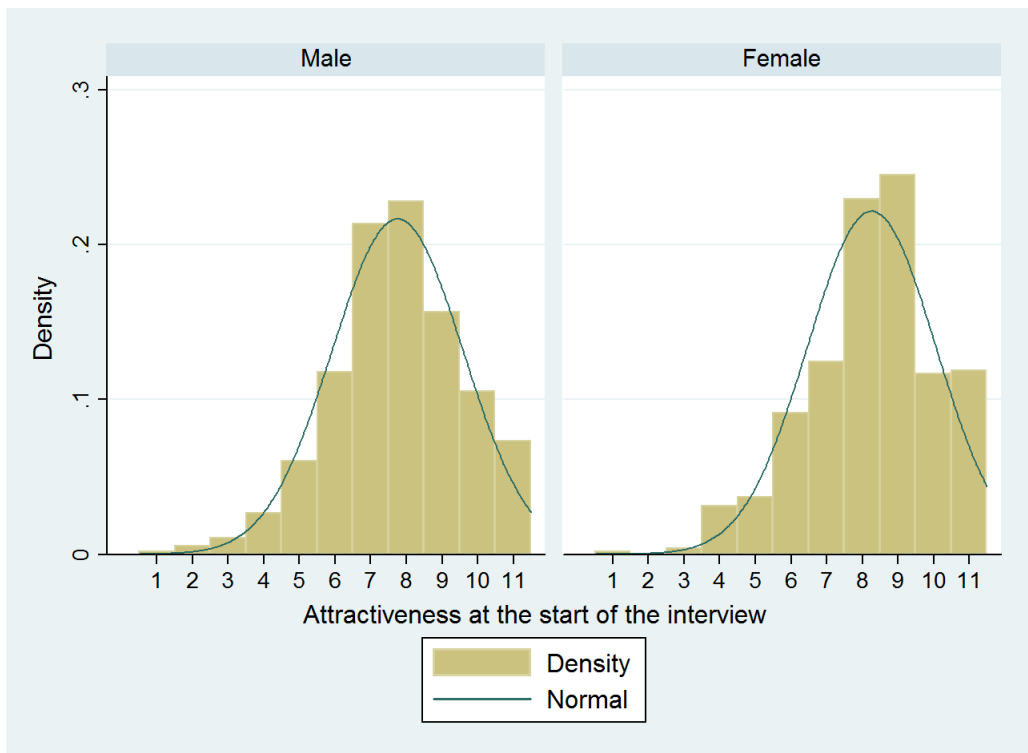


Figure 3: Histograms of attractiveness distributions for men and women.

APPENDIX

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

	OPROBIT	LS	OPROBIT	LS
BMI	−0.066*** (0.017)	−0.116*** (0.026)	−0.071*** (0.021)	−0.119*** (0.034)
BMI × female interviewer			0.012 (0.034)	0.012 (0.059)
BMI × same-age group interviewer			−0.008 (0.035)	−0.018 (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 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 (0.014)	-0.033 (0.024)	0.000 (0.016)	0.000 (0.027)
BMI × female interviewer			-0.050* (0.034)	-0.090* (0.028)
BMI × same-age group interviewer			0.024 (0.044)	0.054 (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 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