VERY PRELIMINARY
December 2010

"BEAUTY IS THE PROMISE OF HAPPINESS"*?

Jason Abrevaya and Daniel S. Hamermesh**

*Stendhal (Marie-Henri Beyle), *De l'Amour*, Ch. 17.

**Professor of Economics, University of Texas at Austin; Sue Killam Professor in the Foundations of Economics, University of Texas at Austin, and professor of labor economics, Maastricht University, IZA and NBER.

ABSTRACT

We measure the impact of individuals' looks on their life satisfaction or happiness. Using four data sets, the Quality of American Life surveys from the 1970s, the British National Child Development Study, and the Wisconsin Longitudinal Survey, we construct the beauty measures in different ways that allow putting a lower bound on the true effects of beauty. Personal beauty raises happiness, with a one standard-deviation change in beauty generating about 0.07 standard deviations of additional satisfaction/happiness among men, 0.09 among women. Accounting for covariates that might be affected by differences in beauty, particularly effects in the labor and marriage markets, these impacts are at least halved. At least the majority of the effect of beauty is through economic outcomes, broadly defined.

I. Introduction

While economists have studied happiness for several generations (Easterlin, 1973; Scitovsky, 1976), it seems fair to say that our interest in it has burgeoned in the last 15 years. The Frey and Stutzer (2002) survey captured part of the literature, but there has been a continuing outpouring of research on happiness from an economic viewpoint (e.g., Clark *et al*, 2008; Stevenson and Wolfers, 2008; Deaton and Kahneman, 2010; Oswald and Wu, 2011). Much of the analysis focuses on measuring the short- and long-run effects of changes in income on happiness, but the relation of happiness to other outcomes that are at least partly economically determined (divorce, fertility and others) has also been subject to discussion.

At the same time a smaller, but also burgeoning literature on the effects of beauty on various outcomes has been created (e.g., Hamermesh and Biddle, 1994, Möbius and Rosenblat, 2006, Mocan and Tekin, 2010). In these studies the economic focus is on such topics as how beauty is traded for income, how it alters occupational choice, and how it affects marital bargaining. The general issue is how human beauty determines outcomes in various markets and shifts the distribution of surpluses in those markets among participants.

Economists have not put these two topics together—have not examined the relationship between happiness and beauty. Some psychologists have correlated subjects' happiness and their self-assessed beauty, but that approach seems flawed. Others have examined simple averages of several measures of happiness among a random sample of people whose beauty was rated differently by interviewers (using one of the data sets we use, Umberson and Hughes, 1987), of college students whose photographs and happiness were examined (Diener *et al*, 1995), and of partial correlations of happiness measures and survey respondents' waist-to-hip ratio, used as a (to us quite unsatisfactory) proxy for beauty (Plaut *et al*, 2009). Here we put the two literatures together, examining how happiness, measured in various ways, is affected by beauty and, as important, measuring the extent to which any effects of beauty on happiness occur indirectly through its effects in labor, marriage and other markets.

We need to be as certain as possible that our analysis does not merely reflect the idiosyncrasies of measuring the subjective concepts of happiness and human beauty. For that reason we use four different surveys, representing three different approaches to measuring human beauty. All four surveys offer different measures of life satisfaction or happiness. Happiness is obviously self-rated by the respondents, but none of the beauty measures is—we are not relating a person's subjective assessment of one aspect of life to his/her assessment of another (Hamermesh, 2004). Our results' validity will depend on their robustness to differing approaches to measuring beauty and to eliciting people's expressions of satisfaction/happiness.

II. Data Sources and Descriptive Statistics

The four data sets that we use are especially diverse in terms of their methods of assessing beauty. The first two are the Quality of American Life (QAL) surveys, undertaken in 1971 and 1978 as random samples of the U.S. population, with the purpose of obtaining information on respondents' subjective views of a large variety of aspects of life. At the end of the interview in each of these surveys, the interviewer assessed the interviewee's looks on a five-to-one scale, with 5 being strikingly handsome or beautiful, 1 being homely. The complete list of descriptions associated with each rating of beauty is shown in the first column of the first panel of Table 1. This measure has been used in a variety of studies linking beauty to economic outcomes (e.g., Hamermesh and Biddle, 1994; Leigh and Susilo, 2009), although typically with the top two categories combined into a category "good looks" and the bottom two combined into "bad looks," because of the paucity of respondents rated 5 or 1.

In all of the analyses we use whatever measures of satisfaction/happiness are available. The analyses are thus data-driven, so that we are not inquiring into the various aspects of satisfaction/happiness that have been identified by psychologists (e.g., Seligman, 2004). Thus the QAL surveys provide the same measures of happiness, each on a three-to-one scale, as the description in column (3) of Table 1 shows. The surveys also provide direct measures of life satisfaction, focused on the current moment and on the person's total experience, measures that are

standard in the life satisfaction literature. Henceforth we distinguish in these data sets and one of the others between the determinants of life satisfaction and those of happiness.

One might argue that beauty, assessed as it is at the same time that the respondent's happiness/life satisfaction is elicited, might be reflecting reverse causation. As one observer noted, "There is no cosmetic for beauty like happiness."[1] Perhaps the interviewer reacts more positively to people who are happier/more satisfied and thus rates their beauty more highly. This reflection of happiness in assessed beauty could stem from the person's overall demeanor, from his/her smiling more, or from some other specific cause. Regardless of its source, however, this possible reverse causation necessitates using other sets of data, with other methods of assessing beauty, to replicate our results.

Another American data set with measures of both beauty and happiness is the Wisconsin Longitudinal Survey (WLS), a study of a cohort of high-school graduates from 1957. The graduation pictures of the participants were assessed in 2004 by panels of raters, nearly all of whom were born earlier than the respondents, and who thus had a feel for what was viewed as good or bad looks in the late 1950s. Each respondent's picture was rated by 12 individuals (6 men and 6 women), with raters' ages ranging between 63 and 91 (in 2004). The ratings were normalized within a given rater (de-meaning and dividing by the standard deviation) and averaged over raters.

WLS respondents were interviewed in 1992 and 2004 (at ages 53 and 65) and asked questions about their general feelings in recent days. In particular, each was asked how many days last week s/he was happy, how many days s/he enjoyed life, and how many days s/he was sad. These happiness measures were thus obtained 35 and 47 years after the photographs from which the respondents' beauty was rated were taken. Since a lot of time has elapsed between the beauty rating

[1] Marguerite Power Farmer Gardiner, Countess of Blessington (1789 –1849).

and happiness assessments, measurement error in beauty (as a proxy for current beauty) could be a concern.[2]

The fourth data set is the British National Child Development Study (NCDS), a longitudinal examination of Britons born March 3-9, 1958. At age 7, and again at age 11, each student's teacher assessed his/her attractiveness, along a scale shown in column (1) of Table 1. We aggregated these into the three categories, good-looking, average-looking and unattractive, similar to previous work relating these ratings to subsequent earnings (Harper, 2000). In various later waves of the survey, including 1991, 1999, 2004 and 2009 (ages 33, 41, 46 and 51), the remaining respondents were asked questions designed to elicit their happiness or life satisfaction, some which have been studied before using these data (e.g., Blanchflower and Oswald, 2008). In the three most recent waves life satisfaction was elicited in a question (column (2) of Table 1) focusing on the respondent's entire life experience. Happiness at age 51 was also measured in a backward-looking manner, while happiness at age 33 was measured with reference to the respondent's current situation only.

Being assessed even earlier than those in the WLS, these beauty measures arguably suffer from more measurement error as proxies for beauty in adulthood. Potentially far more serious is that they are based on assessments by teachers, who presumably knew the students well and are unlikely to be unbiased assessors of their student's attractiveness. Any highly autocorrelated unobservable effect—ability, drive, etc.—is likely to be positively correlated with the assessment of beauty. Whether this positive bias to the estimated effect of this beauty measure on happiness is greater or smaller than the attenuation bias due to classical measurement error is unclear; but for our purposes these data have at least as great difficulties as those in the QAL surveys.

In Tables 2a-2c we present descriptive statistics for the QAL surveys, the WLS and the NCDS respectively. Consider first the QAL. As is usual in assessing beauty, more people are rated

[2]Other studies have assessed beauty from school pictures taken nearly two decades before the outcome to which the assessments were linked (Biddle and Hamermesh, 1998), and one study even showed a high correlation between the assessments of pictures of 10-year-olds and those of the same individuals taken at age 50 (Hatfield and Sprecher, 1986, p. 283).

in the top two categories than in the bottom two; and the majority are rated as average-looking. Also as is usual, women are rated more extremely than men (Hamermesh, 2011, Chapter 2). Consistent with the previous satisfaction/happiness literature, most people are fairly happy and satisfied. There is no consistent gender difference in average satisfaction/happiness.

Because the beauty measures in the WLS were normed, we do not list them in Table 2b. In these data people report being happy on most days (on average, between 5 and 6 days) in the week before the survey. The number of days reported as being sad is typically 20 percent or less than the number of happy days. With one exception—number of days reported happy in the 1992 wave of the survey—female respondents are less happy than males.

Table 2c shows that, as in the QAL data, in the NCDS females' looks (in this case at age 11) were rated more extremely than males'. Perhaps, however, because of their close acquaintance with their charges, the teachers who rated the students' attractiveness included more students in the attractive (good-looking) category than in the excluded category (presumably children viewed as neither attractive nor unattractive). Most of the respondents were fairly happy or satisfied at ages 33-51; but, unlike in the WLS, in three of the five cases women were significantly happier/more satisfied than men.

Before discussing our main regression results, we provide a brief theoretical discussion regarding concerns about measurement and endogeneity problems with the beauty ratings that we use. To focus only on the beauty rating, consider the following simple linear regression model, with H_t (satisfaction/happiness) as the dependent variable and true (latent) beauty B_t^* as the explanatory variable:[3]

$$H_t = \alpha + \beta B_t^* + \varepsilon, \qquad E(\varepsilon) = 0, \qquad Var(\varepsilon) = \sigma_\varepsilon^2.$$

[3]We assume homoskedasticity throughout this section and therefore omit conditioning on B_t^* to simplify notation.

The subscript t indicates the time at which the happiness measure is observed. B_t^* is a measure of beauty that would be obtained by a person who had no contact with the subject but evaluated his/her beauty at the same time the subject's satisfaction/happiness was measured.

We consider three possible sources of difficulty in measuring beauty:

(1) *Measurement error in the beauty rating*: The beauty rating used in the actual regression is an imperfect measure of B_t^*, even if it is obtained at the same time t as the happiness measure.

(2) *Depreciating accuracy in the beauty rating*: Since beauty changes, albeit slowly, over time, the inherent noise in the beauty rating will be larger the more that the rating pre-dates the satisfaction/happiness measure. The variance will be an increasing function of the time interval between observation of the beauty rating and observation of the happiness measure, a problem in both the NCDS and the WLS.

(3) *Bias in the beauty rating*: If the beauty rating is elicited after the rater has spent time interacting with the subject (e.g., in an interview, as in the QAL, or as a teacher, as in the NCDS), we would expect a positive correlation between the beauty rating and the unobservable component ε of the happiness outcome. For instance, an interviewer might have a better opinion of a subject's beauty if the subject projects self- confidence in the interview, which might occur if the subject is happier.

The following stylized model for the observed beauty rating incorporates each of these three possible sources of difficulty:

$$B_s = B_t^* + \nu_{t-s} + \eta,$$

where $s \leq t$ is the time at which the beauty rating B_s is obtained. The depreciation component of the measurement error is ν_{t-s}, which has a variance that is linear in the time interval $(t-s)$:

$$Var(\nu_{t-s}) = (t-s)\sigma_\nu^2, \qquad Cov(\nu_{t-s}, \varepsilon) = 0.$$

The other component of the error, denoted η, is similar to a classical measurement error, except that we also allow it to be correlated with the happiness residual ε:

$$Var(\eta) = \sigma_\eta^2, \qquad Cov(\eta, \varepsilon) = \sigma_{\eta\varepsilon}, \qquad Cov(\eta, \nu_{t-s}) = 0.$$

For this general model, the inconsistency of the least-squares estimator is given by the probability limit of the slope estimate:

$$plim\ \hat{\beta} = \frac{Cov(B_s, H_t)}{Var(B_s)} = \frac{Cov(B_t^* + \nu_{t-s} + \eta, \alpha + \beta B_t^* + \varepsilon)}{Var(B_t^* + \nu_{t-s} + \eta)}$$

$$= \beta \left(\frac{\sigma_{B^*}^2 + \frac{\sigma_{\eta\varepsilon}}{\beta}}{\sigma_{B^*}^2 + (t-s)\sigma_\nu^2 + \sigma_\eta^2} \right), \qquad (1)$$

where $\sigma_{B^*}^2$ denotes the variance of B_t^*. (Note that the textbook case of classical-measurement error is a special case of this formula corresponding to $\sigma_{\eta\varepsilon} = 0$ (no rating bias) and $s = t$ (no depreciation effect), for which $plim\ \hat{\beta} = \beta \left(\frac{\sigma_{B^*}^2}{\sigma_{B^*}^2 + \sigma_\eta^2} \right)$.)

For the QAL datasets, the beauty rating is provided by the interviewer at the same time that the satisfaction/happiness measures are elicited. There is no depreciation effect since $s = t$, but the interview format leads to the possibility of a bias in the beauty rating ($\sigma_{\eta\varepsilon} > 0$). The probability limit of the slope-estimate in this case simplifies to

$$plim\ \hat{\beta} = \beta \left(\frac{\sigma_{B^*}^2 + \frac{\sigma_{\eta\varepsilon}}{\beta}}{\sigma_{B^*}^2 + \sigma_\eta^2} \right).$$

If β is positive, the usual attenuation inconsistency associated with classical measurement error is offset by the inconsistency associated with the beauty-rating bias. The overall direction of the inconsistency depends on whether $\frac{\sigma_{\eta\varepsilon}}{\beta} > \sigma_\eta^2$ (upward inconsistency) or $\frac{\sigma_{\eta\varepsilon}}{\beta} < \sigma_\eta^2$ (downward inconsistency).

For the WLS dataset, the beauty rating is based upon a subject's high-school picture and the happiness measure is elicited during adulthood. A depreciation effect would be expected, since

$s < t$, and the lack of interaction between the rater and the subject eliminates concerns about beauty-rating bias. Entering $\sigma_{\eta\varepsilon} = 0$ into the general formula yields:

$$plim\ \hat{\beta} = \beta\left(\frac{\sigma_{B^*}^2}{\sigma_{B^*}^2 + (t-s)\sigma_v^2 + \sigma_\eta^2}\right).$$

Attenuation inconsistency is expected here due to the mis-measurement and depreciation. The estimated beauty slope from the WLS regressions should therefore be considered as "too low" (a lower bound to the true effect).

Finally, for the NCDS dataset, all three difficulties could arise, since the beauty ratings were assessed in childhood ($s < t$) by teachers who were very familiar with the subjects ($\sigma_{\eta\varepsilon} > 0$). As a result, the general probability-limit formula in (1) would apply. As with the QAL data, the beauty-rating bias acts in an opposite direction from the measurement error. The impact of measurement error here, however, would be expected to be much larger than in the QAL, since the beauty ratings in the NCDS are assessed decades before the expression of happiness is elicited (as captured by $(t-s)\sigma_v^2$). Although it is difficult to say how the values of $\sigma_{\eta\varepsilon}$ would compare between the QAL and NCDS, the large difference in the variance of measurement error between the two studies leads us to believe that the probability limit for the NCDS estimates would be lower than that for the QAL estimates.

III. Basic Results

In this section we use OLS to estimate linear models relating life satisfaction/happiness to beauty in each of the four data sets. For each we first include as regressors only the beauty measure(s) and, in the QAL, the few covariates that cannot be caused by differences in beauty; then we add a number of covariates that have been shown to affect happiness but may not mediate the effect of beauty on satisfaction/happiness. In the next section we report on a large number of robustness checks that include varieties of additional controls, alternative beauty measures and more complex estimation procedures.

Table 3a presents the estimates based on the two QAL surveys. Among women all the coefficients have the expected signs—positive on the indicator for good looks (above-average or beautiful—the upper third of looks), negative on the indicator for bad looks (below-average or homely—the bottom eighth of looks). This is true whether or not we control for age, education, race, number of children and marital status.[4] Indeed, the addition of the vector of controls hardly alters the point estimates of the coefficients among women; and nearly all the estimates are statistically significantly different from zero. Among men almost all of the point estimates have the expected sign, and they are generally statistically significant in the 1978 data. As with women, adding the vector of controls does not greatly alter the point estimates.

The effects of differences in beauty on life satisfaction or happiness are not small, at least in the 1978 data. Using the estimates from the equations expanded to include controls, going from the bottom eighth of women's (men's) looks (those rated below-average) to the top third (those rated above-average) raises satisfaction with life by 0.45 (0.48) standard deviations; the effects on happiness of this difference in beauty are 0.38 (0.48) standard deviations. The impacts of differences in beauty in the 1971 data are smaller, but still average about 0.20 standard deviations.

The results based on the WLS, with number of days happy, enjoyed and/or sad, are presented for the respondents observed at ages 53 and 65 in Table 3b. The upper part of the table contains only the unit-normal measure of beauty, while the bottom part adds years of education, marital status, number of children, BMI observed at high-school graduation, and current BMI. These latter two, which were unavailable in the QAL surveys, allow for possible correlations between ratings of attractiveness and overweight/obesity (although the evidence for the labor market suggests that the correlations, and their impacts on wages, do not affect the estimated effects of beauty on

[4]Whether we should be controlling for marital status here is unclear. There is substantial evidence that married people are happier (e.g., Blanchflower and Oswald, 2004; Oswald and Wu, 2010); but one's gains from marriage are affected by one's looks (Hamermesh and Biddle, 1994). In both of these data sets, however, there are only small differences in marital status by looks: Holding age and gender constant, being above-average in looks has no effect on marital status compared to being average-looking, while being below-average in looks has a marginally significant negative impact on the probability of being married.

outcomes—Hamermesh and Biddle, 1994). As with the results for the QAL, adding this vector of covariates hardly alters the estimated impacts of attractiveness on the measures of satisfaction/happiness.

There is no significant impact of attractiveness on happiness among men at either of the two ages at which these adults are observed. Among women, however, in all the estimated equations the more attractive respondents are significantly happier at age 53 than less attractive respondents. The impacts are smaller relative to the standard deviations of satisfaction/happiness than in the QAL (see below), an unsurprising result of the likelihood of the differences in the nature of the measurement errors in the two samples.

We can explain the disappearance of the results for women as they age by the possibility that the correlation of attractiveness at age 18 with attractiveness at age 53 may be greater than that with attractiveness at age 65. The absence of any relation between attractiveness and happiness among men is harder to explain, especially in light of the fact that labor-market effects are at least as large among men as among women. One possibility is that there is inherently more measurement error in the ratings (assigned over 40 years after the pictures were taken) of men's high-school graduation pictures than of women's, a possibility we examine indirectly in the next section.

Table 3c shows the results of relating measures of happiness and satisfaction in adulthood in the NCDS sample to attractiveness as assessed by a child's teacher at age 11. The first part of the table includes only indicators for being rated as attractive or as unattractive (with a middle category excluded). All of the estimated impacts that are statistically significant are of the expected sign, and there is no obvious difference in the size or significance of the effects between men and women.

The second part of Table 3c reports the estimates of the impacts of the beauty indicators when indicators for educational attainment, marital status, number of children, BMI at age 11 and

current BMI are added to the equations.[5] The estimated effects of attractiveness are typically somewhat attenuated when the control variables are added, although the overall conclusions remain the same: Where significantly nonzero, the beauty measures have the expected effects; and, as in the upper part of the table, the impacts of beauty are roughly the same by gender.

There is a welter of estimates here—24 coefficient estimates for each gender, based on four different samples. Among men in all four samples taken together, only 2 of the 24 have the "incorrect" sign; and of the 22 "correctly" signed, 5 are statistically significantly nonzero. Among women 4 of the estimated effects are incorrectly signed (and not significantly nonzero), but 11 are correctly signed and statistically different from zero.

While these comparisons clearly suggest a positive answer to the titular question of this study, we would like to compare the estimates across the samples, given the differences in the potential biases to our estimates across the samples. To do so we calculate the effect of being at different percentiles of the distribution of beauty on the level of satisfaction/happiness measured in standard deviations. Thus, for example, we assume that the average male among the 12.5 percent rated as below-average in the QAL 1971 is at the 6^{th} percentile of the distribution of looks and is thus 1.53 standard deviations below the mean beauty of men. We use this type of approximation for all the QAL and NCDS results, and for the WLS (where the beauty rating was a unit normal deviate) we impute as percentiles of the distributions the simple averages of the imputed percentiles in the other surveys.

The results of these calculations are shown in Figures 1a and 1b, with each of the 30 points in a Figure representing the fractional change in standard deviations of satisfaction/happiness generated by a movement from the mean beauty to some point in the distribution below or above the mean. Among men (women) the average good-looking respondent is 0.83 (0.71) standard deviations above the mean of beauty, while the average bad-looking respondent is 1.65 (1.58) standard

[5]The categories represented by the vector of education indicators are: cse or equivalent, O-level or equivalent, A-level or equivalent, higher qualification, or university degree or higher, with no qualification the excluded category.

deviations below the mean. On average, among men (women) the gain from being this good-looking is 0.04 (0.04) standard deviations along the satisfaction/happiness index compared to the average male (female), while the loss from being this bad-looking is 0.09 (0.10) standard deviations of satisfaction/happiness. Assuming, as these calculations must, that the effects are linear within the categories above-average and average, or attractive and unattractive, the results in the expanded specifications imply that a one standard-deviation increase in beauty raises satisfaction/happiness by 0.05 (0.06). This is not large, far smaller than the impact of income on happiness in a cross-section (computed from Frey and Stutzer, 2002, Table 1), although that calculation is based on decile averages rather than individual observations. The effect is about the same magnitude as has been reported in other literatures dealing with related subjects (e.g., Carrell *et al*, 2010, on educational production functions).

The relative sizes of the estimates accord with the discussion in Section II of measurement error. They are largest in the regressions based on the QAL, where we expected the simultaneous assessment of beauty and happiness to generate positively correlated measurement errors. They are smallest, and certainly negatively biased, in the estimates based on the WLS, where changes in beauty will have led to classical measurement error that has grown over time. As in the QAL the direction of the bias in the estimates based on the NCDS is unclear, since the errors induce opposite-signed biases, but as expected the estimates are generally below those from the QAL.

Overall the estimates from the four sets of data suggest the following tentative conclusions:

1. There is a positive effect of good looks on satisfaction or happiness, and a negative effect of bad looks, even accounting for a variety of demographic variables that might be correlated with beauty and/or satisfaction/happiness.
2. These effects are not huge, but the true impacts are certainly larger than those implied by the WLS results.
3. The impacts of beauty or its absence on satisfaction/happiness generally seem slightly larger among women than among men, and the estimated effects are more reliable among women.

IV. Robustness Checks and Methodological Extensions

In the previous section we attempted to place a lower bound on the estimated impacts of beauty on happiness. In none of the estimation, however, were we concerned about alternative measures and specifications, nor did we consider alternative approaches to estimation. We do that here, in each case basing the estimates on the expanded specifications with control variables (Specification 2) in Tables 3a-c.

A. Re-specifying Proxies for Beauty and Considering Confounding Variables

No sensible reformulations of the beauty ratings in the QAL survey can be done to check their robustness; but we can use alternative measures in the WLS and NCDS.[6] In the former we re-estimated the expanded specifications using first the normalized beauty ratings given by female raters to pictures of female respondents, and by male raters to male subjects. We then switched and re-estimated the equations using opposite-sex ratings. Most of the estimates are attenuated slightly, just as expected assuming that there is more measurement error in these assessments of beauty when fewer raters are used; but all of those that were statistically significant (women in 1992) remain so.

The NCDS respondents' appearance was assessed by their teachers at age 7 as well as by their teachers at age 11 (the measures used in Section III). To the extent that the measurement error in the variable we used arose from random errors in an individual teacher's assessment of the child's appearance, averaging the teachers' ratings at ages 7 and 11 will reduce that error. Accordingly, we average the indicator variables for appearance at 11 with identically defined variables describing appearance at age 7. These average measures replace the age-11 measures in the estimating equations, and the age-11 BMI is replaced with the average of BMI at ages 7 and 11. In a few cases some previously insignificant parameter estimates in Table 3b become marginally significant, but

[6]In the QAL 1971, for example, only 59 of the respondents are rated as strikingly handsome or beautiful, and only 44 are rated as homely. When we re-estimated the models with measures encompassing each of the five beauty ratings, unsurprisingly, given the cell sizes at the extremes, this extension hardly altered the conclusions.

otherwise there is no change. Implicitly, whatever measurement errors exist in the age-11 proxies are highly positively correlated with those in the age-7 data and thus cannot be eliminated by averaging.[7]

Another concern is that different assessors rate beauty differently, and that their idiosyncrasies may be correlated with the subjects' happiness. With each teacher in the NCDS assessing only one subject's appearance this issue cannot be examined in those data; and we cannot identify the raters in the WLS. In the QAL surveys, however, we know which raters assessed each subject's beauty. Accordingly, we re-estimate the equations in the bottom half of Table 3a adding interviewer fixed effects. With one exception (the impact of bad looks among women in the 1978 data) none of the significant impacts shown in Table 3a became statistically insignificant, nor did any of the estimated effects of looks on happiness reverse sign. The vectors of interviewer fixed effects were themselves always statistically significantly nonzero—some raters were consistently more generous than others—but their addition hardly altered the estimated effects of beauty on happiness.

Another potential difficulty with the results is that there are location-specific determinants of beauty that may also directly affect people's perceived satisfaction/happiness. For example, perhaps living in Los Angeles with its proximity to mountains and ocean makes people happier and also attracts good-looking people. In the two data sets that contain longitudinal data there is strong evidence, consistent with Gautier *et al* (2010), that changes in location are related to beauty.[8] To

[7]We also used the age-7 measures alone—looks and BMI—in place of the age-11 measures. Perhaps unsurprisingly, the results were slightly weaker than with the age-11 measures, and thus somewhat weaker still than the specifications based on the age-7 and age-11 averages.

[8]For example, in the WLS we can compare the beauty of the Wisconsin high-school graduates who remained in the state to those who were not living there at age 51. The average beauty of those still residing in Wisconsin at age 51 was 0.004 (s.e.= 0.045), while the beauty of those who had left was 0.098 (s.e.= 0.030). Those who remained are significantly worse-looking than those who left.
Because the NCDS was national, it allows us to compare the beauty of those who entered, those who stayed and those who left an area. Because the definitions of the British regions were not the same in all the NCDS waves, we cannot examine mobility and beauty for all areas; but southeast England, and Scotland and Wales, are consistently identified at ages 11 and 33. (Since most geographic mobility in the sample occurs between these ages, this is the most useful single comparison.) 0.572 (s.e. = 0.013) of those who moved to the Southeast

investigate this issue in the QAL surveys, we add state fixed effects to the basic equations in Table 3a. As in the other re-specifications in the QAL, this extension hardly altered the results. Although the vector of state effects was itself statistically significant, and its inclusion did increase slightly the absolute values of the estimated effects of the beauty indicators, there were no qualitative changes in their impacts.

In the WLS the only locational information is whether the Wisconsin high-school graduate still resides in the state at the time of the interview. Adding this location variable to the specifications for the three outcomes in Table 3b produces only minute changes in the estimated impacts of beauty. No signs change, and the impacts remain statistically significant only for women observed in 1992.

In the NCDS we can account for regional effects in both the assessments of beauty during childhood and the effect of childhood beauty on adult happiness. There may be regional differences in beauty standards, which using region in childhood as a control would account for; and there may be regional differences in the relationship between beauty and satisfaction/happiness, which using location as an adult could account for. We thus re-specify the equations in Table 3c to include vectors of regional indicators at age 11 and at the time the respondent's happiness/satisfaction is reported. In no case did either of these vectors of fixed effects approach statistical significance; nor did their inclusion qualitatively alter the impacts of beauty on satisfaction/happiness. In these data, at least, regional differences in childhood and adulthood just are not important in affecting the estimated relationships between beauty and satisfaction/happiness.

---

were good-looking at age 11, but only 0.545 (s.e. = 0.013) of those who stayed were, and only 0.512 (s.e.= 0.032) of those who left were good-looking. The Southeast attracted good-looking people, while less good-looking people moved elsewhere in the U.K. While not statistically significant, the differences in Scotland and Wales are exactly opposite those in the Southeast: 0.571 (s.e.= 0.040) of those who entered were good-looking; .587 (s.e.= 0.013) of those who stayed were; and 0.598 (s.e.= 0.031) of those who left were. The small samples of entrants and leavers mean these differences are not statistically significant; but the differences in the fractions considered bad-looking are: 0.078 (s.e.= 0.022) among entrants, 0.097 (s.e.= 0.008), and 0.110 (s.e= 0.020) among leavers.

Before examining additional covariates that we have ignored, consider a conundrum in the WLS results: The effects of beauty on happiness are apparent (in Table 3b) in 1992 (at age 53) but not in 2004 (at age 65). One reason might be that beauty effects on happiness generally diminish as one advances past middle age. To explore this possibility the only way that these data sets allow, we re-estimate the second specifications in Table 3a for the QAL surveys, adding interactions of the quadratic in age with the variables measuring beauty assessments. While in about half of the eight re-specified equations the additional four variables were jointly significant (at the 10-percent level), the effects of beauty on happiness did not vary systematically with age in these equations taken together.

Another possibility is that the nearly 20-percent sample attrition in the WLS between 1992 and 2004 changed the structure of the beauty-satisfaction/happiness relationships in the sample. Re-estimating the models for 1992 with only those respondents who remained in the sample in 2004 generated the same significant results for women as were reported in Table 3b. Excluding these two explanations suggests either that the measurement errors resulting from using beauty assessed at age 18 become increasingly important after middle age, or that some other, unknown effect generates the disappearance of the beauty effects on women's happiness.

An important determinant of happiness in other studies (see Stevenson and Wolfers, 2008, Blanchflower and Oswald, 2008) has been health status. The difficulty here is that in all the data sets the measures of health are subjective, self-assessed, so that they are very likely to be determined by the same factors that determine happiness/satisfaction. Nonetheless, we add self-reported health to the second specifications in each data set.

In the QAL surveys subjective health is based on the response to a question about whether the respondent has health problems (with about 30 percent responding yes). An indicator of this subjective response in the specifications is strongly positively correlated with the happiness and satisfaction measures, as it is in the extensions of the WLS and NCDS results below. The estimated

impacts of above-average and below-average looks on these outcomes, however, change only slightly, with only small decreases in their statistical significance and no changes in their signs.

The WLS contains self-rated health measures on the five-point scale, and we create an indicator for good health at the time of the interview comparing those whose self-rated health is excellent or very good to those whose self-rated health is good, fair or poor. Between 60 and 67 percent of the respondents are classified as being in good health using this indicator. Adding it to each of the specifications again has only small effects on the estimates and their significance.

In the NCDS subjective health was rated on a four-point scale—excellent, good, fair or poor—at ages 33 and 41; on a five-point scale (adding very poor) at age 46, and on a 100-point scale at age 51. We have created an indicator—good health—typically equaling 1 if the individual reports being in excellent or good health, 0 otherwise (with good health coded 1 for 86, 82, 77 and 76 percent of the respondents at ages 33, 41, 46 and 51 respectively). In some of the equations for men the estimated effects of beauty on satisfaction/happiness are attenuated by the addition of subjective health status to the specification, although there are essentially no changes in the estimated effects in the equations among women. Given the subjective, jointly reported nature of satisfaction/happiness and health status, the changes in the parameters we are focusing on seem fairly minor.

A final check includes additional covariates related to the respondent's siblings or parents. For some of WLS respondents the survey obtained measures of their sibling's happiness as well as their own. To the extent that there are family background effects in happiness (Hartog and Oosterbeek, 1997), including these measures could help isolate the effect of beauty independent of any correlation it might have with unmeasured characteristics in the respondent's family background. Own beauty is positively correlated with siblings' beauty, so that adding sibling happiness to the specifications might be expected to have made the estimated effects of own beauty on happiness more negative.

The WLS provides measures of siblings' happiness for about half of the 1992 respondents. Adding these to the specifications nearly doubles the absolute magnitudes of the beauty impacts on

happiness at age 53 (and has little effect in the 2004 data). This appears due to sample selection: The same large effects are produced when we re-estimate the basic equation only over those respondents on whose siblings' happiness information was available. We thus cannot distinguish whether the surprising increase in the effects is due to unexpected correlations of background measures (including the unreported siblings' beauty) with reported happiness, or simply to selection issues.

In the NCDS we know whether the respondent's parents are alive at his/her age 41, 46 and 51. Since parental death has been shown to affect well-being, we thus form an indicator for these latter two waves equal to 1 if a respondent's parent died in the five years preceding the interview (which occurred for 16 percent of the respondents at age 46 and 14 percent at age 51).[9] In each of the six equations describing satisfaction/happiness at ages 46 and 51 a parental death in the quinquennium does have a negative effect; unsurprisingly, given that there is no reason to expect that the assessment of beauty at age 11 will be correlated with parental mortality three decades later, this additional variable leaves the estimated beauty effects essentially unchanged.[10]

B. Methodological Extensions

All of the specifications reported in Tables 3a-3c were estimated by OLS—although in most cases the expressions of satisfaction/happiness take only a small number of values. To examine whether a discrete-variable modeling approach would change the results, we re-estimate the models (except for satisfaction in the QAL 1978, where the questionnaire allowed a large range of discrete responses). Except for the WLS this means estimating ordered probits over these specifications; in the WLS, since the measures are of numbers of days (ranging from 0 to 7), we re-estimated the

[9]A five-year retrospective may be too long to observe an effect on current happiness, but that is the hiatus that the data limits us to.

[10]The effects are not changed greatly in any of the data sets when all the additional variables are included at the same time. This is not surprising, as within each data set these additional variables are typically nearly orthogonal.

model using a count-data method (Poisson estimation). These various estimation methods yielded qualitatively very similar results to those reported in Section III, with significant effects remaining.

In the WLS data the respondents' underlying happiness is measured in three ways—by the days s/he identifies as being happy, as sad, or as having been enjoyed in the past week. Each of these measures can be viewed as proxying the respondent's underlying mental state—each is a noisy measure of that. To remove some of the noise we first simply subtract days sad from days happy and re-estimate the specifications that were presented in the bottom half of Table 3b. Then, since we do not know what the appropriate weights on these particular expressions of happiness might be, we re-estimate these equations using the first eigenvector (first principal factor) to weight the three expressions of happiness—days happy, days sad and days enjoyed—and use this measure as the expression of happiness to be described by the independent variables

We present these re-estimates in Table 4. They do not add much to the basic results. Whether we just subtract sad from happy days, or use information on all three responses by using a dependent variable which weights each using the first eigenvector, the implications are the same as before: Among men the effects of beauty are positive but statistically insignificant; among women they are positive and highly significant in 1992, insignificantly negative in 2004.

V. Inferring the Direct and Indirect Effects of Beauty

The economic question in this study deals with the extent to which the effect of beauty on satisfaction/happiness occurs through markets: How much of the effect is direct—with people who are otherwise identical in every respect being happier or more satisfied than their less good-looking peers? How much is due to the fact that beauty enhances one's outcomes in various markets, including the labor and marriage markets?[11]

Writing the simplest model with no covariates except those that could not be affected by beauty (age and race in the QAL) as:

[11]The effect of a different personal endowment, height, on happiness was decomposed into these components by Deaton and Arora (2009), with the adjustment limited to accounting for the impact of height in the labor market.

$$H_t = \alpha + \beta B_t^* + \varepsilon,$$

we can interpret the estimate of β as the total effect of beauty on satisfaction/happiness If we add as many covariates as we have information—both those included in the expanded specifications in Tables 3, self-reported health, and measures of earnings and spouse quality—we obtain:

$$H_t = \alpha' + \beta' B_t^* + \gamma X_t + \varepsilon'.$$

We definte the direct effect of beauty on satisfaction/happiness as β', and the indirect effect as $\beta - \beta'$.

In the QAL the only additional covariates included in the vector X are self-reported health and the available measures of individual income. No doubt this paucity of additional covariates will generate an additional upward bias in the direct effect beyond the possible overall bias that we already noted. In the NCDS estimates for that age 33, 41 and 51 waves we add self-reported health and own and spouse/partner's weekly earnings, while in the age 46 wave we replace spouse/partner's earnings with family income (as partner's weekly earnings are unavailable). Finally, in the WLS estimates for each of the two years we add health status and variables measuring the respondent's own income and the household income (thus presumably proxying the earnings of a spouse if one is present).

Rather than presenting the estimates of these expanded specifications, in Table 5 we simply list the average effects of going from the mean beauty to being good-looking or bad-looking, measured in standard-deviation units of satisfaction/happiness per standard deviation of beauty. (The statistics listed in Table 5 for Specification 2 are the averages of those shown in Figures 1.) The results show that the direct effects (based on the estimates from Specification 3) are typically around half of the total effect (based on the estimates from Specification 1). While this further expansion of the estimates accounts for some of the indirect effects of beauty on satisfaction/happiness, the available data do not allow us to account for the impacts in other markets. As just one example, there is growing evidence that beauty generates beneficial outcomes in lending markets (Hamermesh, 2011, Chapter 7). Moreover, our proxies for the outcomes in the labor and marriage

markets that are affected by beauty are far from perfect. It thus seems fair to conclude that the direct effect of beauty is at most one-half of the total effect, and perhaps much less. The majority of the impact of beauty on satisfaction/happiness appears to be economic—through its effects on outcomes in various markets.

VI. Conclusions and Extensions

We have examined the relationship between people's life satisfaction/happiness and their beauty. Both are subjective, although in each of our four empirical examples the agent describing his/her satisfaction differs from the agent(s) describing his/her beauty. While the beauty measures introduce difficulties into the inference of the true effect of beauty on happiness, those difficulties, which differ across our data sets, do not result because we make the simple mistake of essentially relating happiness to a proxy for happiness. The difficulties with the beauty measures are more subtle in our context, but they allow us to put a lower bound on the magnitudes of the true impacts of beauty on happiness.

The results suggest that a person's beauty does increase his/her satisfaction/happiness. The effects may be larger among women than among men, but the differences by gender in the relationship are not great. Among both men and women at least half of the increase in satisfaction/happiness generated by beauty is indirect, resulting because better-looking people achieve more desirable outcomes in the labor market (higher earnings) and the marriage market (higher-income spouses).

Overall our findings imply that much of the differences in happiness that exist in a society arise from characteristics that are completely beyond one's control. Substantial evidence (see Hamermesh, 2011, Chapter 2) makes it clear that even radical measures to alter one's looks have fairly small effects. At least along the dimension of this one determinant of happiness, focusing on creating a happier society is not likely to be fruitful.

REFERENCES

Biddle, Jeff, and Daniel Hamermesh, "Beauty, Productivity and Discrimination: Lawyers' Looks and Lucre," *Journal of Labor Economics*, 16 (January 1998): 172-201.

Blanchflower, David, and Andrew Oswald, "Well-Being over Time in Britain and the USA," *Journal of Public Economics*, 88 (2004): 1359-86.

----------, and ----------, "Hypertension and Happiness Across Nations," *Journal of Health Economics*, 27 (2008): 218-33.

Carrell, Scott, and James West, "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," *Journal of Political Economy*, 118 (June 2010): 409-32.

Clark, Andrew, Ed Diener, Yannis Georgellis and Richard Lucas, "Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis," *Economic Journal*, 118 (June 2008): F222-43.

Deaton, Angus, and Raksha Arora, "Life at the Top: The Benefits of Height," *Economics and Human Biology*, 7 (July 2009): 133-6.

Diener, Ed, Frank Fujita and Brian Wolsic, "Physical Attractiveness and Subjective Well-Being," *Journal of Personality and Social Psychology*, 69 (July 1995): 120-129.

Easterlin, Richard, "Does Money Buy Happiness?" *The Public Interest*, 30: 3-10.

Frey, Bruno, and Alois Stutzer, "What Can Economists Learn from Happiness Research?" *Journal of Economic Literature*, 40 (June 2002): 402-435.

Gautier, Pieter, Michael Svarer and Coen Teulings, "Marriage and the City: Search Frictions and Sorting of Singles," *Journal of Urban Economics*, 67 (March 2010): 206-18.

Hamermesh, Daniel, "Subjective Outcomes in Economics," *Southern Economic Journal*, 71 (July 2004): 2-11.

----------. *Beauty Pays*. Princeton, NJ: Princeton University Press, 2011.

---------- and Jeff Biddle, " Beauty and the Labor Market," *American Economic Review*, 84 (Dec. 1994): 1174-1194.

Harper, Barry, "Beauty, Stature and the Labour Market: A British Cohort Study," *Oxford Bulletin of Economics and Statistics,* 62 (2000): 771-800.

Hartog, Joop, and Hessel Oosterbeek, "Health, Wealth and Happiness: Why Pursue a Higher Education?" *Economics of Education Review*, 17 (June 1998): 245-56.

Hatfield, Elaine, and Susan Sprecher, *Mirror, Mirror…: The Importance of Looks in Everyday Life*. New York: SUNY Press, 1986.

Kahneman, Daniel, and Angus Deaton, "High Income Improves Evaluation of Life but not Emotional Well-Being," *Proceedings of the National Academy of Science*, (August 2010):

Leigh, Andrew, and Tirta Susilo, "Is Voting Skin-deep? Estimating the Effect of Candidate Ballot Photographs on Election Outcomes," *Journal of Economic Psychology*, 30 (Feb. 2009): 61-70.

Mocan, H. Naci, and Erdal Tekin, "Ugly Criminals," *Review of Economics and Statistics*, 92 (Feb. 2010): 15-30.

Möbius, Markus, and Tanya Rosenblat, "Why Beauty Matters," *American Economic Review*, 96 (March 2006): 222-35.

Oswald, Andrew, and Stephen Wu, "Well-Being Across America: Evidence from a Random Sample of One Million Americans," *Review of Economics and Statistics*, 93, forthcoming 2011.

Plaut, Victoria, Glenn Adams and Stephanie Anderson, "Does Attractiveness Buy Happiness? 'It Depends on Where You're From," *Personal Relationships*, 16 (2009): 619-30.

Scitovsky, Tibor, *The Joyless Economy*. New York: Oxford University Press, 1976.

Seligman, Martin, Authentic Happiness. New York: Free Press, 2004

Stevenson, Betsey, and Justin Wolfers, "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox," *Brookings Papers on Economic Activity* (Spring 2008): 1-87.

Umberson, Debra, and Michael Hughes, "The Impact of Physical Attractiveness on Achievement and Psychological Well-Being," *Social Psychology Quarterly*, 50 (Sept. 1987): 227-36.

Table 1. Descriptions of Beauty, Satisfaction and Satisfaction Measures, Four Data Sets

	Beauty measure	Satisfaction measure	Happiness
QAL 1971	5-point rating by interviewer at end of interview: Strikingly handsome or beautiful Good-looking (above average for age and sex) Average looks for age and sex Quite plain (below average for age and sex) Homely	How satisfied are you with your life as a whole these days? (7 to 1 scale)	Taking all things together, how would you say things are these days--would you say you're very happy, pretty happy or not too happy these days? (3 to 1 scale)
QAL 1978	Same as QAL 1971	How satisfied are you with your life as a whole ? (100 point scale)	Same as QAL1978
WLS	Constructed from ratings on an 11-point scale, with endpoints labeled as "not at all attractive" (1) and "extremely attractive" (11), based upon an individual's high-school yearbook photo (in 1957); each photo was rated by six men and six women, and the contructed measure is an average of the z-scores across the raters		Age 53: On how many days during the past week did you feel happy? (sad?) (values 0 through 7) Age 65: Same

	Beauty measure	Satisfaction measure	Happiness
NCDS	Teachers' ratings of the student's appearance at age 7, and at age 11. Which best describes the student? Attractive; unattractive; looks underfed; abnormal feature; scruffy and dirty. "Looks underfed and "scruffy and dirty" were coded as missing, "attractive" as good-looking, "unattractive" and "abnormal feature" as bad-looking, others as neither.	Age 41: How satisfied are you with the way your life has turned out so far? (10 to 0 scale, from completely satisfied to completely dissatisfied)	Age 33: All things considered, how happy are you? (4 to 1 scale)

Table 2a. Descriptive Statistics, QAL71 and QAL78, All Observations with Beauty Rating, Satisfaction and Happiness Responses*

	1971		*1978*	
	Men	Women	Men	Women
Good looking	0.271	0.308	0.334	0.348
Bad looking	0.125	0.174	0.101	0.120
Life satisfaction	5.556 (0.041)	5.538 (0.036)	82.313 (0.342)	82.096 (0.319)
Happiness	2.187 (0.020)	2.195 (0.017)	2.232 (0.014)	2.217 (0.013)
N	870	1217	1482	2069

*Sample averages are reported, with their standard errors for non-binary variables in parentheses.

Table 2b. Descriptive Statistics, WLS, All Observations with Beauty Rating, Satisfaction and Happiness Responses

	Men		*Women*	
	1992	2004	1992	2004
# days happy	5.321	5.673	5.466	5.654
	(0.027)	(0.058)	(0.054)	(0.053)
# days enjoyed	5.763	6.040	5.701	5.907
	(0.065)	(0.057)	(0.057)	(0.05b)
# days sad	0.668	0.465	1.115	0.841
	(0.042)	(0.037)	(0.049)	(0.044)
N	801	788	993	952

*Sample averages are reported, with their standard errors for non-binary variables in parentheses.

Table 2c. Descriptive Statistics, NCDS, All Observations with Beauty Rating

	Men	*Women*		
Attractive age 11	0.469	0.592		
Unattractive age 11	0.100	0.103		
N	7886	7450		
	Men	Women	Men	Women
	Age 33:		Age 41:	
Happiness	3.322 (0.009)	3.388 (0.010)		
Life Satisfaction			7.277 (0.028)	7.359 (0.030)
N	3514	3886	3945	4381
	Men	Women	Men	Women
	Age 46:		Age 51:	
Happiness			4.315 (0.020)	4.207 (0.020)
Life Satisfaction	7.562 (0.024)	7.657 (0.024)	7.335 (0.029)	7.314 (0.031)
N	3583	3920	3559	3773

*Sample averages are reported, with their standard errors for non-binary variables in parentheses.

Table 3a. Results from Regressions of Life Satisfaction and Happiness and Beauty Ratings, QAL 1971 and 1978*

	Men		Women	
	Good looks	Bad looks	Good looks	Bad looks
Specification 1: OLS, beauty, age quadratic, race, 1971				
Life Satisfaction	-0.034	**-0.255**	0.116	**-0.203**
	(0.096)	(0.129)	(0.084)	(0.100)
Happiness	0.041	-0.093	**0.097**	**-0.124**
	(0.047)	(0.062)	(0.039)	(0.046)
Specification 1: LPM, beauty, age quadratic, race, 1978				
Life Satisfaction	**2.673**	**-2.290**	**1.423**	**-3.279**
	(0.745)	(1.155)	(0.703)	(1.016)
Happiness	**0.108**	-0.068	**0.120**	**-0.096**
	(0.031)	(0.048)	(0.028)	(0.041)
Specification 2: Add age quadratic, education indicators race, number of children, married, 1971				
Life Satisfaction	-0.088	-0.177	0.113	-0.122
	(0.096)	(0.130)	(0.083)	(0.099)
Happiness	0.019	-0.058	**0.082**	-0.077
	(0.047)	(0.064)	(0.039)	(0.046)
Specification 2: Add age quadratic, education indicators race, number of children, married, 1978				
Life Satisfaction	**2.653**	**-2.405**	**1.323**	**-3.421**
	(0.075)	(1.159)	(0.707)	(1.016)
Happiness	**0.089**	-0.044	**0.101**	**-0.084**
	(0.031)	(0.048)	(0.028)	(0.041)

*Estimates that are significantly non-zero, one-sided 5-percent level, in **bold**.

Table 3b. Regressions of Days Happy, Enjoyed or Sad on Beauty Rating, WLS Ages 53 and 65

	Men		Women	
	1992	**2004**	**1992**	**2004**
Specification 1: OLS, beauty only				
# days happy	-0.012	0.016	**0.128**	-0.026
	(0.056)	(0.044)	(0.048)	(0.044)
# days enjoyed	0.069	0.008	**0.142**	0.006
	(0.056)	(0.045)	(0.050)	(0.049)
# days sad	-0.030	-0.014	**-0.100**	-0.040
	(0.033)	(0.028)	(0.042)	(0.039)
Specification 2: Add completed education, married, number of children, HS BMI, current BMI				
# days happy	-0.028	0.025	**0.118**	-0.052
	(0.056)	(0.046)	(0.049)	(0.045)
# days enjoyed	0.058	0.010	**0.116**	-0.026
	(0.056)	(0.047)	(0.051)	(0.047)
# days sad	-0.035	-0.012	**-0.096**	-0.032
	(0.032)	(0.030)	(0.043)	(0.041)

*Estimates that are significantly non-zero, one-sided 5-percent level, in **bold**.

Table 3c. Results from Regressions of Life Satisfaction and Happiness on Beauty Ratings, NCDS Ages 33, 41, 46 and 51

	Men		Women	
	Attractive Age 11	**Unattractive Age 11**	**Attractive Age 11**	**Unattractive Age 11**
	Specification 1: OLS, beauty only			
	Age 33			
Happiness	**0.036**	-0.023	0.036	-0.034
	(0.020)	(0.035)	(0.022)	(0.036)
	Age 41			
Life Satisfaction	**0.267**	-0.053	**0.146**	-0.043
	(0.058)	(0.104)	(0.068)	(0.114)
	Age 46			
Life Satisfaction	**0.118**	**-0.247**	-0.005	**-0.227**
	(0.049)	(0.088)	(0.055)	(0.092)
	Age 51			
Life Satisfaction	**0.132**	-0.177	**0.203**	**-0.367**
	(0.060)	(0.107)	(0.069)	(0.118)
Happiness	**0.106**	-0.120	0.057	-0.06
	(0.042)	(0.076)	(0.044)	(0.077)

Table 3c, cont.

	Men		Women	
	Attractive Age 11	**Unattractive Age 11**	**Attractive Age 11**	**Unattractive Age 11**
	Specification 2: Add education indicators, number of children BMI 11, BMI current, married/partnered			
		Age 33		
Happiness	0.029 (0.020)	-0.021 (0.034)	0.026 (0.022)	-0.013 (0.036)
		Age 41		
Life Satisfaction	**0.228** (0.062)	-0.006 (0.110)	0.025 (0.073)	-0.073 (0.123)
		Age 46		
Life Satisfaction	0.070 (0.052)	-0.120 (0.093)	0.060 (0.058)	**-0.176** (0.098)
		Age 51		
Life Satisfaction	0.053 (0.064)	-0.091 (0.114)	**0.137** (0.073)	**-0.267** (0.126)
Happiness	**0.099** (0.046)	-0.054 (0.082)	0.028 (0.048)	0.049 (0.084)

*Estimates that are significantly non-zero, one-sided 5-percent level, in **bold**.

Table 4. Combining the Impacts on Days Happy and Day Sad, WLS Ages 53 and 65*

	Men		Women	
	1992	**2004**	**1992**	**2004**
# days happy - # days sad	0.005 (0.072)	0.031 (0.065)	**0.227** (0.074)	-0.027 (0.071)
First eigenvector of # days happy, sad and enjoyed**	0.043 (0.083)	0.026 (0.073)	**0.216** (0.079)	-0.056 (0.075)

*Includes the additional regressors described in Table 3b.
**In 1992 the factor loadings are 0.768, -0.465 and 0.774 on days happy, sad and enjoyed. In 2004 they are 0.762, -0.458 and 0.770.
Estimates that are significantly non-zero, one-sided 5-percent level, in **bold**.

Table 5. Average Effects of Beauty on Happiness/Satisfaction, SDOutcome/SDLooks

	Good Looks	Bad Looks
	MEN	
SD Difference from Mean	0.830	-1.649
Specification No.		
1. (Total effect)	0.054	-0.129
2.	0.040	-0.087
3. (Indirect effect)	0.027	-0.061
	WOMEN	
SD Difference from Mean	0.709	-1.576
Specification No.		
1. (Total effect)	0.057	-0.151
2.	0.044	-0.108
3. (Indirect effect)	0.037	-0.077

Figure 1a. Effects of Beauty on Happiness/Satisfaction, Men, All Data Sets

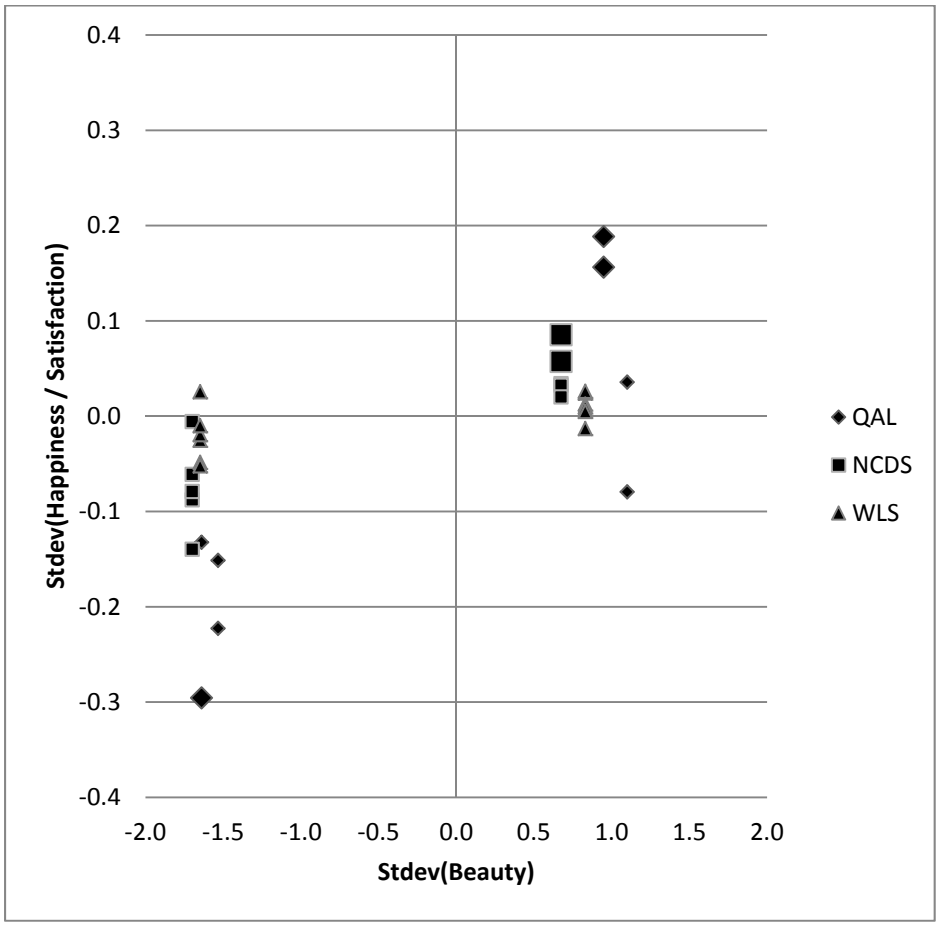

Figure 1b. Effects of Beauty on Happiness/Satisfaction, Women, All Data Sets