Beauty and the Labor Market

By Daniel S. Hamermesh and Jeff E. Biddle*

We examine the impact of looks on earnings using interviewers' ratings of respondents' physical appearance. Plain people earn less than average-looking people, who earn less than the good-looking. The plainness penalty is 5–10 percent, slightly larger than the beauty premium. Effects for men are at least as great as for women. Unattractive women have lower labor-force participation rates and marry men with less human capital. Better-looking people sort into occupations where beauty may be more productive; but the impact of individuals' looks is mostly independent of occupation, suggesting the existence of pure employer discrimination. (*JEL* J71, J10)

> He [Aristotle] used to say that personal beauty was a better introduction than any letter.
>
> —Diogenes Laertius, *The Lives and Opinions of the Eminent Philosophers* (ca. 200 A.D.)

Discrimination in the labor market has generated immense amounts of research by economists. Many alternative theoretical analyses of the nature of discrimination and a vast empirical literature have been produced (see e.g., Glen Cain's [1986] review). In the United States alone, careful empirical studies of possibly discriminatory outcomes involving blacks, Hispanics, women, linguistic minorities, physically handicapped workers, and no doubt others have been produced.[1] Our purpose here is to offer the first study of the economics of discrimination in the labor market against yet another group—the ugly—and its obverse, possible favoritism for the beautiful. We examine whether there is a reduced-form combination of attitudes toward beauty and a distribution of workers among jobs that generates apparently discriminatory labor-market outcomes.

This analysis is interesting in its own right. Every worker brings some physical attractiveness to the labor market along with other attributes, and most are concerned, perhaps inordinately so (Naomi Wolf, 1991), with this aspect of their labor-market characteristics. Interest in "lookism, ... the construction of a standard of beauty/attractiveness," is an expression of a belief that people failing to meet that standard are mistreated. Antidiscrimination legislation has been enacted in the United States to prevent denying employment on the basis of "height, weight and personal appearance," and proposed elsewhere on the basis of "facial features, build and height"; and in the United

*Hamermesh: Department of Economics, University of Texas, Austin, TX 78712, and National Bureau of Economic Research; Biddle: Department of Economics, Michigan State University, East Lansing, MI 48824. We thank Danny Blanchflower, George Borjas, Russell Cooper, Ronald Ehrenberg, Gerald Oettinger, Bob Reed, John Strauss, participants at seminars at the NBER and a number of universities, and two referees for helpful suggestions. We also thank the Interuniversity Consortium for Political and Social Research and the Institute for Social Research (York University) for making the data available. The order of the authors' names is inversely related to their looks.

[1]Examples of each are, respectively, Francine Blau and Andrea Beller (1992), George Borjas and Marta Tienda (1985), David Bloom and Gilles Grenier (1992), and Melissa Famulari (1992).

States a case law in this area is developing and may burgeon under the Americans with Disabilities Act of 1990.[2]

Studying possible discrimination on the basis of looks should also be of broader interest. It is very difficult to construct a research design that allows one to distinguish labor-market outcomes arising from discrimination against a group from those produced by intergroup differences in unobserved (by the researcher) productivity. In the case of looks we may have a better chance of doing so, for we can identify activities in which looks are likely to be more important, and thus where the payoff to beauty (or penalty for homeliness) reflects differences in productivity. In the literature on wage discrimination, attempts to sort out the importance of alternative sources of measured discrimination are quite rare (but see Alan Dillingham et al., 1994).

In Section I we examine some relevant results of social-psychological studies of beauty and human behavior, aiming toward considering whether it is possible to use measures of beauty as if they were objective descriptions. Section II discusses how beauty might be rewarded in the labor market and how it affects workers' choice of occupations. Section III describes the three sets of microeconomic data that we use to analyze the role of looks. Section IV tests for the presence of earnings differentials based on looks; Section V examines possible causes of male–female differences in the effects of beauty; and Section VI conducts tests aimed at distinguishing the sources of wage differences by looks.

[2]Quoted by Fred Siegel, "The Cult of Multiculturalism," *New Republic*, 18 February 1991, p. 38, from an official document from Smith College. The city of Santa Cruz, California, enacted and subsequently repealed an ordinance banning such discrimination (*New York Times*, 13 February 1992, p. A18). The foreign legislation was proposed in the Philippine Congress, reported by the Associated Press, 13 December 1992. The case law and the Americans with Disabilities Act are discussed by Tony McAdams et al. (1992). A recent case is *Hodgdon* v. *Mt. Mansfield Company*, 6 November 1992, in which the Vermont Supreme Court ruled that a chambermaid's lack of upper teeth qualified as a handicap protected under the state's Fair Employment Practices Act.

I. Background

If there is no common agreement on what constitutes beauty, it makes no sense to consider the role of looks in the labor market. Fortunately, a huge literature exists on this subject, including research by anthropologists, sociologists, and social psychologists, that has recently been ably summarized (Elaine Hatfield and Susan Sprecher, 1986). It seems quite clear that there are few consistent standards of beauty across cultures. Hugely distended lower lips are considered attractive by Ubangi men as were women's bound feet by Manchu dynasty men; and other less extreme examples of differences in standards of beauty across cultures could easily be cited.

What is perhaps a bit less obvious is that standards of beauty change over time within the same culture, changes that go beyond preferences and fads in clothing to the question of body type. The Rubens ideal looks much different from her Northern European counterpart walking down the runway at a modern Paris salon. Today's ideal lean Western male would have been viewed as potentially or actually consumptive and a bad match in both labor and marriage markets in 19th-century America. The crucial issue for our purposes is whether standards of attractiveness change slowly enough to allow labor-market decisions related to beauty to be planned for a horizon as long as a person's expected working life.

The evidence seems quite clear on this issue: within a culture at a point in time there is tremendous agreement on standards of beauty, and these standards change quite slowly. For example, respondents ranging in age from 7 to 50 who were asked to rank the appearance of people depicted in photographs showed very high correlations in their rankings. Moreover, the ratings of the appearances of a group of individuals photographed at different stages of their adult lives were highly autocorrelated (Hatfield and Spreecher, 1986 pp. 282–83). Today the same facial types are even pre-

TABLE 1—PERSISTENCE IN RATINGS OF BEAUTY, CANADIAN QUALITY OF LIFE, 1977, 1979, AND 1981 (PERCENTAGE DISTRIBUTIONS)

A. *Distribution of Ratings, 1977–1979 and 1979–1981 Combined*

Men ($N = 1{,}504$):

First-year rating	Second-year rating: 1	2	3	4	5
1) Strikingly handsome	0.2	0.9	1.0	0.0	0.0
2) Above average (good looking)	1.4	14.9	15.9	0.7	0.0
3) Average	0.9	15.1	37.5	4.8	0.1
4) Below average (plain)	0.1	0.4	4.0	1.7	0.1
5) Homely	0.0	0.1	0.1	0.1	0.1

1977–1979: $X^2_{[16]} = 151.78$ 1979–1981: $X^2_{[16]} = 142.67$

Women ($N = 2{,}147$):

First-year rating	Second-year rating: 1	2	3	4	5
1) Strikingly handsome	0.4	1.4	0.6	0.0	0.0
2) Above average (good looking)	1.0	14.3	15.8	1.0	0.0
3) Average	0.7	13.3	37.0	4.3	0.4
4) Below average (plain)	0.0	0.8	6.2	2.0	0.2
5) Homely	0.0	0.1	0.2	0.2	0.1

1977–1979: $X^2_{[16]} = 231.13$ 1979–1981: $X^2_{[16]} = 169.17.$

B. *Summary of Ratings Across Three Years*

Both Genders ($N = 1{,}330$):

	Absolute deviations from 1977 rating								
			1,1				2,2		
1977 rating	0,0	0,1	Same	Different	0,2	1,2	Same	Different	2,3
1) Strikingly handsome	0.1	0.2	0.8	0.0	0.1	1.1	0.2	0.0	0.1
2) Above average	8.1	13.2	10.4	0.6	0.6	0.9	0.2	0.0	0.0
3) Average	26.3	19.7	6.8	1.0	0.7	0.4	0.1	0.0	0.0
4) Below average	0.3	2.9	3.8	0.2	0.2	0.8	0.2	0.0	0.1
5) Homely	0.0	0.1	0.2	0.0	0.0	0.1	0.2	0.0	0.0
Total:	34.8	36.0	21.9	1.7	1.5	3.2	0.7	0.0	0.2

ferred by people of different races on different continents, perhaps because of the increasing internationalization of media images (*New York Times*, 22 March 1994, p. A6).

Some explicit evidence that, while "beauty is in the eye of the beholder," beholders view beauty similarly is provided by the tabulations in Table 1. This Canadian survey was conducted in 1977, 1979, and 1981, with different interviewers in each year asked to "categorize the respondent's physical appearance" into one of the five rubrics: strikingly handsome or beautiful; above average for age (good looking); average for age; below average for age (quite plain); and homely. The data have some aspects of a panel, so that many of the respondents were interviewed in two adjacent years, and some appear in all three years.

The matrices of ratings for pairs of adjacent years in the upper part of Table 1 are highly nonrandom, as shown by the chi-square statistics based on the contingency tables implicit in them. In each there is much more clustering along the prime diagonal than would arise randomly. The lower part of Table 1 provides information on the constancy of the interviewers' ratings over three biennia. Thirty-five percent of the sample is rated identically in all three years; and nearly 93 percent of the respondents are rated identically in at least two years and only one rating level different in the third year.[3] There is substantial positive correlation in how people rate others' looks.

There has been some examination of some of the labor-market correlates of beauty. The best of these is probably Robert Quinn (1978), who generated simple correlations of interviewers' ratings of the looks of respondents who were full-time employees with their incomes using one of the data sets we employ. Incomes were higher among both men and women the higher was the assessment of the respondent's looks, based on a three-point rating of beauty. The results held for both genders, and there was no evidence of asymmetry in the effect on income of departures from the middle category. A similar study (Patricia Roszell et al., 1989) used the Canadian data underlying Table 1 to regress 1981 income on 1979 income and a variable rating the respondent's looks, with results implying faster income growth among better-looking respondents.

Several studies have examined correlations of earnings with the appearance of workers in a narrow age or occupational cohort. A recent example is Irene Frieze et al. (1991), who studied earnings of MBA's over the first ten post-degree years. Ratings of beauty based on photographs of the students while in school were correlated positively with both starting and subsequent salaries for males. Among females there was no correlation with starting salary, but more attractive women experienced more rapid salary growth.[4]

A related larger literature has offered photographs and hypothetical résumés of potential workers and asked experimental subjects to choose among these workers for various jobs (Hatfield and Sprecher, 1986). Among men, beauty enhanced the worker's likelihood of being chosen for both clerical and professional/managerial jobs. Beauty helped the women's chances of being selected only for the higher-level clerical jobs.

We can be fairly sure that within the modern industrial world standards of beauty are both commonly agreed upon and stable over one's working life. The evidence also suggests that women's and men's beauty/ugliness might be treated differently in the labor market, so that any empirical study must analyze genders separately. Most important, an examination of the literature makes it clear that there has been little systematic thought about the role of beauty in the labor market; that the empirical analysis of this issue has almost exclusively dealt with very narrow samples; and that it has been limited to tabulations and regressions holding at most one or two variables (usually age) constant.

II. Models of Beauty in the Labor Market

One approach to modeling looks-based differences in labor-market outcomes is to assume that in at least some occupations attractive workers are more productive than unattractive ones. This advantage could arise from consumer discrimination, with customers preferring to deal with better-looking individuals; or there may be occupations in which physical attractiveness enhances the worker's ability to engage in productive interactions with coworkers. Prima facie evidence supporting this as-

[3]Given the distributions across the five categories in 1977 in this three-year sample, 20 percent would randomly be classified identically in all three years, and 79 percent would be randomly classified identically in two years and only one category different in the third. With a sample of 1,330 people, the probabilities of observing the outcomes in this part of Table 1 are infinitesimally tiny.

[4]An unpublished work in the late 1970's by Robert Frank of Cornell University correlated earnings of recent Cornell graduates with ratings of their appearance (from pictures) by a group of current undergraduates.

sumption is provided by a recent survey of employers (Harry Holzer, 1993), who were asked, "How important or unimportant is attractive physical appearance [for the job most recently filled]?" Eleven percent responded that appearance was very important, while 39 percent believed that it was somewhat important.

To explore the implications of such a model, consider an economy where each worker i is endowed with a vector of productivity-enhancing characteristics $\mathbf{X}_i$ and can be classified as either attractive or unattractive. In each of a number of occupations j the wage is given by

$$w_{ij} = \mathbf{a}_j \mathbf{X}_i + b_j \theta_i$$

where the $\mathbf{a}_j$ is a vector of parameters, b_j is positive in some occupations and zero in others, and θ_i equals 1 if the worker is attractive and 0 otherwise. Workers are assumed to choose the occupation offering the highest wage.

One set of empirical implications of this model involves the distribution of workers across occupations. There will be sorting, in that attractive workers will be observed in greater proportions in those occupations where attractiveness is rewarded. However, segregation by looks will be incomplete; both attractive and unattractive people might be found in any occupation. For example, unattractive workers may choose an occupation where attractiveness adds to productivity if they happen to be well endowed with other characteristics that are valued in that occupation. Likewise, an attractive worker might choose an occupation where attractiveness has no payoff if the choice provides a high relative reward for the worker's particular bundle of other characteristics.

A second set of implications concerns the earnings of attractive versus unattractive workers. If the distribution of $\mathbf{X}_i$ is uncorrelated with beauty, attractive workers will on average earn more, whether or not one controls for $\mathbf{X}$. Within occupations we will observe a difference between the average earnings of attractive and unattractive people only in those occupations where attractiveness is productive.

An obvious alternative to a model with productivity differences (including those associated with consumer discrimination) generating looks-based differences in outcomes has them resulting from employer discrimination against the unattractive. A Becker-type model involving employers' distaste for unattractive employees produces a looks differential in earnings, but no systematic sorting of workers into occupations on the basis of attractiveness. Further, there is no reason to expect the wage differences between attractive and unattractive workers to differ across occupations.[5]

It thus may be possible to distinguish empirically between a model with looks-based labor-market outcomes driven by productivity differences and one in which they arise because of employer discrimination. A practical obstacle to this task is identifying those occupations where attractiveness might plausibly lead to greater productivity. Assuming that a reasonable criterion for identification can be found, however, evidence that attractive people are more heavily represented in such occupations would support the productivity model.

Another test involves a regression like

$$(1) \quad w_i = \beta_0 + \boldsymbol{\beta}_1 \mathbf{X}_i + \beta_2 \theta_i + \beta_3 \mathrm{OCC}_i + \beta_4 \theta_i \mathrm{OCC}_i + \varepsilon_i$$

where $\mathrm{OCC}_i = 1$ if the worker's occupation has been identified as one where looks are productive, the ε_i are residuals, and the β's are parameters. This regression nests a simple view of occupational crowding, in which confining unattractive workers in certain

[5]Such a model would predict a disproportionate representation of attractive workers in certain industries (i.e., those shielded from competition). The literature on occupational crowding has often assumed that preference-based employer discrimination is occupation-specific, which in our case would imply that employers experience a visceral reaction only when contemplating the presence of an attractive or unattractive employee in certain occupations. It is hard to know how one might identify such occupations a priori (although the employee's physical proximity to the employer at work might be one factor).

occupations depresses the wages of all workers in those occupations and thus implies that $\beta_3 > 0$.[6] The productivity model implies $\beta_4 > 0$ and $\beta_2 = \beta_3 = 0$ (i.e., that the worker's looks matter only in those occupations where beauty is important). The employer-discrimination model implies that $\beta_2 > 0$ and $\beta_3 = \beta_4 = 0$.

The main focus of our empirical work is to determine whether standard earnings equations yield evidence of a pay difference based on looks. We then try to identify occupations where beauty might be productive in order to examine the extent of labor-market sorting by looks and to implement the tests that are implicit in (1).

III. Data

Two broad household surveys for the United States and one for Canada provide data on the respondents' looks as well as on the usual labor-market and demographic variables of interest to economists. The 1977 Quality of Employment Survey (QES) contains information on 1,515 workers. This survey has the advantage of including great detail about labor-market behavior, but the disadvantage of including only labor-force participants. The 1971 Quality of American Life survey (QAL) contains interviews of 2,164 respondents. For our purposes this study has the advantage of having substantial background information on the respondents, but the disadvantage of containing relatively few variables describing the worker's job. The 1981 Canadian Quality of Life study (QOL) contains 3,415 observations. This study has none of the disadvantages of the two American data sets and has the additional attraction of providing (for a much smaller subsample that constitutes a three-year panel) three observations on each respondent's looks.

In all three surveys, the interviewer, who visited the respondent in his or her abode, had to "rate [or categorize] the respondent's physical appearance" on the five-point scale shown in Table 1, along which looks range from strikingly handsome or beautiful to homely.[7] The distributions of the ratings in the three surveys are shown in Table 2. (For the Canadian data we present averages based on all the respondents included in the three-year study.) Among both men and women, roughly half are rated as average, and many more are rated above-average than are viewed as below-average. Either Canadians are better-looking than Americans, or Canadian interviewers (perhaps the populace generally) are less willing to describe someone as having below-average looks. What is most interesting is that the ratings of women are more dispersed around the middle category. This is a common finding in the social-psychological literature: women's appearances evoke stronger reactions, both positive and negative, than men's (Hatfield and Sprecher, 1986).

In these samples very few people are rated as strikingly beautiful (handsome) or as homely. We assign these to the nearest category and base all of our estimation on the three-category distinction among above-average, average, and below-average. Even this means that the cell sizes for some of the categories (e.g., people with below-average looks in the QAL) are not very large.

All three surveys offer a variety of measures of earnings. In all of them we chose to calculate hourly earnings as annual earnings

[6] It is not clear what an occupational-crowding model would imply about β_2 and β_4. The literature usually presumes that occupational segregation will be incomplete; but it has not produced a rigorous, canonical model that generates predictions about the relative wages of different types of workers in the same occupation.

[7] These are the only broadly based surveys we could find that contain information on looks and earnings. A number of other surveys, including one interesting proprietary data set used in a (racial) discrimination case by Mark Killingsworth, contain information on the worker's general appearance. This measure seems more likely to be influenced by income than the physical-appearance measures that are available in our samples.

TABLE 2—DISTRIBUTION OF LOOKS: QUALITY OF EMPLOYMENT SURVEY (QES), 1977; QUALITY OF AMERICAN LIFE, (QAL), 1971; CANADIAN QUALITY OF LIFE (QOL), 1977, 1979, AND 1981 (PERCENTAGE DISTRIBUTIONS)

	QES		QAL		QOL (pooled)	
Category	Men	Women	Men	Women	Men	Women
1) Strikingly beautiful or handsome	1.4	2.1	2.9	2.9	2.5	2.5
2) Above average for age (good looking)	26.5	30.4	24.2	28.1	32.0	31.7
3) Average for age	59.7	52.1	60.4	51.5	57.9	56.8
4) Below average for age (quite plain)	11.4	13.7	10.8	15.2	7.2	8.3
5) Homely	1.0	1.7	1.7	2.3	0.4	0.7
N:	959	539	864	1,194	3,804	5,464

divided by 52 times weekly hours.[8] In the analyses involving hourly earnings, all respondents who worked less than 20 hours per week and who earned less than \$0.75 per hour in the QAL (\$1 per hour in the QES and the QOL) are excluded, as are the self-employed individuals and all those for whom data on the various control variables are unavailable.[9] The empirical work includes only people aged 18–64.

Other variables defined for the analyses of hourly earnings and included in all three data sets are: marital status (which we measure as a zero–one dummy variable, married or not); education, defined as a vector of dummy variables measuring high-school completion, some college, or a college degree or more; and one-digit industry. Self-reported health status is included in all the regressions. Most important, anyone whose health status in the QES is listed as "totally and permanently disabled" or the next most severe category on a seven-point subjective scale is excluded from all the empirical work. In the QAL, a respondent is excluded if health "prevents him/her from doing lots of things," while in the QOL anyone whose self-reported health status is not at least rated as "fair" is excluded.[10] These exclusions minimize any spurious results stemming from a possible correlation between physical appearance and major physical disabilities that reduce productivity in the market.

Our purpose is to isolate the effect of beauty on earnings by controlling for as many other causes of variation in earnings as possible. Inferentially we are thus asking: what is the marginal effect of looks after accounting for all the other causes of variations in earnings that are usually measured? We define the set of regressors quite broadly and try to make them comparable across the three sets of data. In the QES and QOL the data allow the construction of actual labor-market experience, years of tenure with the firm, and an indicator of union status. In the former, establishment size is included, while the latter includes firm size. In the QAL, experience is measured as age $-$schooling-6. In estimates based on the two American data sets we include dummy

[8]All the equations were reestimated using annual earnings, with weekly hours included as an independent variable. None of our conclusions is changed qualitatively by this modification.

[9]Note that in 1971 in the United States the minimum wage was \$1.60 per hour, and in 1977 it was \$2.30. In Canada in 1981 the federal minimum was \$3.50, and some provincial minima were even higher. The disqualifications on the wage rate are thus designed to exclude those observations for which measurement errors are likely. Excluding the small fraction of workers whose estimated hourly wage is far below statutory minima does not imply any selectivity on a characteristic that is correlated with looks. In the QAL, for example, there is no relation at even the 20-percent level of significance between the beauty measures and the probability of exclusion from the sample for this reason. Even if there were, the fraction of people so excluded is below 5 percent of the sample.

[10]Of the respondents in the QES between the ages of 18 and 64 this disqualified 10; from the QAL, 126; and from the QOL, 18.

variables for race and for location in the South, while in the QOL we include a vector of variables for Canada's regions and an indicator of whether or not the person speaks English at home. Finally, the QAL allows us to include measures of the respondents' fathers' occupations, of their early childhood background, and of their immigrant status and that of their parents and grandparents.

IV. Looks and Earnings

The most interesting economic question involving beauty is probably its relation to an individual's economic success. In Section II we suggested three possible reasons for a premium for beauty or a penalty for ugliness in the labor market: pure employer discrimination, customer discrimination/productivity, and occupational crowding. In order to examine these we need to know first whether earnings differentials based on beauty even exist.

We make no claim to be able to estimate a structural model of a hedonic market for looks. Rather, in the first part of this section we present estimates of standard earnings equations that allow for the possibility of differences in earnings related to looks. In the final part we synthesize the findings to infer what we have learned from this approach about the existence of such earnings differentials. We consider whether such problems as unobservable influences on earnings are correlated with the measures of beauty; whether measurement error clouds our results; and how severe potential problems of simultaneity between earnings and beauty might be.

A. *Estimates of the Relationship of Looks and Earnings*

Columns (i) and (iii) of Table 3 present estimates of earnings equations based on the data from the QES. Columns (i) and (iv) of Table 4 do the same using data from the QAL, as do columns (i) and (v) of Table 5 for the QOL. In these and subsequent tables we present the probabilities (p) related to the F statistic testing the joint significance of the variables reflecting individuals' beauty.

TABLE 3—THE IMPACT OF LOOKS ON EMPLOYEES' EARNINGS: QES, 1977

	Men		Women	
Variable	(i)	(ii)	(iii)	(iv)
Looks:				
Below average	−0.164	−0.162	−0.124	−0.107
	(0.046)	(0.046)	(0.066)	(0.071)
Above average	0.016	0.010	0.039	0.035
	(0.033)	(0.034)	(0.048)	(0.049)
Obese		0.119		−0.122
		(0.172)		(0.134)
Overweight		−0.024		−0.016
		(0.038)		(0.058)
Tall		0.027		0.104
		(0.045)		(0.114)
Short		−0.105		−0.017
		(0.060)		(0.124)
$\bar{R}^2$:	0.403	0.404	0.330	0.327
p on F statistic for beauty variables:	0.001	0.001	0.069	0.173
N:	700	700	409	409

Notes: The dependent variable is log(hourly earnings); standard errors are shown in parentheses. The equations here also include continuous and indicator variables measuring actual experience (and its square), union membership, health status, marital status, race, years of vocational school, and region, and vectors of indicator variables for educational attainment, tenure with the firm, plant size, city size, and industry. The regressions exclude observations for which data were not available to form these measures and for which weekly hours worked < 20, hourly earnings < \$1, and age > 64 or age < 18.

Among the six equations, the pair of beauty variables is jointly significantly nonzero at some conventional level in four cases. Moreover, in all six groups people with above-average looks receive a pay premium, ranging from as little as 1 percent to a high estimate of 13 percent (for women in the QAL). In five groups (excluding only women in the QAL), workers with below-average looks receive a pay penalty, ranging from 1 percent to as much as 15 percent. Not all of these individual coefficients are significantly different from zero. However, many are, and the consistency of the

TABLE 4—THE IMPACT OF LOOKS ON EMPLOYEES' EARNINGS: QAL, 1971

	Men			Women		
Variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Looks:						
Below average	−0.078	−0.138	−0.079	0.069	0.122	0.061
	(0.069)	(0.081)	(0.069)	(0.073)	(0.095)	(0.073)
Above average	0.065	0.109	0.064	0.128	0.129	0.118
	(0.045)	(0.052)	(0.045)	(0.056)	(0.076)	(0.056)
Short			0.095			0.235
			(0.101)			(0.109)
Tall			0.018			0.251
			(0.066)			(0.214)
Interviewer effects:	no	yes	no	no	yes	no
$\bar{R}^2$:	0.371	0.471	0.370	0.283	0.332	0.293
p on F statistic for beauty variables:	0.124	0.014	0.130	0.072	0.174	0.108
N:	476	476	476	307	307	307

Notes: The dependent variable is log(hourly earnings); standard errors are shown in parentheses. Also included are continuous and indicator variables measuring experience (age − education − 6) and its square, health status, race, marital status, and region, and vectors of indicator variables measuring educational attainment, city size, rural background, immigrant status of the individual and his or her parents and grandparents, father's occupational status, and industry. The regressions exclude observations for which data were not available to form these measures and for which weekly hours worked < 20, hourly earnings < $0.75, and age > 64 or age < 18.

TABLE 5—THE IMPACT OF LOOKS ON EMPLOYEES' EARNINGS: CANADIAN QOL, 1981

	Men				Women			
Variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Looks below average:								
1981	−0.012	−0.027	−0.110		−0.058	−0.072	−0.042	
	(0.052)	(0.054)	(0.104)		(0.063)	(0.067)	(0.096)	
Average of three years				−0.148			−0.053	
				(0.172)			(0.120)	
Looks above average:								
1981	0.073	0.059	0.019		0.013	0.010	0.016	
	(0.028)	(0.030)	(0.056)		(0.027)	(0.029)	(0.039)	
Average of three years				0.123			0.068	
				(0.084)			(0.056)	
Interviewer effects:	no	yes	no	no	no	yes	no	no
$\bar{R}^2$:	0.302	0.306	0.222	0.228	0.394	0.389	0.487	0.491
p on F statistic for beauty variables:	0.023	0.099	0.498	0.147	0.540	0.492	0.821	0.348
N:	887	887	350	350	883	883	282	282

Notes: The dependent variable is log(hourly earnings); standard errors are shown in parentheses. Also included are continuous and indicator variables measuring actual experience and its square, health status, union status, non-English speaker, and marital status, and vectors of indicator variables measuring educational attainment, tenure with the firm, firm size, region, and industry. The regressions exclude observations for which data were unavailable to form these measures and for which weekly hours worked < 20, hourly earnings < $1, and age > 64 or age < 18.

pattern across three independent samples suggests that the finding of pay premia and penalties for looks is robust.

The estimates based on the QES indicate that more attractive people are paid more. However, the premia for good looks are considerably smaller than the penalties for bad looks and are not statistically significant. The results for men are corroborated by the QAL results in Table 4, with positive estimated coefficients for above-average looks categories and (larger) negative wage penalties for those in below-average looks categories. They are, however, contradicted by the estimates from the QOL in Table 5. In that sample there is a significant premium for good-looking men, but a tiny and insignificant penalty for men of below-average looks. A similar disagreement exists in the estimates for women. The large penalties for ugliness in the QES are replicated in the Canadian QOL, but are contradicted by a positive coefficient for below-average-looking women in the QAL. There are small premia for above-average-looking women in the QES and QOL, and a large significant premium in the QAL.

The similarity of the premia and penalties across the two genders is also interesting. In the results from the QES they are nearly identical. In the QAL there is a larger penalty for below-average-looking men than for women, but a larger premium for good-looking women. The opposite pattern holds in the QOL. Among people who choose to work at least half time, beauty does not generate very different effects on the earnings of women and men.

While the results are qualitatively similar in the three samples, one might worry still more about the robustness of the estimates. One concern is that each interviewer might have a different standard for beauty. These differences could be regarded as a form of measurement error, lowering the efficiency of our estimates and biasing them to the extent that interviewer standards were spuriously correlated with respondents' earnings. To account for any potential problems this might cause, columns (ii) and (v) of Table 4 and columns (ii) and (vi) of Table 5 reestimate these reduced-form earnings equations using interviewer-specific fixed effects for the QAL and QOL, respectively. Among men, the penalty for ugliness increases slightly in both samples; but the changes in the premium for good looks are in opposite directions. Among women the unexpected positive effect of below-average looks in the QAL becomes larger, but none of the other estimates of penalties and premia is affected much. Taken together, the results suggest clearly that the relation between looks and earnings does not arise from idiosyncratic ratings by particular interviewers.[11]

Another worry is about variables that are necessarily excluded from some or all of the samples because they are unavailable. Obviously, variables in the latter group cannot be examined here. But in the former group we can consider the impact of excluding the worker's family background and intelligence. Including the family background measures from the QAL, as in Table 4, lowered the absolute values of the estimated looks premia and penalties by less than 0.005 for men, and by less than 0.02 for women. Had we also included in columns (i) and (iv) of Table 4 a dummy variable for workers whose intelligence was perceived by the interviewer as being in the top 7 percent, the absolute values of the coefficients for men would fall by 0.002 each, and those for women would fall by 0.006 each. Despite the positive correlation between the subjective measures of intelligence and beauty, the changes are tiny. Adding father's and mother's educational attainment to family background measures in Table 4 alters the

[11] One related possibility is that interviewers of different sexes rate respondents differently. This possibility is also handled by using interviewer fixed effects. It is not likely to be a problem in any case, since 95 percent of the respondents in the two American samples were interviewed by women. A related problem is that there may be differences in the interviewers' ability to classify workers of different races. Not surprisingly, given that the overwhelming majority of the respondents are white, the estimates in Tables 3 and 4 change only minutely when African-Americans are deleted from the sample.

coefficients on the beauty measures by less than 0.001.

Beauty may alter other attributes people bring to the labor market that are not ordinarily considered in economic models. While these effects are difficult to measure, our data permit some exploration of this additional omitted-variable problem. The QOL asks respondents six questions designed to measure their self-esteem, with answers on a four-point scale indicating agreement/disagreement with statements such as, "Those who are always trying to get ahead in life will never be happy." A simple average of these six responses is (weakly) positively correlated with the three-category rating of individuals' looks; and the same measure generates significant positive coefficients when added to the equations underlying columns (i) and (v) in Table 5. It hardly alters the impacts of the beauty measures, however: In column (i) the estimates become -0.003 and 0.068, while in column (v) they become -0.053 and 0.014. Bad looks may produce low self-esteem before the person enters the labor market, and low self-esteem is associated with lower wages; but the measured direct impact of looks on wages is hardly affected by any pre-labor-market effects through self-esteem.

A long, large, and still growing literature (e.g., Paul Taubman, 1975; Robert McLean and Marilyn Moon, 1980; Susan Averett and Sanders Korenman, 1993) has studied the relation between weight or height and earnings. We can test whether our results merely demonstrate the effect on earnings of these few bodily characteristics by including measures of height and weight in the earnings equations. In the QES the interviewer rated the respondent's weight on a five-point scale and estimated the respondent's height in inches, while only height is available in the QAL.[12] For both samples we formed dummy variables based on height, categorizing women as tall if they exceeded 5′9″ (6′ for men) or short if they were below 5′ (5′6″ for men). Self-explanatory dummy variables for people who are obese, or only overweight, were constructed for the QES sample.

The results of adding these measures to the earnings regressions are shown in columns (ii) and (iv) of Table 3 and columns (iii) and (vi) of Table 4. Other than wage premia for both short and tall women in the QAL and a penalty for short men in the QES, none of these variables has a coefficient that exceeds its standard error. Most important, including these measures of body type has only a small effect on the coefficients on the ratings of beauty in all four samples—much too small to suggest that the relationship between looks and earnings arises from correlations between appearance and height or weight.

The Canadian data allow us to examine the effect of the measurement error associated with using only one rating of the beauty of each respondent. For the subset of respondents included in the bottom part of Table 1 the study provides three independent estimates of beauty. One approach to using this information would create a set of dummy variables for each of the ten combinations of looks ratings based on the threefold classification for each of the three years. This has difficulties in that it produces a few very sparsely occupied cells and generates a different metric from the other results in Tables 3–5. An alternative, very simple approach averages the dummy variables for above- and below-average looks for each year. Thus, for example, a person who is rated above-average in all three years would have a value of 1 for the combined dummy variable indicating above-average looks and 0 for the below-average variable; for someone rated below-average in one year and average in the other two, the above-average variable equals 0, while the below-average variable equals one-third.

Columns (iii) and (vii) of Table 5 present estimates of the same equations as in columns (i) and (v), but now based on the smaller longitudinal sample. Columns (iv)

[12] The rating scale for weight (in descending order) was: "obese," "overweight," "average for height," "underweight," and "skinny." Among women (men), 3.2 (0.7) percent were rated obese, 19.6 (17.4) percent were rated overweight, 65.8 (72.7) percent were considered average, 11.2 (8.5) percent were rated underweight, and 0.2 (0.7) percent were rated skinny.

TABLE 6—STACKED ESTIMATES OF THE IMPACT OF LOOKS ON HOURLY EARNINGS

Sample	Penalty for below-average looks	Premium for above-average looks	$\hat{\beta}_{\text{above}} - \hat{\beta}_{\text{below}}$	p on F statistic for looks	p on intersample equality of looks effects
Men:					
All three samples	−0.091	0.053	0.144	0.0001	0.246
	(0.031)	(0.019)	(0.040)		
Two U.S. samples	−0.132	0.036	0.168	0.0003	0.443
	(0.039)	(0.027)	(0.051)		
Women:					
All three samples	−0.054	0.038	0.092	0.042	0.163
	(0.038)	(0.022)	(0.048)		
Two U.S. samples	−0.042	0.075	0.117	0.041	0.123
	(0.049)	(0.037)	(0.069)		
Men and women combined:					
All three samples	−0.072	0.048	0.120	0.0001	0.106
	(0.024)	(0.015)	(0.031)		
Two U.S. samples	−0.092	0.046	0.138	0.0002	0.051
	(0.031)	(0.022)	(0.041)		

Notes: The dependent variable is log(hourly earnings); standard errors are shown in parentheses.

and (viii) replace the one-year dummy variables with the three-year averages. This substitution adds to the significance of the equations for both men and women. Moreover, all four estimated coefficients increase in absolute value, as we would expect if each year's rating contained some degree of measurement error.[13] Obtaining additional information on a worker's beauty provides additional information about his or her earnings.

[13]There are other ways of combining the three ratings. For example, assume that each interviewer assigns a rating along the five-point scale based on her estimate of underlying beauty, B. For a homely person, for example, the data in Table 2 imply that the person is in the lowest 2.5 percent of the population. Assuming that B is normally distributed, the best estimate of that person's B is $\hat{B} = E(B|B < N^{-1}(0.025))$. Similar inferences can be drawn based on partitioning the normal density for each of the other ratings using the population percentages in Table 2. An estimate of a respondent's true beauty is B^*, the average of the three independent estimates of $\hat{B}$. Using B^* rather than $\hat{B}$ as a measure of beauty generates improvements in goodness of fit and increases in the absolute values of estimated coefficients similar to those associated with columns (iii) and (iv) and with columns (vii) and (viii).

B. *Synthesis of the Basic Results, Some Criticisms, and an Initial Interpretation*

Tables 3–5 stand on their own and provide the basic evidence for the existence of earnings differentials based on beauty. Nonetheless, it is useful to summarize the results in order to infer what the three sets of data imply are the best estimates of the penalties and premia associated with looks, especially since the individual sets of data are relatively small.

Table 6 presents these summaries for each gender separately and for the entire set of observations, and for all three samples combined and for the two U.S. samples alone. The estimates are from regressions that pool the samples in Tables 3–5 (or Tables 3 and 4 only) and that allow the coefficients on all variables other than the beauty measures to differ across the samples (i.e., analyses that "stack" the regressions). The last column shows that constraining the estimated effects of beauty on earnings to be the same across samples for men and women separately is not rejected by the data; and for each gender both the earnings penalty and premium are significantly nonzero. Indeed, even constraining the effects to be the same for both genders for all three samples is not rejected; and the penalties and premia in both sets of pooled equations are all significantly nonzero.

The results make it clear that there is a significant penalty for bad looks among men. The 9 percent of working men who are viewed as being below average or homely are penalized about 9 percent in hourly earnings, other things equal. The 32 percent who are viewed as having above-average looks or even as handsome receive an earnings premium of 5 percent. Among women there is some evidence of a premium for good looks, with an average effect of about 4 percent; the penalty for bad looks (for the lowest 8 percent of working women) is 5 percent. Among women, neither effect alone is highly significant, though they are jointly significant. The combined results in the bottom two rows suggest a 7–9-percent penalty for being in the lowest 9 percent of looks among all workers, and a 5-percent premium for being in the top 33 percent. While the absolute values of the point estimates of the penalties generally exceed the estimates of the earnings premia, these differences are not significant. There is only weak evidence of asymmetry in how the labor market treats ugliness and beauty.[14]

The third column in Table 6 combines the premia and penalties from these stacked regressions to estimate the hourly earnings gain to moving from below- to above-average looks. The estimate of 0.120 for the three samples including both men and women is equivalent to the effects on earnings in these (and most other studies) of an extra 1.5 years of schooling. Viewed differently, moving from average to below-average looks would shift the worker from the median of the distribution of earnings to the 43rd percentile; moving to above-average looks would shift him or her to the 53rd percentile. Clearly, while the impacts on earnings of differences in looks are not so great as those of differences in gender, education, or race, they are not trivial.

No doubt there are unobserved factors that might affect productivity and be correlated with looks. For example, greater attractiveness and higher earnings in adulthood may be joint products of a privileged family background. Only the QAL contains variables that allow us to attempt to control for such effects. If family background in general were important, one would expect these partial indicators of it to have a noticeable effect on the estimates. We saw in the previous subsection that they did not, suggesting that the unobservable background measures are unlikely to be biasing our results seriously.[15] This observation and the robustness of the estimated effects of beauty suggest, though they do not prove, that adding still more variables to the list is not likely to alter our conclusions.

There are also potential simultaneity problems with the results. One might argue that they may merely show that the unobserved determinants of productivity generate extra earnings that are used to improve a worker's beauty. This is the conventional problem associated with any hedonic estimation (i.e., holding constant the observables, people with higher wages will choose to invest more in beauty). Alternatively, perhaps the interviewers in these data sets

[14]Remember that hourly earnings were calculated using actual weekly hours, but assuming that all workers spent the same number of weeks employed. The QES and QOL provide data on weeks of layoff (in the last year in the QES, two years in the QOL). We estimated Tobit regressions of the determinants of weeks of layoff (for the roughly 7 percent of males who reported having been laid off) including controls for education level, experience, union status, tenure with the firm, and firm or establishment size. In both samples the t statistics on the dummy variable for above-average looks were below 0.5 in absolute value. Bad looks raised the probability of layoff and lengthened its duration, with t statistics of 1.54 in the QES (1.40 in the QOL). This provides additional evidence for the conclusion that there is some asymmetry in the effect of looks on earnings. However, note from Table 2 that below-average looks are much less frequent than above-average looks in these ratings, so that any asymmetry in our results may be due more to how beauty is rated than to how the market treats beauty.

[15]We are indebted to Bob Willis for suggesting this point.

subconsciously bias their ratings of the respondents' beauty because they know, or can intuit, the respondent's earnings.

Three pieces of evidence suggest that these simultaneity problems are not crucial here. First, the social-psychological evidence we mentioned in Section I showed how little individuals' relative physical appearances change during adulthood. This implies that there is limited scope for using unexplained earnings differences to "buy" differences in beauty. Second, if differences in unexplained earnings were used to affect beauty, their persistence over a working life should lead to a greater simultaneity bias among older workers than among younger workers, and thus smaller apparent penalties and premia if we restrict the samples in Tables 3–5 to workers aged 18–30. In fact, all beauty premia and penalties in the QES are larger in this subsample than in the basic estimates in Table 3. In the other two samples, half the estimates increase in absolute value, while half decrease. There is no evidence of a weaker relation between earnings and beauty among younger workers.

The third bit of evidence addresses the potential problem of interviewers assigning higher ratings to more prosperous respondents. Using the longitudinal data on which columns (iii) and (iv) and columns (vii) and (viii) of Table 5 are based, we replace the three-year averages of the dummy variables with averages only of the 1977 and 1979 ratings of beauty. If there is a problem of reverse causation from 1981 earnings, it should be less severe when these instruments for beauty in 1981 are used. The estimates become -0.076 and 0.138 for men, and -0.027 and 0.071 for women. For both genders the $\bar{R}^2$ values increase by 0.001 compared to the estimates in columns (iv) and (viii). This standard simultaneity correction does not alter our basic results.[16]

All of these tests reinforce the conclusion that, whatever the causes, people who are better-looking receive higher pay, while bad-looking people earn less than average, other things equal. It is crucial to stress that these penalties and premia reflect the effects of beauty in all its aspects, not merely one of its many components, such as facial structure, bearing, height, weight, or complexion.

V. The Absence of Differences by Gender

Particularly surprising in light of some popular discussion (e.g., Wolf, 1991) is the absence of significantly larger penalties and premia, especially the latter, for women than for men. If anything, the evidence goes in the opposite direction: men's looks may have slightly larger effects on their earnings than do women's. One simple explanation might be that our results are a statistical artifact produced because the beauty ratings are a noisier signal of women's physical appearance than of men's. The evidence contradicts this: in the longitudinal part of the QOL the beauty ratings of women are slightly *less* variable over the three years than those of men.

One way that beauty can affect women's labor-market success is by influencing their labor-force participation. To examine this possibility we estimate probits relating participation to measures of attractiveness for married women in both the QAL and the QOL, and in the longitudinal subsample of the QOL. The coefficients on the beauty measures are shown for the QAL in column (i) and for the QOL in columns (ii)–(iv) of Table 7. Except when we use the three-year average ratings of beauty in the QOL, the t statistics on the above-average looks ratings are tiny, and the coefficients are always nearly zero. There is only very weak evidence that good-looking women are more likely to be in the labor force than otherwise identical average-looking women.

[16] The same conclusion is reached if we replace 1981 ratings of beauty with ones predicted from regressions using all the information contained in the 1977 and 1979 ratings. Despite this evidence, one might still argue that serial correlation in earnings creates a simultaneity between current earnings and lagged ratings of beauty. Under usual assumptions about serial correlation in earnings, however, one would not find, as we do, that the results using the 1977 and 1979 ratings in the small longitudinal sample are at least as strong as those using the 1981 data only.

TABLE 7—THE IMPACT OF LOOKS ON MARRIED WOMEN'S LABOR-FORCE PARTICIPATION (QAL, 1971; QOL, 1981) AND ON HUSBAND'S EDUCATION (QES, 1977)

	Probits of participation				Regression of husband's education, QES
	QAL	QOL			
Variable	(i)	(ii)	(iii)	(iv)	(v)
Looks below average:					
1971 (or 1981 or 1977)	−0.168	−0.310	−0.429		−1.043
	(0.176)	(0.153)	(0.245)		(0.369)
Average of three years in the QOL				−0.206	
				(0.318)	
Looks above average:					
1971 (or 1981 or 1977)	−0.034	−0.010	0.020		0.077
	(0.131)	(0.078)	(0.115)		(0.308)
Average of three years in the QOL				0.245	
				(0.169)	
Pseudo-R^2 or $\bar{R}^2$:	0.148	0.067	0.082	0.082	0.402
Mean of dependent variable:	0.401	0.524	0.514	0.514	12.63
N:	583	1,287	603	603	199

Notes: Numbers in parentheses are standard errors. In the QAL, the dependent variable equals 1 if the women was employed at the time of the interview. In the QOL, it is whether she stated she was in the labor force on the interview date. Also included in the probits in both samples are indicator variables measuring educational attainment, health status, and age. In the probits based on the QAL, indicator variables for race and the age of the youngest child are also included, as is a measure of family income less the woman's income. In probits based on the QOL, indicator variables describing the number of children are included. In the regression on husband's education from the QES, his age and the wife's educational attainment, age, and health status are also included in the regression.

The effects of below-average looks on women's participation are negative (though insignificantly so) in the QAL; and in the QOL these effects are significantly negative when the current rating of beauty is used (and insignificantly negative when we use the three-year average). The effects are not small. In the QAL the 6 percent of married women with below-average looks are 3-percent less likely to participate than are above-average-looking women. In the QOL the difference in participation rate is 8 percent based on the estimates in column (iii) and 11 percent based on the estimates in column (iv) (again illustrating how using several years of ratings of beauty reduces potential downward biases arising from measurement error).[17]

[17]Not surprisingly, similar probits on men's labor-force participation yielded no relationship between looks and the probability of participation. These results and those for women are qualitatively the same when we use linear regressions instead of probits to describe participation.

There is thus some evidence that women select themselves out of the labor force if they are particularly unattractive. However, this selectivity has no important impact on the basic estimates of the effects of looks on earnings [in column (iv) of Table 4 and column (v) of Table 5]. Correcting for selectivity in the QAL changes the estimated premium associated with above-average looks from 0.128 to 0.130. Accounting for this form of selectivity does not alter the premium in the QOL and changes the earnings penalty from −0.058 to −0.036.

Another possibility is that looks affect women's economic success by altering their opportunities for marriage. Holding constant a woman's age and educational attainment, in all three samples her looks are completely unrelated to her likelihood of being married. They are, however, related to the quality of the husband whom she marries. We use data on husband's education in the QES to estimate regressions that include our standard pair of measures of looks of the married woman (and also her husband's age and her health status, age, and education, to account for assortative mating).[18]

The results, presented in column (v) of Table 7, also show that above-average looks have essentially no effect on the outcome, in this case on the quality of the husband to whom the woman is matched. However, all else equal, below-average-looking women marry men whose educational attainment is one year less than what the women's own characteristics, including her educational attainment, predict.[19] Women face an additional economic penalty for bad looks in the form of marriage to husbands whose potential earnings abilities are lower.

The results show that the economic penalties facing below-average-looking women are not limited to hourly earnings. Both their success in the marriage market and their likelihood of working outside the home are reduced by their bad looks. No such effects exist for below-average-looking men; and there is no apparent premium in the marriage market or extra effect on participation for either good-looking women or men.

VI. Sorting, Productivity or Discrimination?

Having demonstrated that the labor market does reward beauty, we now consider the sources of the penalties and premia. The discussion in Section II suggested that to examine these issues we must learn how workers are sorted into occupations and discover how the earnings regressions of Tables 3–5 are affected when the model in (1) is estimated.

A test for sorting requires prior determination of the occupations where looks are likely to enhance productivity. In the absence of a widely accepted objective measure for determining this, we use three independent subjective methods. The first is based on the *Dictionary of Occupational Titles* (DOT) (1977). We assign each worker to a DOT occupation using three-digit occupational codes in both the QES and the QAL and note the DOT measure of "the job's relationship to people." Since physical attractiveness can affect productivity through the worker's interactions with customers or coworkers, we classify jobs with DOT measures that suggest an important role for interpersonal communication as ones where looks are important.[20]

The second method relies on the opinions of eight adults with at least one year of full-time labor-market experience who were asked to rate each of the three-digit occupations on a three-point scale: 0, looks are probably not important; 1, looks might be important; and 2, looks are definitely important.[21] If the average rating of the occupation exceeds 0.5, we treat looks as being important in the occupation and form a dummy variable reflecting this average of the subjective ratings.

The third measure uses a survey (Holzer, 1993) of employers' views of the importance of an applicant's appearance in filling the most recent job vacancy. The vacancy's occupational category was also recorded, as was the gender of the applicant hired. We first divided the survey data on the basis of the gender of the worker hired, then com-

[18] The underlying data on education are listed in seven categories, not single years of schooling. We assign years of schooling to these categories (5, 8, 10, 12, 14, 16, and 17) and base the regressions on these. Ordered probits based on the seven categories yield the same qualitative conclusions.

[19] Regressions for the education of wives in the QES generated estimated effects of -0.11 and 0.13 years on the dummy variables indicating their husbands' looks, with t statistics below 0.8 in absolute value. The sorting of economic outcomes in marriage appears to be related to beauty only for women.

[20] We rely on the fifth digit of the DOT code, which can take nine different values according to whether the job involves "mentoring," "negotiating," "instructing," "supervising," "diverting," "persuading," "speaking, signaling," "serving," or "taking instructions, helping." We treat all but the last as indications that interpersonal interaction is an important aspect of the occupation.

[21] The 28 pairwise correlations of the ratings of the 504 occupations ranged from 0.36 to 0.61, with a mean of 0.47.

piled for each gender a list of occupations that seemed fairly homogeneous with respect to the importance of appearance and for which there were at least ten observations. For each occupation/gender cell we calculated the percentage of employers responding that appearance was very important or somewhat important and matched these percentages, where possible, to workers from the QES and the QAL.[22]

To split the samples roughly in half, for men we define an occupation as one with "looks important" if more than 40 percent of the employers responded that appearance was important; for women the dividing line is 44 percent. In general, occupations with higher percentages have more contact between workers and customers: sales occupations top the list for men; for women looks are deemed most important in hiring cashiers, receptionists, and waitresses.[23]

If workers sort themselves among occupations/employers based in part on the relative productivity of their beauty, we should observe the highest average rating of individuals' looks in those occupations where our indexes suggest that looks matter most. Table 8 presents the fractions of workers in each of the three categories of individuals' looks who work in occupations where looks are important. With three rating schemes for the occupations, two samples, and both genders, we have constructed 12 tests for occupational sorting. Formal tests for sorting yield significant chi-square statistics in only four of the 12 rows. A good way to summarize the results is that all three rating schemes yield a significant relationship between our measures of the importance of beauty in an occupation and the beauty of workers in that occupation in the QAL, but not in the QES. However, in seven of the 12 rows the percentage of workers in jobs where looks are important increases monotonically along the scale of individuals' looks. More important, in ten of them, above-average-looking people are the most likely to be working in occupations where looks are important.

The results in Table 8 provide some evidence of sorting across occupations by beauty, but it is certainly not strong enough to suggest that occupational crowding is a major factor explaining the looks differential in earnings. It is unclear whether the weakness of the evidence is due to imperfections in our proxies for differences in the importance of beauty among occupations or to the relatively minor role that sorting by beauty plays.

Following (1), we augment the earnings regressions of Tables 3 and 4 with a dummy variable signifying whether or not looks are important in an occupation and with interactions between this variable and the two dummy variables indicating the individual's own looks. As in Table 8, we base the results on all three measures of occupational beauty. An attempt to capture the spirit of occupational crowding would predict that the occupational dummy variable will have a significant coefficient. A model based on the productivity of beauty in certain occupations implies that the interaction terms will capture the looks differential. The employer-discrimination model predicts that coefficients on all of these additional terms will equal zero, but that individuals' own beauty will affect their wages regardless of occupation.

The results of this test are shown in Table 9, which presents equations analogous to those in columns (i) and (iii) of Table 3 [columns (i) and (iv) of Table 4]. For the DOT and subjective measures, the samples are identical to those used in Tables 3 and 4. The coefficients on the main effects representing the respondents' own beauty are not greatly different from what they were in those tables; and the p values on the F statistics testing the pair of variables also differ little from the corresponding estimates in those tables. Even holding constant occupational beauty, below-average-

[22]The survey targeted employers of low-education workers. As a result there were too few observations in several broad occupation cells to calculate occupational beauty ratings, preventing many QES and QAL sample members from being included in this part of the analysis.

[23]Complete information on the occupational rankings is available upon request from the authors.

TABLE 8—OCCUPATIONAL SORTING: PERCENTAGE OF SAMPLE IN OCCUPATIONS WITH LOOKS IMPORTANT

	Own looks					
Looks are important	Below average	Average	Above average	Total	X^2	N
QES, men:						
DOT	62.6	63.5	64.7	63.7	0.14	700
Subjective	13.2	13.3	11.1	12.7	0.65	700
Employer	46.5	52.2	44.3	49.3	2.14	428
QES, women:						
DOT	76.4	76.2	80.9	77.8	1.16	409
Subjective	21.8	26.2	28.7	26.4	0.96	409
Employer	45.9	45.2	47.1	45.9	0.10	309
QAL, men:						
DOT	40.0	55.6	64.5	56.9	9.00	476
Subjective	17.8	12.9	22.4	16.4	6.50	476
Employer	33.3	61.2	63.3	59.3	7.48	268
QAL, women:						
DOT	67.4	73.9	81.1	75.6	3.61	307
Subjective	34.9	35.3	40.5	37.1	0.87	307
Employer	44.1	44.5	62.6	51.1	8.30	270

Note: Critical $\chi^2_{[2]}$ values are 5.99 (5-percent level of significance) and 4.60 (10-percent level).

looking workers receive substantial penalties (except, as before, for women in the QAL), and above-average-looking workers receive earnings premia (especially women in the QAL). In the samples using the employer-based estimates of occupational looks, which contain roughly 40-percent fewer observations, the effects of the workers' own looks are significant at least at a low level in three of the four cases.

The main effects of occupational looks exceed their standard errors in six of the 12 equations. The coefficients on the interaction terms exceed their standard errors in ten of the 24 cases. The $\bar{R}^2$ values here are higher for the QES men, lower for the QES women, and higher in one case, lower in the other for both QAL samples than in Tables 3 and 4, while in the reduced samples using the employer-based indexes the $\bar{R}^2$ values are increased in three of the four cases.[24]

Taken together, the estimates provide a hint that occupational requirements for beauty may produce independent effects on earnings; but we cannot reject the possibility that they have no effect.

This final exercise demonstrates one thing very clearly: the effects of an individual's own looks on his or her earnings are very robust. That there are earnings premia and penalties for looks independent of occupation suggests that employer discrimination on the basis of looks may lie behind those premia and penalties. That there is some evidence of sorting implies that pure em-

[24]A more straightforward test simply includes a vector of dummy variables for one-digit occupations in the basic equation for both samples and genders. The coefficients on the dummy variables for below- and above-average looks are hardly altered in size or significance. Among the QES men (women), the coefficients [analogous to those in columns (i) and (iii) of Table 3] become −0.156 and 0.014 (−0.100 and 0.026). Among the QAL men (women), the coefficients [analogous to those in columns (i) and (iv) of Table 4] become −0.059 and 0.062 (0.068 and 0.115). Taking this approach to its logical extreme (and losing between one-fourth and one-half of the degrees of freedom in each model), we reestimated the equations with separate dummy variables for each three-digit occupation. The results for men in both samples are essentially unchanged in this extension; for women the parameter estimates maintain their signs, but their absolute values are cut in half.

TABLE 9—SORTING, LOOKS, AND THE DETERMINATION OF EARNINGS: QES, 1977; QAL, 1971

Sample and occupation index	Looks below average	Looks below average × occupation index	Looks above average	Looks above average × occupation index	Occupation index	$\bar{R}^2$	p on F statistic on main effects
QES, men:							
DOT	−0.177	−0.036	0.041	0.072	0.052	0.405	0.002
	(0.058)	(0.095)	(0.042)	(0.069)	(0.041)		
Subjective	−0.162	0.007	0.012	0.051	0.124	0.405	0.003
	(0.049)	(0.127)	(0.035)	(0.097)	(0.072)		
Employers	−0.187	−0.112	−0.095	0.103	−0.066	0.410	0.026
	(0.076)	(0.107)	(0.057)	(0.084)	(0.049)		
QES, women:							
DOT	−0.174	−0.218	0.023	−0.068	0.032	0.329	0.036
	(0.075)	(0.157)	(0.054)	(0.119)	(0.085)		
Subjective	−0.115	−0.037	0.050	−0.036	0.083	0.326	0.130
	(0.074)	(0.151)	(0.055)	(0.096)	(0.093)		
Employers	−0.078	−0.013	0.152	−0.312	0.216	0.315	0.064
	(0.107)	(0.158)	(0.076)	(0.111)	(0.077)		
QAL, men:							
DOT	−0.102	−0.057	0.070	0.011	0.093	0.373	0.224
	(0.107)	(0.142)	(0.056)	(0.089)	(0.055)		
Subjective	−0.097	0.078	0.045	0.089	0.085	0.371	0.223
	(0.076)	(0.177)	(0.048)	(0.099)	(0.102)		
Employers	0.145	−0.107	0.124	−0.072	−0.006	0.213	0.449
	(0.150)	(0.250)	(0.121)	(0.152)	(0.095)		
QAL, women:							
DOT	0.049	−0.056	0.166	0.175	−0.066	0.282	0.031
	(0.088)	(0.159)	(0.063)	(0.130)	(0.088)		
Subjective	0.130	−0.172	0.075	0.142	−0.053	0.287	0.266
	(0.090)	(0.152)	(0.068)	(0.099)	(0.099)		
Employers	0.253	−0.304	0.261	−0.355	0.218	0.272	0.058
	(0.153)	(0.229)	(0.127)	(0.162)	(0.117)		

Notes: The dependent variable is log(hourly earnings); standard errors are shown in parentheses. Each regression includes the same additional variables as in the corresponding regression in Table 3 or 4. Those using the occupational indexes based on the DOT and subjective measures also use the same samples. Those using the survey of employers are based on smaller samples: $N = 428$, 309, 265, and 259.

ployer discrimination alone does not describe the role of beauty in the labor market; beauty may be productive in some occupations perhaps as a result of customers' preferences.

VII. Conclusions and Implications

In separate empirical analyses using three sets of household data, we find some evidence of a positive impact of workers' looks on their earnings. The evidence in each sample alone is suggestive but not very strong. When the three samples are combined, however, sample sizes become sufficient to make some fairly clear inferences about the role of beauty in the labor market. Other things equal, wages of people with below-average looks are lower than those of average-looking workers; and there is a premium in wages for good-looking people that is slightly smaller than this penalty. The penalty and premium may be higher for men, but these gender differences are not large. There is also some evidence that the labor market sorts the best-looking people into occupations where their looks are productive.

It is difficult to disentangle the effects of alternative sources of earnings differentials in the data. Nonetheless, our finding that earnings penalties and premia are essentially unaffected when we account for workers' occupations suggests no support for a model of occupational crowding along the dimension of beauty. That there is some occupational sorting by looks provides support for productivity-related discrimination; but the evidence is fairly weak. A related explanation, that there are inherent productivity differences that we do not capture because of omitted variables, cannot be ruled out, though there is some evidence against it. The strongest support is for pure Becker-type discrimination based on beauty and stemming from employer/employee tastes. More light could be shed on these questions by comparative examinations of the relationship between looks and earnings within particular narrowly defined occupations.

Our demonstration shows the magnitude of the incentives that the labor market in North America provides to expend resources on beauty and the mechanisms by which those incentives arise. Whether the same incentives exist in other economies is an obvious topic of interest. The results also lead naturally to further examination of the sources of wage differentials and possible discrimination along various other dimensions, such as physical and mental handicaps. In each case, the method we have developed to aid in distinguishing between productivity/discrimination and occupational sorting can be applied mutatis mutandis to discover the source of other apparently discriminatory outcomes.

REFERENCES

Averett, Susan and Korenman, Sanders. "The Economic Reality of 'The Beauty Myth'." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 4521, 1993.

Blau, Francine and Beller, Andrea. "Black–White Earnings over the 1970s and 1980s: Gender Differences and Trends." *Review of Economics and Statistics*, February 1992, *74*(1), pp. 276–86.

Bloom, David and Grenier, Gilles. "Earnings of the French Minority in Canada and the Spanish Minority in the United States," in Barry Chiswick, ed., *Immigration, language, and ethnicity*. Washington, DC: American Enterprise Institute Press, 1992, pp. 373–409.

Borjas, George and Tienda, Marta, eds. *Hispanics in the U.S. economy*. New York: Academic Press, 1985.

Cain, Glen. "The Economic Analysis of Labor Market Discrimination: A Survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of labor economics*. Amsterdam: North-Holland, 1986, pp. 693–785.

Dictionary of Occupational Titles, 4th Ed. Washington, DC: U.S. Government Printing Office, 1977.

Dillingham, Alan; Ferber, Marianne and Hamermesh, Daniel. "Gender Discrimination by Gender: Voting in a Professional Society." *Industrial and Labor Relations Review*, July 1994, *47*(4), pp. 622–33.

Famulari, Melissa. "The Effects of a Disability on Labor Market Performance: The Case of Epilepsy." *Southern Economic Journal*, April 1992, *58*(4), pp. 1072–87.

Frieze, Irene; Olson, Josephine and Russell, June. "Attractiveness and Income for Men and Women in Management." *Journal of Applied Social Psychology*, July 1991, *21*(3), pp. 1039–57.

Hatfield, Elaine and Sprecher, Susan. *Mirror, mirror*...: *The importance of looks in everyday life*. Albany, NY: State University of New York Press, 1986.

Holzer, Harry. "Multi-City Study of Urban Inequality." Unpublished mauscript, Michigan State University, 1993.

McAdams, Tony; Moussavi, Farzad and Klassen, Michael. "Employee Appearance and the Americans with Disabilities Act: An Emerging Issue?" *Employee Responsibilities and Rights Journal*, December 1992, *5*(4), pp. 323–38.

McLean, Robert and Moon, Marilyn. "Health, Obesity, and Earnings." *American Journal of Public Health*, September 1980, *70*(9), pp. 1006–9.

Quinn, Robert. "Physical Deviance and Occupational Mistreatment: The Short, the Fat and the Ugly." Unpublished manuscript, Institute for Social Research, University of Michigan, 1978.

Roszell, Patricia; Kennedy, David and Grabb, Edward. "Physical Attractiveness and Income Attainment Among Canadians." *Journal of Psychology*, November 1989, *123*(6), pp. 547–59.

Taubman, Paul. *Sources of inequality in earnings*. Amsterdam: North-Holland, 1975.

Wolf, Naomi. *The beauty myth*. New York: Anchor, 1991.