Cornell University
ILR School

ILRReview

Volume 60 | Number 1 Article 1

10-1-2006

Estimating the Effect of Personality on Male and Female Earnings

Gerrit Mueller
Centre for European Economic Research

Erik J.S. Plug
University of Amsterdam

Follow this and additional works at: http://digitalcommons.ilr.cornell.edu/ilrreview

Estimating the Effect of Personality on Male and Female Earnings

Abstract
The authors adopt the Five-Factor Model of personality structure to explore how personality affected the earnings of a large group of men and women who graduated from Wisconsin high schools in 1957 and were re-interviewed in 1992. All five basic traits--extroversion, agreeableness, conscientiousness, neuroticism, and openness to experience--had statistically significant positive or negative earnings effects, and together they appear to have had effects comparable to those commonly found for cognitive ability. Among men, substantial earnings advantages were associated with antagonism (the obverse of agreeableness), emotional stability (the obverse of neuroticism), and openness to experience; among women, with conscientiousness and openness to experience. Of the five traits, the evidence indicates that agreeableness had the greatest influence on gender differences in earnings: men were considerably more antagonistic (non-agreeable) than women, on average, and men alone were rewarded for that trait.

Keywords
personality and earnings, male and female eranings

This article is available in ILRReview: http://digitalcommons.ilr.cornell.edu/ilrreview/vol60/iss1/1

ESTIMATING THE EFFECT OF PERSONALITY ON MALE AND FEMALE EARNINGS

GERRIT MUELLER and ERIK PLUG*

The authors adopt the Five-Factor Model of personality structure to explore how personality affected the earnings of a large group of men and women who graduated from Wisconsin high schools in 1957 and were re-interviewed in 1992. All five basic traits—extroversion, agreeableness, conscientiousness, neuroticism, and openness to experience—had statistically significant positive or negative earnings effects, and together they appear to have had effects comparable to those commonly found for cognitive ability. Among men, substantial earnings advantages were associated with antagonism (the obverse of agreeableness), emotional stability (the obverse of neuroticism), and openness to experience; among women, with conscientiousness and openness to experience. Of the five traits, the evidence indicates that agreeableness had the greatest influence on gender differences in earnings: men were considerably more antagonistic (non-agreeable) than women, on average, and men alone were rewarded for that trait.

It is clear that individuals' cognitive abilities play a vital role in generating labor market success. Almost all empirical studies that focus on cognition and earnings find that returns to cognitive ability, measured by standardized test scores, are positive and statistically significant. But we know little about the role of noncognitive traits. Empirical studies of the earnings effects of noncognitive traits not only are relatively scarce, but also vary widely in their specific focus—from Machiavellianism (Turner and Martinez 1977), to self-esteem (Goldsmith, Veum, and Darity 1997), to aggression-withdrawal (Osborne 2003), for example. The diversity of the traits studied, of their measures, and of the corresponding returns makes it difficult to identify a consistent pattern or to make any generalizations about the role of noncognitive traits in the labor market.

*Gerrit Mueller is Post-Doctoral Researcher in the Department of Labor Economics, Human Resources, and Social Policy, Centre for European Economic Research (ZEW), and Research Affiliate, Institute for the Study of Labor (IZA). Erik Plug is Senior Researcher at the Research Center for Education and Labor, Amsterdam School of Economics, University of Amsterdam, and Research Fellow at the Tinbergen Institute and IZA. For helpful comments, the authors thank Catherine Weinberger, Arthur Goldberger, and numerous seminar and conference participants. Gerrit Mueller gratefully acknowledges the hospitality of the University of Chicago Graduate School of Business, where part of this research was carried out. Erik Plug is grateful to the Dutch National Science Foundation for its support under VIDI Grant 452.03.309. Support for collection and dissemination of data from the Wisconsin Longitudinal Study was provided by the National Institute on Aging (AG-9775), the National Science Foundation (SBR-9320660), the Spencer Foundation, and the Center for Demography and Ecology and Vilas Estate Trust at the University of Wisconsin–Madison.

Data and documentation from the Wisconsin Longitudinal Study are available at http://dpls.dacc.wisc.edu/WLS/wlsarch.htm. Copies of the computer programs used to generate the results presented in this paper are available upon request from the authors. E-mail: gerrit.mueller@zew.de; e.j.s.plug@uva.nl.

Industrial and Labor Relations Review, Vol. 60, No. 1 (October 2006). © by Cornell University.
0019-7939/00/6001 $01.00

This paper investigates the link between personality and earnings by incorporating traits from the Five-Factor Model (FFM) of personality structure (Digman 1990; Goldberg 1990) into models of wage determination using data from the Wisconsin Longitudinal Study (WLS). The five personality traits in the FFM are extroversion, agreeableness, conscientiousness, neuroticism, and openness to experience. In addition to estimating how yet another five different personality traits affect earnings, this paper offers three advantages over previous studies.

The first advantage concerns the comprehensiveness of the model we use to estimate the role of personality. Psychologists argue that virtually any personality construct can be mapped onto the FFM. Therefore, the five-factor taxonomy may also be of interest to an audience of economists. It may serve as a framework for integrating the existing body of evidence, as well as for structuring future research efforts. The second advantage is our application of the FFM to the gender wage gap. We explicitly allow for gender differences, both in personality traits and in the corresponding premia/penalties. This enables us to examine to what extent gender differences in earnings are due to differences in masculine and feminine personality traits, as opposed to differences in estimated returns to such traits. The third advantage is the direct comparison that our data allow us to draw between returns to the five personality traits and returns to cognitive ability.

Although we believe our approach has some distinct strengths, it is not without an important limitation—one that it shares with most other studies in the field and that is related to availability of reliable and credibly exogenous measures of personality. Reviewing the existing literature on the importance of noncognitive traits, Carneiro and Heckman (2004) expressed our concern. They noted that most personality determinants of earnings studied so far are self-reported ex-post assessments and are likely to be both causes and consequences of labor market outcomes. However, they also emphasized that such studies are valuable for the light they shed on the importance of personality traits. Given that research on personality traits is still in its infancy, there is ample room for exploratory studies of the kind presented here.

The Five-Factor Model of Personality Structure

According to the Five-Factor Model (FFM), five independent categories are sufficient to describe individual personality differences at the broadest level of abstraction (Costa and McCrae 1992; Goldberg 1990). As noted above, the dimensions of the FFM are labeled extroversion, agreeableness, conscientiousness, neuroticism, and openness to experience. This categorization does not imply that all personality attributes can be fully reduced to five traits. Rather, the "big five" should be viewed as broad factors underlying a number of related personality facets and sets of even more specific attributes. To provide a better idea of what they are, in Table 1 we list a number of characteristics related to each of the five personality dimensions.

The five-factor categorization of personality is pervasive in the current personality and social psychology literature. The FFM was first suggested by studies that tried to organize trait adjectives commonly used to describe people, available from dictionaries of a natural language, into a taxonomic structure (Allport and Odbert 1936; Norman 1963; Tupes and Christal 1961). The factorial structure has since been replicated in a large number of studies, cross-validated using a variety of questionnaire scales, and found to generalize across languages and cultures (see, for example, McCrae and Costa 1997).[1]

In this paper, we adopt the standard economic viewpoint of personality as a bundle of productive attributes valued in the labor market. Earnings follow, as usual, from the kind and amount of traits possessed, and the return that each trait receives in the market. We thereby implicitly assume that personal-

[1]For comprehensive reviews of the historical roots of the FFM, as well as the more recent developments, we refer the reader to Digman (1989, 1990), John and Srivastava (1999), McCrae and Costa (1999), and McCrae and John (1992).

ity affects behavior. This view closely corresponds to that of trait theorists who believe that personality traits constitute basic determining tendencies (for example, McCrae and Costa 1999). Determining tendencies are psychological dispositions that evoke "recurrent patterns of acting and reacting that simultaneously characterize individuals and differentiate them from others" (McCrae and Costa 1999:140). This interpretation does not imply that traits predispose an individual to behave in exactly the same way, irrespective of the situation. It merely holds that traits make certain behaviors more likely—and therefore more frequently observed across a multitude of situations.[2]

Using the FFM as a comprehensive framework to organize traits, multiple studies by organizational and industrial psychologists have examined how the big five personality dimensions relate to labor market outcomes, including job performance (Barrick and Mount 1991; Tett, Jackson, and Rothstein 1991), job satisfaction (Judge, Heller, and Mount 2002), firm performance (Welbourne, Cavanaugh, and Judge 1998), and, most closely related to our focus, executive career success (Boudreau, Boswell, and Judge 2001) and occupational attainment across the life span (Judge, Higgins, Thoresen, and Barrick 1999).

We stress that the FFM is not the last word on personality, and researchers should not, "seduced by convenience and seeming consensus, act as if they can obtain a complete portrait of personality by grabbing five quick ratings" (Funder 2001:201). Nevertheless, it is certainly fair to say that the FFM is the most comprehensive categorization of personality traits available to date. Moreover, for the time being, economists may find the FFM useful for the same reason psychologists did: it can help integrate findings on a variety of traits studied in isolation.

[2]The trait perspective, like every theory, is not without its critics. For brevity, we refer to Funder (2001), who reviewed all of the major research paradigms in personality psychology, such as the behaviorist and social-cognitive paradigms, as well as psychoanalytic, biological, and evolutionary perspectives.

Table 1. The Big Five Personality Traits.

Dimension	*Facet (and correlated trait adjective)*
Extroversion vs. Introversion	Gregariousness (sociable) Assertiveness (forceful) Activity (energetic) Excitement-seeking (adventurous) Positive emotions (enthusiastic) Warmth (outgoing)
Agreeableness vs. Antagonism	Trust (forgiving) Straightforwardness (not demanding) Altruism (warm) Compliance (not stubborn) Modesty (not showing off) Tender-mindedness (sympathetic)
Conscientiousness vs. Lack of Direction	Competence (efficient) Order (organized) Dutifulness (not careless) Achievement striving (thorough) Self-discipline (not lazy) Deliberation (not impulsive)
Neuroticism vs. Emotional Stability	Anxiety (tense) Angry hostility (irritable) Depression (not contented) Self-consciousness (shy) Impulsiveness (moody) Vulnerability (not self-confident)
Openness vs. Closedness to Experience	Ideas (curious) Fantasy (imaginative) Aesthetics (artistic) Actions (wide interest) Feelings (excitable) Values (unconventional)

Note: This table is adapted from John and Srivastava (1999) and shows Costa and McCrae's (1992) NEO-PI-R Facets.

Personality and Earnings Differentials

In this section we briefly discuss how personality may affect earnings. We follow a very simple framework outlined in most economic studies on earnings differentials and distinguish three alternative sources to explain differences in pay: differences in skills, differences in preferences, and a discriminating labor market. Within this framework, we will then draw on a much

bigger literature in psychology, especially the empirical part of it, to understand which particular personality traits matter for earnings, and in what way.

Differences in skills. Human capital theory features prominently in the analysis of wage differentials (Becker 1975; Mincer 1958). In this framework, systematic variation in earnings arises from differences in productive skills. Productive skills are individual human capital attributes providing a direct input into the production process, and may be either innate abilities or skills developed through investments in education, training, and work experience. Individuals sell their bundle of skills to firms at an equilibrium market price per unit of skill. Therefore, an individual's overall compensation depends on the kind and amount of skills possessed, and the return that each subcomponent of the skill vector earns in a given occupation. In this vein, one may think of personality as part of an individual's set of productive traits, valued in the market, with some components that are exogenous (innate) and others that are endogenous (developed over time). Of course, this does not mean that personality traits are equally productive across all occupations. If some traits are valued in certain occupations but not in others, we expect to find occupational sorting based in part on personality, assuming that workers choose those occupations that offer the highest rewards for their traits.

There have been a number of studies in occupational psychology in which personality traits are linked to job performance (Barrick and Mount 1991; Tett, Jackson, and Rothstein 1991). Since job performance is closely related to the economist's notion of productive output, we may associate these personality effects directly with higher earnings. Tett et al. (1991) found that neuroticism and job performance were negatively related across all occupations. Barrick and Mount (1991) reported a robust positive correlation between conscientiousness and job performance across all occupation groups. Extroversion was positively linked to job performance in management and sales occupations, that is, jobs involving a strong interpersonal performance component. Thus, the traits that increase performance, and thereby wages, depend on the requirements of the job. This evidence is well in line with sorting theories suggesting that some of the five personality dimensions may predict "extrinsic career success" (as measured, for example, by salary) if personality traits fit the psychological requirements of the job (Bretz and Judge 1994; Holland 1997; Tharenou 1997).

Any group differences in personality traits between men and women will translate into gender differences in earnings either directly, through productivity differences, or indirectly, through occupational segregation (for example, Polacheck 1981). In this regard, we expect the agreeableness and neuroticism dimensions to be of importance. In a recent literature review, Bouchard and Loehlin (2001) concluded that agreeableness and neuroticism are the two traits most consistently showing the largest gender differences.

There are alternative theories of earnings determination according to which certain traits/skills receive a return in a competitive market even though they are not productive skills in the sense defined above. One example, which we shall explore in this paper, is the wage bargaining model. In *Women Don't Ask: Negotiation and the Gender Divide,* Babcock and Laschever (2003) argued that personality differences between men and women may lead to differences in pay: women shy away from negotiations, and if they do start negotiating, they ask for less in their opening offer than men do, and tend to concede too quickly.[3]

Differing preferences. In addition to differences in skills, individuals may also have different preferences or tastes that are work-related. If these differences are related to personality, it is possible that personality affects earnings indirectly through occupa-

[3]These hypotheses are related to recent experimental work on behavioral differences between men and women finding that women try to avoid competitive environments, and that they perform worse than men within such environments (Gneezy, Niederle, and Rustichini 2003; Niederle and Vesterlund 2005).

tional choice processes. In particular, Tokar, Fischer, and Subich (1998) reported results that indicate some overlap between FFM personality traits and vocational preferences. Statistically significant positive associations were generally found between openness and artistic and investigative interests and between extroversion and enterprising and social interests. These findings generalized across genders. Judge and Cable (1997), who studied a sample of college graduates applying for jobs with various employers, found that agreeableness was positively related to preferences for supportive, team-oriented organizational cultures, and negatively related to aggressive, decisive, and outcome-oriented cultures. Conscientiousness related positively to preferences for detail and outcome orientation, and negatively to preferences for innovative cultures.

Labor market discrimination. In light of the large part of the gender wage gap left unexplained by productive traits and vocational preferences, it has been argued that differences in occupational structure and pay may also be a result of labor market discrimination (for example, England 1982). In fact, one may even conceive of discrimination against women that starts before they enter the labor market. This is because subject choice in schools plays a major role in determining subsequent occupational choices and thereby earnings. Women may be discouraged from entering gender-nontraditional fields of study such as engineering, physics, and mathematics. Such gender stereotyping may later confine them to traditional service-oriented female-type occupations with, generally, lower wages. Although discrimination certainly plays an important role as a determinant of the gender pay gap, it is difficult to separate empirically the differences in pay that are due to discrimination from differences in unobserved preferences and productive traits (for example, Bertrand and Hallock 2001).

With this in mind, the main focus of our empirical work is to determine whether standard earnings equations yield evidence of a pay difference based on personality, and to what extent these differences in pay relate to labor market sorting. In addition, we explicitly focus on the gender wage gap and ask whether such earnings differentials occur because of differences in males' and females' prototypical personality traits or because men and women face different returns for the same traits.

Method

Data and Sample

Our analysis employs the Wisconsin Longitudinal Study (WLS) of 10,317 randomly sampled Wisconsin students in the 1957 graduating class. After the initial wave of data collection, primary respondents were re-interviewed in 1975 and 1992. Together with the 1964 interviews with their parents, these waves provide information on, among other things, educational attainment, mental ability, socio-economic background, family formation, and labor market histories. The original sample is broadly representative of white men and women who had completed at least twelve years of schooling.[4]

Unlike other large longitudinal studies of school-based samples, the WLS contains personality measures together with information on respondents' labor market careers. This allows us to work with a much larger sample than comparable studies do in the psychological literature. We use data on personality traits from the 1992 mail questionnaire sent to 8,493 members of the original survey. This questionnaire collected information on respondents' personality traits based on the Big Five Inventory (BFI) developed by John, Donahue, and Kentle (1991). The BFI was specifically designed to facilitate the collection of personality data in surveys using a short test instrument that allows efficient assessment of the five dimensions when there is no need for a more differentiated measurement of individual facets (John and Srivastava 1999).

Of the initial 10,317 randomly sampled high school graduates, 8,493 received the

[4]For more detailed information on the WLS, see Sewell, Hauser, Springer, Hauser (2001) and the references therein.

1992 mail questionnaire; 6,875 responded; and 6,692 gave at least two complete answers to the separate items that correspond to each personality trait. Thus, nonresponse is a potential threat to the validity of our analysis, although, compared to other studies covering a similarly long time span, the response rate is fairly high. The population under study was then restricted to men and women who were employed in 1992, which reduced our sample to 6,062 observations. We further excluded all workers who were self-employed, worked less than 20 hours per week, or earned less than one dollar per hour, as well as all those for whom data on the various control variables were unavailable. In the end, our sample contained 5,025 observations. Descriptive statistics for both the male (N = 2,424) and female (N = 2,601) subsamples are provided in Table 2.

Measurement

The personality test instrument used in the WLS assesses the various dimensions by means of self-ratings on 29 questionnaire items. It is an abbreviated version of the original 44-item BFI (John et al. 1991). Each dimension is assessed by 6 items, except for neuroticism–emotional stability, which is assessed by 5 items. Items are statements such as "I see myself as someone who is talkative" or "I see myself as someone who is easily distracted." Individuals are asked to rate to what extent these statements apply to them on a 6-point scale ranging from "agree strongly" (1) to "disagree strongly" (6). The single item responses are then coded into average scores.

Any research based on measurement must confront the question of the reliability of its measures. We quantify the size of the measurement error by calculating reliability coefficients for the BFI scales, often referred to as Cronbach's alpha reliabilities (Cronbach 1951). Derivations are provided in the Appendix. We found some notable differences: extroversion, .76; agreeableness, .71; conscientiousness, .66; neuroticism, .77; and openness, .60. The reliabilities of the abbreviated scales averaged .70, which suggests that a considerable fraction of the variability in the reported traits is due to measurement error.

It is possible to compare these numbers with previous estimates of reliability ratios. John and Srivastava (1999) reported that reliabilities of the original 44-item BFI scales typically lie between .75 and .90 and are on average above .80. These estimates indicate that the internal consistency of the BFI scales is high on average, and about ten percentage points higher than the numbers we find for our own data. This does not necessarily mean, however, that the original BFI scales are better. Reliability ratios increase with the number of items. The abbreviated BFI scales we use include only five to six items. If we had eight to ten items, as the original scales do, the estimated reliability ratios would range from .69 to .83, average out at .78, and thus be very similar to the ones previously estimated.[5]

To measure general intelligence, we use test scores on the *Henmon-Nelson Test of Mental Ability* that respondents took in 1957 while attending high school. Unfortunately, we do not have access to individual test items, so we are unable to estimate reliability coefficients ourselves. In contrast to the BFI, however, the Henmon-Nelson test was implemented as originally designed. We can therefore safely rely on estimates available in the literature. The psychometric properties of the test are

[5]The reliability ratio R_0 of any given personality measure is defined by

$$R_0 = \frac{k_0\rho}{1 + (k_0 - 1)\rho},$$

where k_0 and ρ represent the number of items and the average inter-item correlation, respectively. Let R_1 be the reliability ratio of the same personality measure but measured with Δk additional items. With ρ fixed, we can express R_1 in terms of k_0, R_0, and Δk as follows:

$$R_1 = R_0\left(\frac{k_0 + \Delta k}{k_0 + R_0\Delta k}\right).$$

It is easy to see that reliability ratios increase with the number of items. For example, if a reliability ratio of .71 were obtained with 9 instead of 6 items, like our original agreeableness measure, the ratio would rise to .79.

[6]Zax and Rees (2002:603) reported that, based on publicly unavailable WLS data, Robert Hauser estimated the reliability ratio to be between .92 and .95. In our later calculations we impose the average of .935.

Table 2. Summary Statistics.

	Males (N = 2,424)		*Females (N = 2,601)*	
Independent Variable	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*
Labor Market Outcomes:				
Log Hourly Wages	2.886	*0.569*	2.299	*0.542*
Hourly Wages	21.891	*21.642*	11.827	*11.521*
Personality Traits:				
Extroversion	3.751	*0.878*	3.857	*0.898*
Agreeableness	4.597	*0.737*	4.887	*0.701*
Conscientiousness	4.875	*0.674*	4.904	*0.670*
Neuroticism	3.081	*0.956*	3.277	*0.981*
Openness	3.626	*0.770*	3.675	*0.807*
Individual Characteristics, Human Capital, and Region:				
Henman-Nelson IQ-Scores	102.225	*14.870*	102.666	*14.332*
Married	0.861		0.774	
No. of Children	2.483	*1.490*	2.660	*1.609*
Years of Education	14.076	*2.507*	13.474	*2.089*
Experience	17.773	*2.293*	15.372	*4.358*
Tenure	17.723	*10.972*	10.912	*8.602*
State of Residence Wisconsin	0.679		0.688	
Occupations:				
Professional and Technical	0.230		0.256	
Executive and Managerial	0.180		0.085	
Sales and Trade	0.098		0.092	
Clerical	0.062		0.368	
Production and Crafts	0.176		0.015	
Operatives	0.164		0.060	
Service	0.056		0.106	
Laborers	0.033		0.017	
Other	0.001		0.001	
Industries:				
Agriculture and Mining	0.014		0.006	
Construction	0.064		0.008	
Manufacturing	0.371		0.139	
Transportation	0.097		0.042	
Wholesale and Retail Trade	0.106		0.166	
Finance	0.048		0.088	
Services	0.221		0.504	
Administration	0.076		0.046	
Other	0.003		0.001	
Public Sector	0.242		0.277	
Part-Time	0.019		0.226	

well established, with reliability ratios ranging from .87 to .94 (Buros 1959).[6] In addition, potential measurement error due to time and cohort effects is ruled out by the very nature of our data.

Estimation

We first estimate a standard log-linear earnings equation separately for men and women in the form

$$(1) \qquad Y_{im} = X'_{im} b_m + \varepsilon_{im},\ Y_{if} = X'_{if} b_f + \varepsilon_{if},$$

where the subscripts *i, m,* and *f* indicate, respectively, individual, male, and female groups, *Y* denotes the logarithm of hourly earnings, *X* is a vector of covariates includ-

ing the five personality measures and control variables assumed to affect earnings, and ε is the remaining error. The parameter vector b contains the estimates of how the labor market would value different characteristics.

While this equation estimates whether personality traits matter for generating earnings, and whether they affect earnings differently for men and women, it does not tell us how large a role these differences play in explaining the gender gap in earnings. To address that question, we decompose the gender gap into two components: one that can be attributed to differences in observable personality traits between men and women, and a second that can be attributed to differences in trait premia/penalties between men and women. To calculate the latter component, we must choose which set of coefficients to use as the standard of comparison (male or female). We follow Neumark (1988) and Oaxaca and Ransom (1994) and use the common coefficients estimated from a pooled regression of men and women ($\hat{b}_p$). If we express the difference in earnings between men (m) and women (f) in terms of averages,

$$\bar{Y}_m - \bar{Y}_f = \bar{X}'_m \hat{b}_m - \bar{X}'_f \hat{b}_f, \tag{2}$$

where $\hat{b}_m$ and $\hat{b}_f$ are the estimates from (1), the earnings differential can be separated into two components,

$$\bar{Y}_m - \bar{Y}_f = [\bar{X}_m - \bar{X}_f]'\hat{b}_p + [\bar{X}'_m(\hat{b}_m - \hat{b}_p) - \bar{X}'_f(\hat{b}_f - \hat{b}_p)]. \tag{3}$$

The first term can be interpreted as the part of the earnings differential that is due to differences in observed characteristics, and the second as the part due to differences in estimated parameters. Decompositions like these have a long tradition in studies of wage differentials, beginning with the work of Oaxaca (1973), who interpreted the differences in returns between men and women with similar characteristics as a measure of labor market discrimination. We do not argue for that particular interpretation. As noted earlier, different parameters could just as easily stem from (unobserved) differences in preferences and productive skills.

Results

Personality and Male vs. Female Earnings

Table 3 reports OLS estimates of the relationship between our measures of the five personality traits and the log of hourly earnings, separately for men and women. For reasons of comparability, we have standardized each trait scale on the full estimation sample to have zero mean and unit variance. The same transformation is applied to IQ scores. Panels A and B of Table 3 show results for samples of employed men and women, respectively. Each panel contains four OLS estimates, with varying sets of covariates.

We begin by discussing the effects of personality traits on men's earnings. In column (i) we estimate a baseline specification in which the five personality traits are the only right-hand-side variables. We find that antagonistic, emotionally stable, and open men enjoyed statistically significant and substantial earnings advantages. Of the five personality traits, openness to experience seems to have been the most rewarding, whereas extroversion and conscientiousness generated no returns at all. In column (ii) we add the childhood IQ test scores to the regression to control for the respondents' cognitive ability. With this IQ measure added, the returns to being antagonistic, emotionally stable, and open to experience fall, but remain positive and statistically significant.[7] In column (iii) we add several other covariates, including

[7]Based on psychometric and experimental studies, psychologists argue that there is no meaningful relation between personality and intelligence. However, there is evidence that actual performance on IQ tests is related to some dimensions of personality. It has been found, for example, that introverts show more vigilance and less fatigue during extended tests. Also, feelings of anxiety (a facet of neuroticism) are known to affect test performance if the test subjects the individual to considerable stress (for example, time pressure). Our proxy variable for intelligence might be picking up this performance effect to some extent. For an exhaustive treatment of the relation between personality and intelligence, see Sternberg and Ruzgis (1994).

Table 3. The Effects of Personality on Male-Female Earnings.

Traits	*(i)*		*(ii)*		*(iii)*		*(iv)*	
A. Males, Log Hourly Earnings (N = 2,424)								
Personality Traits:								
Extroversion	–0.002	*0.012*	0.019	*0.012*	0.014	*0.011*	0.009	*0.010*
Agreeableness	–0.064	*0.012****	–0.047	*0.012****	–0.036	*0.011****	–0.037	*0.010****
Conscientiousness	–0.006	*0.012*	0.009	*0.012*	0.003	*0.011*	–0.002	*0.010*
Neuroticism	–0.050	*0.013****	–0.032	*0.012***	–0.022	*0.011***	–0.020	*0.011**
Openness	0.104	*0.012****	0.058	*0.012****	0.033	*0.011****	0.024	*0.011***
IQ Scores	—		0.179	*0.011****	0.098	*0.011****	0.065	*0.011****
Adjusted R^2	0.05		0.14		0.29		0.45	
F-Test Personality Traits	24.39		11.35		5.37		4.39	
B. Females, Log Hourly Earnings (N = 2,601)								
Personality Traits:								
Extroversion	–0.034	*0.011****	–0.022	*0.011***	–0.004	*0.010*	0.005	*0.009*
Agreeableness	–0.031	*0.012****	–0.023	*0.011***	–0.005	*0.010*	–0.008	*0.009*
Conscientiousness	0.030	*0.011****	0.028	*0.011***	0.025	*0.010****	0.023	*0.009****
Neuroticism	–0.035	*0.012****	–0.017	*0.011*	–0.018	*0.010**	–0.006	*0.009*
Openness	0.122	*0.011****	0.092	*0.011****	0.043	*0.010****	0.027	*0.010****
IQ Scores	—		0.127	*0.011****	0.066	*0.010****	0.051	*0.010****
Adjusted R^2	0.06		0.11		0.31		0.40	
F-Test Personality Traits	36.14		18.76		7.88		4.52	
Controls:								
Individual, Human-Capital, Region	—		—		×		×	
Occupation, Industry, Job Characteristics	—		—		—		×	

Notes: Standard errors in italics. F-tests indicate whether estimated coefficients for the big five personality traits are jointly significant.

*Statistically significant at the .10 level; **at the .05 level; ***at the .01 level.

years of schooling, work experience, tenure, region, and other individual and family characteristics. Including these variables reduces the estimated coefficients for non-agreeableness (antagonism), emotional stability, and openness, yet the results remain qualitatively similar.

It is not clear whether respondents' occupation and industry affiliations are appropriate variables to include in our wage regressions. If we believe that workers are selected into certain jobs on the basis of specific personality profiles, we would probably not want to control for occupations when estimating the effect of personality on earnings. When job-holding is controlled for, we expect it to partially mediate the effects of personality on earnings. In column (iv), which presents the results of an estimation that adds 8 one-digit industry and occupation dummies to the previous specification, returns to non-agreeableness and emotional stability remain virtually identical. Returns to openness fall somewhat, but remain statistically significant.[8] It appears that antagonistic

[8]For information concerning jobs held by individuals, we rely on the standard classification system of the U.S. Bureau of the Census, using 1990 occupation and industry definitions at the one-digit level. Of course, it is possible that personality results would change if we used more fine-tuned information on industry or occupation characteristics. As a simple test, we replaced the one-digit–level dummies with sets of occupation and industry indicators defined at the two-digit and three-digit level. We find that our parameter estimates are robust with respect to the inclusion of more detailed indicators of respondents' job-holding.

and open men always earned more. Across all specifications, the agreeableness-antagonism dimension has the most persistent effect on earnings. A one–standard deviation increase in antagonism raises the hourly earnings for male workers on average by 4% to 6%.

For measured cognitive ability, we find strong positive effects that are reduced substantially when we add human capital characteristics in the third column. The magnitude of the change in earnings induced by a one–standard deviation change in IQ scores is large—7% to 17%—conditional on the particular set of covariates entered. Thus, although the premia/penalties of the traits viewed in isolation range only from 0 to 6%, a favorable personality profile potentially has as strong an earnings effect as cognitive ability if personality is viewed as a bundle of traits. Nonetheless, personality does not predict as well as our cognitive ability measure. In isolation, the five personality measures explain about 5% of the variance in earnings. The addition of IQ test scores improves the R^2 by almost 10 percentage points.

And what about women? In column (i) of Table 3, which presents results of estimating the earnings specification without controls, we find that all five personality traits have statistically significant associations with earnings. Three personality estimates are very similar to those found for men: antagonistic women, emotionally stable women, and open women earned higher wages. Two personality estimates are different. Unlike men, women appear to have been penalized for being extroverted, while they received a premium for being conscientious. If we include IQ test scores in column (ii), we find that all partial personality effects decrease only marginally. The results, however, appear to be sensitive to adding more controls. In columns (iii) and (iv), only returns to openness and conscientiousness are consistently statistically significant and positive for women.[9] A one standard deviation increase in either of the two traits is associated with a 2–3% increase in hourly earnings. The combined premia of openness and conscientiousness are comparable in magnitude to the earnings effect of measured cognitive ability. With the full set of controls entered, a one–standard deviation increase in IQ raises hourly earnings by about 5%. A comparison of columns (i) and (ii) further shows that personality and cognitive ability are roughly equally important in explaining the variance in earnings. The five personality measures together account for approximately 6% of the earnings variation; adding IQ test scores to the baseline specification in column (ii) improves the R^2 by 5 percentage points.

Overall, these results suggest that personality predicts earnings for both men and women. A favorable personality profile—a distinct bundle of traits rewarded in the market—appears to have an impact on earnings that is comparable to that of cognitive ability. Of course, we do not claim to have identified causal effects of personality traits on earnings. We merely show that personality adds explanatory power to our model. All our F-tests indicate that the big five traits are jointly statistically significant. In terms of the magnitude of additional variance explained, the contribution is similar to that of cognitive ability—statistically significant, but modest.[10]

There are, to our knowledge, two studies in the psychological literature that also have investigated the relationship between the big five traits and earnings (Boudreau, Boswell, and Judge 2001; Judge, Higgins, Thoresen, and Barrick 1999). Both studies employed American data and are therefore valuable for comparative purposes.[11] Bou-

[9]As for men, results are insensitive to replacing the one-digit occupation and industry dummies with indicators defined at the more detailed two- or three-digit level.

[10]The magnitudes are, for example, much smaller than those conventionally found for education. According to the specification in column (iii), the traditional "return to education," as measured by the earnings effect of an additional year of schooling, is estimated to be .064 for men and .066 for women. When we replace the years of schooling variable with its standardized equivalent, we find point estimates of .149 [13.73] and .153 [13.18] for men and women, respectively (*t*-ratios in brackets).

[11]Nyhus and Pons (2005) analyzed the effect of big five personality traits on earnings using a sample of about

dreau, Boswell, and Judge (2001) studied the effects of personality on intrinsic and extrinsic career success based on samples of American and European executives. For the American sample, consisting primarily of white men in their late forties, they found that agreeableness and neuroticism were negatively related to remuneration, with extroversion and conscientiousness having little or no impact and openness to experience exhibiting a positive association. The highly selective nature of the sample places limits on possible generalizations, most importantly with respect to the effects one should expect for women. Judge et al. (1999) also found that agreeableness and neuroticism had a negative effect on earnings. Extroversion and conscientiousness were positively associated with earnings, but the positive effect of the openness dimension disappeared when the full set of conditioning variables was entered. Some caution is again in order when generalizing these findings to broader populations, as they are based on a sample only slightly in excess of 100 observations.[12]

Limitations

While our results are comparable to those obtained by previous studies, we should treat them with care. The parameter estimates presented in Table 3 may be subject to a number of sources of bias: measurement error in the BFI scales, selective non-response, misspecification, and simultaneity between wages and personality traits.

Measurement error. Our first concern relates to the possible attenuating effects of measurement error. If personality effects seem only modestly important, it is quite possible that our personality traits are measured with error. After all, random error will bias any estimated effect toward zero. One way to correct for such error is to adjust the parameter estimates and standard errors by imposing reliability ratios in estimation (see Appendix B). Panel A of Table 4 presents parameter estimates that are adjusted for the effects of measurement error.[13] The estimated effects remain qualitatively very similar, except that they are almost all larger than the corresponding point estimates in Table 3. The increase is substantial and often statistically significant. Assuming that there is no serial correlation among the measurement errors across the five personality scales, our results suggest that unreliability in trait measurement indeed leads to a considerable underestimation of the corresponding premia and penalties.[14]

Selective non-response. A related concern derives from the fact that respondents are kept in the sample if they provided at least two complete answers to the question sets that correspond to each personality trait. It is possible that selective non-response introduces inconsistencies when we estimate our regression models. In order to see how sensitive our results are, we calculate reliability ratios and run OLS regressions on a sample of workers who responded to all items. As

800 male and female workers in the Netherlands. In a specification comparable to our model in column (iii), they mainly found that emotional stability was positively associated with wages of both women and men. They did not analyze a gender wage decomposition.

[12]A number of recent studies have investigated whether other non-cognitive traits account for differences in labor market success (Bowles, Gintis, and Osborne 2001; Duncan and Dunifon 1998; Dunifon, Duncan, and Brooks-Gunn 2001; Goldsmith, Veum, and Darity 1997; Osborne 2003), building on earlier work by Andrisanni (1978), Filer (1981, 1986), Jencks (1979), and Turner and Martinez (1977). In Mueller and Plug (2004) we provide more details on how some of these particular traits can be mapped onto the FFM.

[13]We only allow for unreliability in the measurement of the five personality traits and the *Henmon-Nelson* IQ scores. Note further that in Table 3 it was useful to see how coefficients changed as additional covariates were added. This is not as important when we test for the effects of measurement error. We therefore show only two specifications that correspond to columns (ii) and (iii) of Table 3.

[14]We have skirted the more subtle issue of subjectivity in self-reported data. Bertrand and Mullainathan (2001) discussed how cognitive factors and the social nature of the survey procedure may affect the way people answer questions, and how subjectivity may be treated in a measurement-error framework. We cannot explore this issue empirically with the data at hand, but refer the reader to Costa and McCrae (1988), who presented evidence that the convergence between self ratings and peer or expert ratings is, on average, between .80 and .90.

Table 4. Sensitivity Analyses.

	Men				*Women*			
Personality Traits:	*(i)*		*(ii)*		*(i)*		*(ii)*	
A. Effects of Personality on Earnings Corrected for Measurement Error								
Extroversion	0.009	*0.019*	0.009	*0.018*	–0.067	*0.017****	–0.028	*0.016*
Agreeableness	–0.085	*0.021****	–0.067	*0.019****	–0.035	*0.020**	–0.014	*0.018*
Conscientiousness	0.023	*0.023*	0.011	*0.021*	0.055	*0.021****	0.045	*0.018***
Neuroticism	–0.042	*0.020***	–0.033	*0.018**	0.005	*0.018*	–0.007	*0.017*
Openness	0.103	*0.025****	0.063	*0.025***	0.189	*0.025****	0.100	*0.026****
IQ Scores	0.179	*0.013****	0.102	*0.013****	0.111	*0.013****	0.063	*0.011****
R^2	0.16		0.30		0.15		0.33	
F-Test Personality Traits	10.81		5.28		18.17		7.72	
B. Effects of Personality on Earnings Using the Full-Response Sample								
Extroversion	0.006	*0.020*	0.008	*0.019*	–0.054	*0.020****	–0.023	*0.018*
Agreeableness	–0.087	*0.022****	–0.067	*0.021****	–0.042	*0.022**	–0.021	*0.020*
Conscientiousness	0.021	*0.024*	0.011	*0.022*	0.057	*0.022***	0.051	*0.020***
Neuroticism	–0.043	*0.021***	–0.036	*0.019**	0.003	*0.020*	–0.009	*0.018*
Openness	0.108	*0.027****	0.065	*0.026***	0.177	*0.028****	0.090	*0.030****
IQ Scores	0.180	*0.04****	0.105	*0.013****	0.111	*0.04****	0.066	*0.012****
R^2	0.16		0.30		0.13		0.31	
F-Test Personality Traits	9.66		4.86		14.17		6.29	
N	2,149		2,149		2,225		2,225	
C. Testing for Nonlinear Effects of Personality on Earnings								
Extroversion								
Bottom 25%	–0.016	*0.026*	–0.012	*0.024*	0.015	*0.025*	–0.004	*0.023*
Top 25%	0.050	*0.028**	0.023	*0.026*	–0.026	*0.025*	–0.010	*0.022*
Agreeableness								
Bottom 25%	0.067	*0.024****	0.060	*0.026****	0.049	*0.025**	0.000	*0.022*
Top 25%	–0.050	*0.033*	–0.029	*0.030*	0.008	*0.025*	0.009	*0.022*
Conscientiousness								
Bottom 25%	–0.004	*0.027*	–0.000	*0.024*	–0.043	*0.025**	–0.049	*0.022***
Top 25%	0.023	*0.028*	0.021	*0.025*	0.014	*0.025*	0.003	*0.022*
Neuroticism								
Bottom 25%	0.058	*0.036***	0.036	*0.024*	0.007	*0.026*	0.006	*0.023*
Top 25%	–0.050	*0.026**	–0.038	*0.027*	–0.047	*0.025**	–0.045	*0.022***
Openness								
Bottom 25%	–0.070	*0.026****	–0.033	*0.023*	–0.112	*0.025****	–0.074	*0.022****
Top 25%	0.082	*0.028****	0.049	*0.027**	0.147	*0.026****	0.055	*0.024***
IQ Scores								
Bottom 25%	–0.217	*0.028****	–0.131	*0.024****	–0.158	*0.024****	–0.081	*0.022****
Top 25%	0.268	*0.028****	0.119	*0.026****	0.175	*0.025****	0.083	*0.023****
Adjusted R^2	0.12		0.29		0.10		0.31	
F-Test Personality Traits	5.93		2.72		25.07		7.83	
Controls	—		×		—		×	

Notes: Standard errors in italics. Reliability ratios imposed in the estimation in Panel A: extroversion, .76; agreeableness, .68; conscientiousness, .63; neuroticism, .77; openness to experience, .60; Henman-Nelson IQ scores, .94. In Panel B the sample is restricted to workers who responded to all personality items. Corresponding reliability ratios imposed in the estimation: extroversion, .77; agreeableness, .69; conscientiousness, .64; neuroticism. 77; openness to experience, .60. In Panel C the (omitted) reference categories are the 2nd and 3rd quartiles of the respective trait distribution; F-tests indicate whether estimated coefficients for the big five personality traits are jointly statistically significant. The set of controls includes all variables on individual, human-capital, and region characteristics as detailed in Table 2.

*Statistically significant at the .10 level; **at the .05 level; ***at the .01 level.

Panel B of Table 4 shows, this restriction induces a loss of 651 observations but does not affect the results substantially. This means that the trait coefficients, as well as the estimated reliability ratios, are not sensitive to item non-response.

Nonlinearities. The third issue relates to whether or not the relationship between the big five traits and hourly earnings is nonlinear. When it comes to personality, it is not *a priori* clear that more is necessarily better. If, for example, the labor market values people who are only moderately extroverted and punishes those who are too introverted or too extroverted, it is possible that the linear specification is pushing estimated average returns to zero. In Panel C of Table 4, we test for nonlinear personality effects by replacing the reported trait scores with sets of trait-level dummies. For each personality trait, we transform the average reported scores into quartiles and create three corresponding dummy variables for whether or not the personality scores are in the top or bottom 25% of the distribution, with the middle 50% being the omitted category. With personality traits measured in levels, we observe that not all of the individual dummy coefficients are significantly different from zero. However, for those personality traits that mattered in the linear specifications, we find that many individual dummy variables are statistically significant and show a consistent monotonic pattern. These results suggest that for the traits that mattered previously, a linear representation is a fairly accurate approximation of the overall relationship. For the traits that did not affect earnings in previous specifications, the fluctuations we observe are difficult to reconcile with any consistent pattern.

Reverse causation. Our fourth and biggest concern is the potential for endogeneity in the personality measures. Since in our data personality traits were assessed at the same time as hourly earnings, we do not know whether personality is the cause or the consequence of earnings. To the extent that trait measures are endogenous, our parameter estimates will be upwardly biased because they capture both cause and effect. In what follows, we will argue that the extent of the bias may not be as severe as it appears at first sight.

It is well understood that personality traits are both inherited and formed. Bouchard and Loehlin (2001), for example, reviewing a large number of twin studies, found that 40–60% of the variation in personality is attributable to genetic differences between individuals. The inherited part, which is substantial, can be treated as predetermined with respect to earnings. The concern about endogeneity bias derives, of course, from the fact that part of one's personality is developed over time and shaped by labor market experiences.

The current state of knowledge indicates that the formation of personality occurs primarily during early childhood and adolescence, that personality is largely set by age 30, and that it remains fairly stable thereafter (see the reviews by Caspi and Roberts 1999; Costa and McCrae 1994, 1997; Digman 1989). The evidence of stability of mean levels in big five traits, "absolute continuity," is strong and consistent. Moreover, measures of personality traits are found to exhibit strong "differential continuity," meaning that individuals tend to preserve their relative position within the respective trait distribution as they age (see, for example, Costa and McCrae 1988). Overall, the big five personality traits are heritable and enduring individual predispositions, second in stability only to measures of cognitive ability (Conley 1984).

Our study is based on a single cohort of equal-age individuals in their early fifties. It therefore offers the clear advantage that sample members are homogeneous in terms of age and timing of personality measurement. We have remarked evidence that mean trait levels change only imperceptibly over time and that individuals generally maintain their own rank order within the group. This evidence implies that even if personality changes as people age, it is unlikely that the corresponding estimates are driven much by the simultaneity between wages and our personality regressors. We are aware that these arguments do not prove that endogeneity bias is absent. With the data at hand, it is impossible to remove this bias. It is appropriate, however, to inter-

Table 5. Male-Female Differences in Personality Traits and Coefficients.

	Characteristics $(X_m - X_f)$		*Coefficients* $(b_m - b_f)$					
Traits	*(i)*		*(ii)*		*(iii)*		*(iv)*	
Personality Traits:								
Extroversion	–0.120	*0.028****	0.075	*0.025****	0.037	*0.024*	0.012	*0.023*
Agreeableness	–0.396	*0.027****	–0.050	*0.029**	–0.052	*0.027***	–0.049	*0.025**
Conscientiousness	–0.043	*0.028*	–0.033	*0.030*	–0.034	*0.028*	–0.041	*0.027*
Neuroticism	–0.202	*0.028****	–0.047	*0.027**	–0.025	*0.025*	–0.036	*0.024*
Openness	–0.062	*0.028***	–0.086	*0.035***	–0.037	*0.036*	–0.016	*0.036*
IQ Scores	–0.030	*0.028*	0.068	*0.018****	0.039	*0.017***	0.016	*0.016*
Controls:								
Individual, Human-Capital, Region			—		×		×	
Industry, Job Characteristics			—		—		×	

Note: Standard errors in italics.
*Statistically significant at the .10 level; **at the .05 level; ***at the .01 level.

pret our estimates as upper bounds of true personality effects.

To summarize, we find that our estimated personality returns (a) increase substantially when adjusted for measurement error; (b) are not caused by selective non-response; (c) are not an artifact of the linear specification; but (d) remain upper bounds of actual personality effects due to the endogeneity of our personality regressors. The sensitivity tests we performed reinforce our earlier conclusion. The impact of personality on earnings is statistically significant and comparable to the impact of differences in cognitive ability, and though it is not large, it is not trivial either.

Decomposing the Gender Gap

In the tradition of most empirical work on wage differentials, we focus on two major mechanisms explaining differences in pay: differences in characteristics, and differences in the corresponding premia and penalties.

Gender differences in personality traits. In the first column of Table 5 we test for average gender differences in personality characteristics using standardized trait scales to make the gender differences more readily visible. We find that women scored significantly higher than men along the agreeableness, neuroticism, extroversion, and openness dimensions, with gender differences in the first two traits being the largest. Gender means for neuroticism and agreeableness lie 20–40% of a standard deviation apart. These findings are consistent with evidence from the psychological literature (for example, Bouchard and Loehlin 2001).

Gender differences in the earnings premia and penalties. The next three columns of Table 5 present gender differences in personality returns based on error-corrected estimates.[15] According to the baseline model in column (ii), women received higher returns than men for the traits introversion, agreeableness, conscientiousness, neuroticism, and openness. As we include more controls (columns iii and iv), agreeableness remains the only dimension that shows statistically significant differences in labor market valuation across genders, with the estimated penalties for men driving these differences. With regard to measured cognitive abilities, we find that differences in returns to IQ scores are not statistically significant. Since the magnitude and direction of the effects happen to be

[15]Since it is interesting to see what happens to the estimated differences when we take into account other sources of variation, we compute differences in parameter estimates based on the specifications shown in columns (ii)–(iv) of Table 3.

Table 6. Decomposition Results for Personality Traits.

Description	*(i)*		*(ii)*		*(iii)*	
Difference:						
(1) Log Hourly Earnings	0.587					
Differences Due To:						
(2) Characteristics	0.091	*15.6%*	0.307	*52.3%*	0.403	*68.7%*
(3) Coefficients	0.495	*84.4%*	0.280	*47.7%*	0.184	*31.3%*
Differences Due To:						
(4) Personality Characteristics	0.095	*16.2%*	0.060	*10.3%*	0.043	*7.3%*
(5) Personality Coefficients	–0.074	*12.7%*	–0.045	*7.7%*	–0.027	*4.5%*
Differences Due to Characteristics:						
(6) Extroversion	0.006	*1.0%*	0.002	*0.3%*	–0.000	*0.1%*
(7) Agreeableness	0.076	*12.9%*	0.048	*8.1%*	0.035	*5.9%*
(8) Conscientiousness	–0.003	*0.5%*	–0.002	*0.3%*	–0.001	*0.2%*
(9) Neuroticism	0.023	*4.0%*	0.016	*2.8%*	0.011	*1.9%*
(10) Openness	–0.007	*1.2%*	–0.003	*0.6%*	–0.001	*0.2%*
(11) IQ Scores	–0.004	*0.7%*	–0.002	*0.3%*	–0.002	*0.3%*
(12) Other Characteristics	0.248	*42.2%*	0.361	*61.7%*		
Differences Due to Coefficients:						
(13) Extroversion	–0.003	*0.4%*	–0.001	*0.1%*	0.000	*0.1%*
(14) Agreeableness	–0.052	*8.8%*	–0.031	*5.3%*	–0.017	*2.9%*
(15) Conscientiousness	0.001	*0.2%*	0.001	*0.1%*	0.000	*0.0%*
(16) Neuroticism	–0.019	*3.3%*	–0.012	*2.1%*	–0.008	*1.4%*
(17) Openness	–0.002	*0.3%*	–0.002	*0.3%*	–0.002	*0.4%*
(18) IQ Scores	–0.000	*0.1%*	0.001	*0.1%*	–0.000	*0.0%*
(19) Other Characteristics			0.139	*23.6%*	0.242	*41.3%*
(20) Intercept	0.570	*97.2%*	0.187	*31.9%*	–0.032	*5.4%*
Controls:						
Individual, Human-Capital, Region	—		×		×	
Occupation, Industry, Job Characteristics	—		—		×	

Note: Earnings effects as a proportion (percentage share) of the gross differential in italics.

very similar across genders, the difference is virtually zero.

It is further interesting to note that, except for extroversion, the penalties (returns) to men tend to be larger for those personality traits for which men have the lower (higher) means; the converse holds for women.[16] This can easily be verified by checking that the following interaction term $(\overline{X}_m - \overline{X}_f)'(\hat{b}_m - \hat{b}_f)$ is strictly positive. Apparently, it is not universally better—an absolute advantage—to be masculine, but individuals with masculine traits have a comparative advantage under a male wage structure; and, similarly, individuals with feminine traits have a comparative advantage under a female wage structure.

Decomposition results. In Table 6 we report earnings decompositions based on error-corrected parameter estimates from male and female earnings equations for the same three specifications as before. The overall differential, that is, the difference in logarithms of hourly wages between men and women, amounts to .58. The magnitude of the gap is large, but is not unusual for the particular generation of men and women under study.

In rows 2 and 3 we start decomposing the wage gap into the share that is attributable to differences in characteristics (included in the model) versus differences in coefficients. It is clear that the part of the wage

[16]The finding that women were slightly more extroverted than men appears odd, at first sight. However, extroversion has both dominance and sociability facets, and our abbreviated test instrument may be picking up the sociability component to a larger extent.

gap explained increases as more regressors are added. Our primary focus, of course, is on the decomposition results for personality traits in the fourth and fifth rows. Based on the first specification (column i), we find that 16% of the gender gap can be attributed to differences in mean personality traits, and about 13% to differences in labor market rewards/penalties. When additional variables are introduced in columns (ii) and (iii), these numbers fall to 10% and 8%, then to 7% and 5%. Note that the effects work in opposite directions, such that the overall differential is only moderately affected. Overall, only 3–4% of the gender gap is explained by differences in personality, including differences in traits and trait returns. This number is about the same in every column.

In the remaining rows we report in more detail how much of the total difference is attributable to each of the five personality traits separately. We find that the decomposition results for personality are primarily driven by one single dimension: agreeableness-antagonism. Rows 6 to 10 indicate that most of the share explained by differences in personality characteristics comes from mean differences in agreeableness. Rows 13 to 17 describe a similar pattern, showing that most of the gender differences in personality returns are due to the differences in returns to non-agreeableness.

Interpreting the decomposition results. Our findings thus suggest that, among the five personality traits, differences in agreeableness are the most important factor explaining differences in male-female earnings. Two channels are responsible for this result: men are much less agreeable than women; and men are those who receive a reward for being less agreeable. The first channel does not require an economic explanation, while the second one does. In what follows, we consider labor market sorting, discrimination, and bargaining as possible explanations for why the market rewards agreeableness differently for men and women.

The sorting argument is one of the standard arguments to explain particular differences in pay. With our focus on agreeableness, sorting implies that less agreeable workers (primarily men) select into occupations where being less agreeable is required and rewarded as a productive trait. In regression models that do not control for occupation choice, it is possible that the return estimates for the agreeableness trait are picking up sorting effects. A simple test for sorting would therefore be to add variables to our models that measure characteristics of occupations in which non-agreeableness is possibly productive. As it is difficult to identify such occupations *a priori,* we revert to the equivalent procedure of conditioning directly on occupation codes at the one-, two-, and three-digit level. Consistent with our earlier findings, we conclude that there is little evidence that sorting explains the large difference between men and women in the returns to agreeableness.

An interpretation that is partly consistent with ours comes from Badget and Fobre (2003), who argued as follows. In the presence of societal expectations about gender-appropriate traits and behavior, it is possible that the market rewards men and women who conform to traditional gender roles and punishes those who deviate. With this type of discrimination, we expect that agreeable men and non-agreeable women are punished for being perceived as too feminine and too masculine, respectively. In our empirical analysis, men's returns to agreeableness behaved as the model predicts: the market punished those who were too considerate and cooperative, by male standards. In the case of women, however, our evidence does not square with the model predictions. Women's returns to agreeableness were either negative or close to zero, but never positive.

An alternative argument that has received little attention in the empirical literature on gender wage differentials is related to wage bargaining. Most of the work on gender differences in bargaining has been done by psychologists, who find that women are more cooperative than men in bargaining (Walters, Stuhlmacher, and Meyer 1998). We do observe the extent to which men and women are cooperative in terms of our agreeableness measure, and our data and estimation results corroborate the finding that women are, on average, more cooperative than men

when bargaining over their wages.[17] This is an argument based on differences in traits and does not explain why, among equally agreeable men and women, only men benefit from being uncooperative. Where does the difference in parameters then come from? One possible explanation could be treatment differentials in the form of discriminating employers offering lower wages to women (Säve-Söderbergh 2003). Another explanation, based on differences in tastes (Niederle and Vesterlund 2005), could be that men simply tend to like bargaining more than do women who are equally (un)cooperative and skilled in bargaining.

Concluding Remarks

We have estimated the effect of personality on male-female earnings using the Five-Factor Model of personality structure as a comprehensive organizing framework. The personality traits we examined are extroversion, agreeableness, conscientiousness, neuroticism, and openness to experience. Our results indicate that (a) men who were antagonistic, open, and, to a lesser extent, emotionally stable enjoyed earnings advantages over otherwise similar men; (b) women received a premium for being more conscientious and open; (c) returns to non-agreeableness were very different for men and women; but (d) the positive returns to openness were very similar across gender, suggesting that being creative, unconventional, and artistic was equally important for the men and women we observed.

We would like to emphasize once more that our empirical findings require careful interpretation. The main reason is that causality may be reversed. We assume that personality affects earnings, but we cannot rule out the possibility that earnings also affect personality. Despite this well-founded endogeneity concern, we still believe that an exploratory study like this one can help to develop our understanding of the role of personality in the labor market.

Last but not least, we should stress that our results are specific to a highly educated group of mainly white men and women, raised in Wisconsin, who were in their early fifties about a decade ago. Traits that are important for this particular group are not necessarily relevant for current generations in the labor market.

Those caveats stated, let us take one step back and evaluate what we have found. Our results indicate that personality matters and that its impact on earnings is comparable to that of cognitive ability. Its contribution in explaining the variance in observed hourly earnings, much like measured intelligence, is rather modest. Our analyses show that the joint influence of the FFM trait variables is considerably weaker than that of education, for example. This holds even though we have made considerable allowance for unreliability in the measurement of personality and despite the fact that our estimates are likely to be upper bounds of true effects. Nonetheless, the effects we find are not trivial. Their comparability in magnitude to the earnings effects of cognitive ability is noteworthy, given the recognized place in the literature of the latter.

When economists talk about the importance of abilities, usually they are referring to unobserved abilities that may bias the estimated return to schooling or discussing measures of cognitive ability and their effects on outcomes like schooling and earnings. Such incidental regard is clearly less than personality traits merit. They are interesting in their own right, and not just as confounding factors in estimates of the returns to schooling.

Appendix A
Measuring Unobserved Traits and Classical Measurement Error

$A_1, \ldots, A_k$ are observed scores on k items, all designed to measure the same but unobserved trait A. The following relationship is used to link these observed variables to the unobserved trait:

$$A_i = A + e_i \text{ for } i = 1, \ldots, k$$

The observed measure is decomposed into its true value A and a classical measurement error e_i that is uncorrelated with A and e_j. If $\mathrm{Var}(e)$ is the variance of the measurement error, assumed to be identical for all i, and if $\mathrm{Var}(A)$ represents the variance of the true trait, the covariance matrix can be written as

$$\begin{array}{cccc} & A_i & \ldots & A \\ A_i & \mathrm{Var}(A) + \mathrm{Var}(e) & & \mathrm{Var}(A) \\ \vdots & & \ddots & \\ A_k & \mathrm{Var}(A) & & \mathrm{Var}(A) + \mathrm{Var}(e) \end{array}$$

The reliability ratio of any available measure of A represents the fraction of the variance in the observed measure of A that is due to the true variation in A,

$$\frac{\mathrm{Var}(A)}{\mathrm{Var}(A_i)} = \frac{\mathrm{Var}(A)}{\mathrm{Var}(A) + \mathrm{Var}(e)},$$

which, in this model, is identical to the correlation between any two measures

$$\rho_{ij} = \frac{\mathrm{Var}(A)}{\mathrm{Var}(A) + \mathrm{Var}(e)} = \rho.$$

The reliability ratio of the average score $\bar{A} = (A_1 + \ldots + A_k)/k$ is defined by

$$\frac{\mathrm{Var}(A)}{\mathrm{Var}(\bar{A})} = \frac{\mathrm{Var}(A)}{\mathrm{Var}(A) + (\mathrm{Var}(e)/k)} = \frac{kp}{1 + (k-1)\rho}.$$

It is easy to see that the impact of measurement error is reduced when we use all available measures of A, not just one. If we can consistently estimate ρ, we also obtain a consistent estimate of the reliability ratio by simply substituting the estimated ρ in the previous equation.

Appendix B
Correcting the OLS Estimates

Consider the simple model

$$Y = \beta A + \varepsilon,$$

where Y represents a measure for earnings, β measures the effect of A on earnings, and ε is an error independent of A. For simplicity, we ignore other covariates and suppress all subscripts that indicate that variables are measured for individuals. We are interested in parameter estimation when A is an unobserved variable. We observe $\bar{A}$ instead. The effect of regressing outcome Y on $\bar{A}$ rather than on A,

$$Y = \beta \bar{A} + \varepsilon,$$

provides the following least square estimator:

$$\hat{\beta}_{OLS} = \frac{\mathrm{Cov}(Y,\bar{A})}{\mathrm{Var}(\bar{A})} = \beta_{OLS} \frac{\mathrm{Var}(A)}{\mathrm{Var}(\bar{A})},$$

which is inconsistent. The least squares regression coefficient is attenuated by an amount equal to the reliability ratio. We already mentioned that data on all observed measures $A_1, \ldots, A_k$ allow us to measure the reliability ratio and therefore to identify the effect of A on earnings.

REFERENCES

Allport, Gordon W., and Henry S. Odbert. 1936. "Trait-Names: A Psycho-Lexical Study." *Psychological Monographs,* Vol. 47, No. 211 (entire number).

Andrisanni, Paul J. 1978. *Work Attitudes and Labor Market Experience.* New York: Praeger.

Babcock, Linda, and Sara Laschever. 2003. *Women Don't Ask: Negotiation and the Gender Divide.* Princeton: Princeton University Press.

Badgett, M. V. Lee, and Nancy Folbre. 2003. "Job Gendering: Occupational Choice and the Marriage Market." *Industrial Relations,* Vol. 42, No. 2 (April), pp. 270–98.

Barrick, Murray R., and Michael K. Mount. 1991. "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis." *Personnel Psychology,* Vol. 44, No. 1 (Spring), pp. 1–26.

Becker, Gary S. 1975. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* New York: Columbia University Press.

Bertrand, Marianne, and Kevin F. Hallock. 2001. "The Gender Gap in Top Corporate Jobs." *Industrial and Labor Relations Review,* Vol. 55, No. 1 (October), pp. 3–21.

Bertrand, Marianne, and Sendhil Mullainathan. 2001. "Do People Mean What They Say? Implications for Subjective Survey Data." *American Economic Review,* Vol. 91, No. 2 (May), pp. 67–72.

Bouchard, Thomas J., Jr., and John C. Loehlin. 2001. "Genes, Evolution, and Personality." *Behavioral Genetics,* Vol. 31, No. 3 (May), pp. 243–73.

Boudreau, John W., Wendy R. Boswell, and Timothy A. Judge. 2001. "Effects of Personality on Executive Career Success in the United States and Europe." *Journal of Vocational Behavior,* Vol. 58, No. 1 (February), pp. 53–81.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature,* Vol. 39, No. 4 (December), pp. 1137–76.

Bretz, Robert D., and Timothy H. Judge. 1994. "Person-Organization Fit and the Theory of Work Adjustment: Implications for Satisfaction, Tenure, and Career Success." *Journal of Vocational Behavior,* Vol. 44, No. 1 (February), pp. 32–54.

Buros, Oscar K. 1959. *The Fifth Mental Measurements Yearbook.* Highland Park, N.J.: Gryphon.

Carneiro, Pedro, and James J. Heckman. 2004. "Human Capital Policy." In James J. Heckman, Alan B. Krueger, and Benjamin M. Friedman, eds., *Inequality in America: What Role for Human Capital Policy?* Cambridge, Mass.: MIT Press.

Caspi, Avshalom, and Brent W. Roberts. 1999. "Personality Continuity and Change across the Life Course." In Lawrence A. Pervin and Oliver P. John, eds., *Handbook of Personality: Theory and Research.* New York: Guilford.

Conley, James J. 1984. "The Hierarchy of Consistency: A Review and Model of Longitudinal Findings of Adult Individual Differences in Intelligence, Personality, and Self-Opinion." *Personality and Individual Differences,* Vol. 5, No. 1, pp. 11–25.

Costa, Paul T., and Robert R. McCrae. 1988. "Personality in Adulthood: A Six-Year Longitudinal Study of Self-Reports and Spouse Ratings on the NEO Personality Inventory." *Journal of Personality and Social Psychology,* Vol. 54, No. 5 (May), pp. 853–63.

____. 1992. *NEO PI-R Professional Manual.* Odessa, Fl.: Psychological Assessment Resources.

____. 1994. "Set Like Plaster? Evidence for the Stability of Adult Personality." In Todd F. Heatherton and Joel L. Weinberger, eds., *Can Personality Change?* Washington, D.C.: American Psychological Association.

____. 1997. "Longitudinal Stability of Adult Personality." In Robert Hogan, John Johnson, and Stephen Briggs, eds., *Handbook of Personality Psychology.* San Diego, Calif.: Academic.

Cronbach, Lee J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika,* Vol. 16, No. 3 (September), pp. 297–334.

Digman, John M. 1989. "Five Robust Trait Dimensions: Development, Stability, and Utility." *Journal of Personality,* Vol. 57, No. 2 (June), pp. 195–214.

____. 1990. "Personality Structure: Emergence of the Five-Factor Model." *Annual Review of Psychology,* Vol. 41, pp. 417–40.

Duncan, Greg J., and Rachel Dunifon. 1998. "Soft-Skills and Long-Run Labor Market Success." *Research in Labor Economics,* Vol. 17, pp. 123–49.

Dunifon, Rachel, Greg J. Duncan, and Jeanne Brooks-Gunn. 2001. "As Ye Sweep, So Shall Ye Reap." *American Economic Review,* Vol. 91, No. 2 (May), pp. 150–54.

England, Paula. 1982. "The Failure of Human Capital Theory to Explain Occupational Sex Segregation." *Journal of Human Resources,* Vol. 17, No. 3 (Summer), pp. 358–70.

Filer, Randall K. 1981. "The Influence of Affective Human Capital on the Wage Equation." In Ronald Ehrenberg, ed., *Research in Labor Economics,* Vol. 4, pp. 367–416. Greenwich, Conn.: JAI Press.

____. 1986. "The Role of Personality and Tastes in Determining Occupational Structure." *Industrial and Labor Relations Review,* Vol. 39, No. 3 (April), pp. 412–24.

Funder, David C. 2001. "Personality." *Annual Review of Psychology,* Vol. 52, pp. 197–221.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics,* Vol. 118, No. 3 (August), pp. 1049–74.

Goldberg, Lewis R. 1990. "An Alternative Description of Personality: The Big Five Factor Structure." *Journal of Personality and Social Psychology,* Vol. 59, No. 6 (December), pp. 1216–29.

____. 1992. "The Development of Markers for the Big-Five Factor Structure." *Psychological Assessment,* Vol. 4, No. 1 (March), pp. 26–42.

Goldsmith, Arthur H., Jonathan R. Veum, and William Darity, Jr. 1997. "The Impact of Psychological and Human Capital on Wages." *Economic Inquiry,* Vol. 35, No. 4 (October), pp. 815–29.

Heckman, James J., and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review,* Vol. 91, No. 2 (May), pp. 145–49.

Holland, John L. 1997. *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments.* Odessa, Fl.: Psychological Assessment Resources.

Jencks, Christopher. 1979. *Who Gets Ahead?* New York: Basic Books.

John, Oliver P., Eileen M. Donahue, and Robert L. Kentle. 1991. *The Big Five Inventory—Versions 4a and 54.* Berkeley: University of California, Institute of Personality and Social Research.

John, Oliver P., and Sanjay Srivastava. 1999. "The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives." In Lawrence A. Pervin and Oliver P. John, eds., *Handbook of Personality: Theory and Research.* New York: Guilford.

Judge, Timothy A., and Daniel M. Cable. 1997. "Applicant Personality, Organizational Culture, and Organization Attraction." *Personnel Psychology,* Vol. 50, No. 2 (Summer), pp. 359–94.

Judge, Timothy A., Daniel Heller, and Michael K. Mount. 2002. "Five-Factor Model of Personality and Job Satisfaction: A Meta-Analysis." *Journal of Applied Psychology,* Vol. 87, No. 3 (June), pp. 530–41.

Judge, Timothy A., Chad A. Higgins, Carl J. Thoresen, and Murray R. Barrick. 1999. "The Big Five Personality Traits, General Mental Ability, and Career Success across the Life Span." *Personnel Psychology,* Vol. 52, No. 3 (Autumn), pp. 621–52.

McCrae, Robert R., and Paul T. Costa. 1997. "Personality Trait Structure as a Human Universal." *American Psychologist,* Vol. 52, No. 2 (May), pp. 509–16.

____. 1999. "A Five-Factor Theory of Personality." In Lawrence A. Pervin and Oliver P. John, eds., *Handbook of Personality: Theory and Research.* New York: Guilford.

McCrae, Robert R., and Oliver P. John. 1992. "An Introduction to the Five-Factor Model and Its Applications." *Journal of Personality,* Vol. 60, No. 2 (June), pp. 175–215.

Mincer, Jacob. 1958. "Investment in Human Capital and Personal Income Distribution." *Journal of Political Economy,* Vol. 66, No. 4 (August), pp. 281–302.

Mueller, Gerrit, and Erik J. S. Plug. 2004. "Estimating the Effect of Personality on Male-Female Earnings." IZA discussion paper 1254, Bonn.

Neumark, David. 1988. "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination." *Journal of Human Resources,* Vol. 23, No. 3 (Summer), pp. 279–95.

Niederle, Muriel, and Lise Vesterlund. 2005. "Do Women Shy Away from Competition? Do Men Compete Too Much?" NBER working paper 11474.

Norman, Warren T. 1963. "Toward an Adequate Taxonomy of Personality Attributes: Replicated Factor Structure in Peer Nomination Personality Ratings." *Journal of Abnormal and Social Psychology,* Vol. 66, No. 6, pp. 574–83.

Nyhus, Ellen K., and Empar Pons. 2005. "The Effects of Personality on Earnings." *Journal of Economic Psychology,* Vol. 26, No. 3 (June), pp. 363–84.

Oaxaca, Ronald L. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review,* Vol. 14, No. 3 (October), pp. 693–709.

Oaxaca, Ronald L., and Michael R. Ransom. 1994. "On Discrimination and the Decomposition of Wage Differentials." *Journal of Econometrics,* Vol. 61, No. 1 (March), pp. 5–21.

Osborne, Melissa. 2003. "The Power of Personality: Labor Market Rewards and the Transmission of Earnings." Doctoral diss., University of Massachusetts–Amherst.

Polachek, Solomon W. 1981. "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure." *Review of Economics and Statistics,* Vol. 63, No. 1 (February), pp. 60–69.

Säve-Söderbergh, Jenny. 2003. "Are Women Asking for Too Low Wages? – Gender Wage Differentials and Individual Wage Bargaining." Dissertation Series, No. 59, Swedish Institute for Social Research, Stockholm University.

Sewell, William H., Robert M. Hauser, Kristen W. Springer, and Taissa S. Hauser. In press. "As We Age: The Wisconsin Longitudinal Study, 1957–2001." In Kevin T. Leicht, ed., *Research in Social Stratification and Mobility.* Greenwich, Conn.: JAI Press. (Available as CDE Working Paper 2001–09, Center for Demography and Ecology, University of Wisconsin-Madison.)

Sternberg, Robert J., and Patricia Ruzgis. 1994. *Personality and Intelligence.* New York: Cambridge University Press.

Tett, Robert P., Douglas N. Jackson, and Mitchell Rothstein. 1991. "Personality Measures as Predictors of Job Performance: A Meta-Analytic Review." *Personnel Psychology,* Vol. 44, No. 4 (Winter), pp. 703–42.

Tharenou, Phyllis. 1997. "Managerial Career Advancement." *International Review of Industrial and Organizational Psychology,* Vol. 12, pp. 39–93.

Tharenou, Phyllis, Shane Latimer, and Denise Conroy. 1994. "How Do You Make It to the Top? An Examination of Influences on Women's and Men's Managerial Advancement." *Academy of Management Journal,* Vol. 37, No. 4 (August), pp. 899–931.

Tokar, David M., Ann R. Fischer, and Linda Mezydlo Subich. 1998. "Personality and Vocational Behavior: A Selective Review of the Literature, 1993–1997." *Journal of Vocational Behavior,* Vol. 53, No. 2 (October), pp. 115–53.

Tupes, Ernest C., and Raymond E. Christal. 1961. "Recurrent Personality Factors Based on Trait Ratings." USAF ASD Tech. Rep. No. 61–97. Lackland Air Force Base, Tex.: U.S. Air Force.

Turner, Charles F., and Daniel C. Martinez. 1977. "Socioeconomic Attainment and the Machiavellian Personality." *Sociometry,* Vol. 40, No. 4 (December), pp. 325–36.

Walters, Amy E., Alice F. Stuhlmacher, and Lia L. Meyer. 1998. "Gender and Negotiator Competitiveness: A Meta-Analysis." *Organizational Behavior and Human Decision Processes,* Vol. 76, No. 1 (October), pp. 1–29.

Wellbourne, Theresa M., Marcie A. Cavanaugh, and Timothy A. Judge. 1998. "Does the Leader Make a Difference? Relationship between Executive Leader Personality and Entrepreneurial Firm Performance." CAHRS working paper, Cornell University.

Zax, Jeffrey S., and Daniel I. Rees. 2002. "IQ, Academic Performance, Environment, and Earnings." *Review of Economics and Statistics,* Vol. 84, No. 4 (November), pp. 600–616.