Labour Economics 51 (2018) 170–183

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Personality, IQ, and lifetime earnings

Miriam Gensowski [a,b,1]

[a] Center for Economic Behavior and Inequality (CEBI) and Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Copenhagen K 1353, Denmark
[b] IZA, Institute of Labor Economics, Germany

ARTICLE INFO

JEL classification:
J24
I24
J16

Keywords:
Personality traits
Socio-emotional skills
Cognitive skills
Returns to education
Lifetime earnings
Big Five
Human capital
Factor analysis

ABSTRACT

This paper estimates the effects of personality traits and IQ on lifetime earnings of the men and women of the Terman study, a high-IQ U.S. sample. Age-by-age earnings profiles allow a study of *when* personality traits affect earnings most, and *for whom* the effects are strongest. I document a concave life-cycle pattern in the payoffs to personality traits, with the largest effects between the ages of 40 and 60. An interaction of traits with education reveals that personality matters most for highly educated men. The largest effects are found for Conscientiousness, Extraversion, and Agreeableness (negative), where Conscientiousness operates partly through education, which also has significant returns.

A growing economics literature documents effects of socio-emotional skills, often called non-cognitive skills, on life outcomes ranging from wages to health—see summaries in Borghans et al. (2008) or Almlund et al. (2011). Labor market outcomes, in particular, have been shown to be influenced by skills such as self-control (Moffitt et al., 2011), Conscientiousness (Prevoo and ter Weel, 2015; Uysal and Pohlmeier, 2011), Self-esteem, or Locus of Control (Caliendo et al., 2015; Heckman et al., 2006b). This paper contributes to the body of work that studies how earnings are affected by personality traits. Personality trait measures, such as the Big Five (McCrae and John, 1992), are a popular way of proxying socio-emotional skills. This paper provides evidence on *when* in the life cycle personality traits are most important and *for whom* they have the largest effects.

The data that make this analysis possible come from the seminal Terman study (Terman, 1992). This survey was initiated in 1922 in California and followed a group of high-IQ men and women from childhood to old age. While it has been widely used for research in psychology, this paper is the first to have generated earnings profiles for ages 18 to 75 from the different waves. It combines measures on IQ and personality traits in early waves with a very long follow-up. The Terman study also contains rich background information on each participant.

The question of *when* personality traits matter can be addressed with the detailed age-by-age earnings measures. For most traits, the earnings effects have a hump-shaped pattern: early in these men's careers, the effects of personality traits are barely visible, but become large in their prime working years. Insofar as this life-cycle pattern is due to general mechanisms that are not specific to high-IQ individuals, these analyses could, for example, inform the forecasting of lifetime effects of skill-building interventions that target socio-emotional skills related to the personality traits observed here.

To test for heterogeneous effects of traits on earnings, I interact personality traits with education. I find statistically and economically meaningful interactions. The payoffs to two important traits, Conscientiousness and Extraversion, are more than twice as large for men with a graduate degree than for men with a bachelor's or less. Another interpretation of this interaction is that the earnings gain from higher education is larger for men who possess stronger socio-emotional skills. Most of the existing studies do not allow for a trait-education interaction, and

E-mail address: Miriam.Gensowski@econ.ku.dk

[1] The activities of CEBI are financed by a grant from the Danish National Research Foundation. This paper and other versions of it have benefited immensely from discussions with numerous seminar participants. I am very grateful to everyone who provided helpful feedback, especially James J. Heckman, Steven N. Durlauf, and Gary S. Becker, Rémi Piatek, Mathilde Almlund, Katarína Borovičková, Deborah Cobb-Clark, Pia Pinger, Stefanie Schurer, Jeffrey Smith, Lucas Threinen and Philipp Eisenhauer. I am also indebted to Peter Savelyev, who generously shared his knowledge of the Terman data, and to Min Ju Lee and Molly Schnell who provided outstanding research assistance in the project's earlier stages.

https://doi.org/10.1016/j.labeco.2017.12.004
Received 9 May 2016; Received in revised form 11 December 2017; Accepted 11 December 2017
Available online 14 December 2017
0927-5371/

may therefore over- or under-estimate these effects of personality traits conditional on education.

With the Terman survey, the relationship between personality traits and earnings can be studied in a more detailed way than is possible elsewhere, and it fills out our understanding of the age-pattern and the interaction with education. Yet the Terman sample is not representative of the general population, and was never intended to. Therefore, this study also adds to our knowledge about what determines lifetime earnings of individuals with top IQs—usually, sample sizes are too small to identify the intellectual elite. For example, it is not clear whether high-IQ children would benefit from improving socio-emotional skills. Many socio-emotional-skill building interventions are targeted at disadvantaged, and sometimes low-IQ, populations (s.a. Grossman and Tierney, 1998; Heckman et al., 2010; 2013; University of Chicago Crime Lab, 2012). This paper shows that high-IQ children also significantly benefit from positive personality traits later in life, and that they can expect positive returns to education.

The study of the high-IQ women in the Terman data is of interest as well, albeit of a rather historical nature, as they can be less easily compared to current cohorts. Only about half of the women of the Terman sample were securely attached to the labor force, and many relied on husbands as bread-winners. I therefore study women's family earnings, and demonstrate how women's own and husband's earnings reacted differently to women's personality traits.

Methodologically, this paper addresses a common problem to research on personality traits: whenever personality scores are used as regressors, measurement error bias is introduced because true personality is always unobserved. Instead, predicted scores of personality traits are used, and the prediction will contain some noise. An adjustment has been suggested (Bolck et al., 2004; Croon, 2002), which I develop further to apply it to a setting where the variable measured with error is interacted with an indicator for education.

1. The Terman survey

The Terman survey was initiated by the prominent psychologist Lewis Terman to study the life outcomes of high IQ children. His team canvassed all schools in California, grades 1–8, in 1921–1922, to enrol children who scored in the top 0.5% of the IQ distribution. The sample consists of 856 boys and 672 girls, born around 1910, and who were followed until 1991, with surveys every 5–10 years.[2] It is the longest prospective cohort study that also has data on earnings.

The Terman data have been used extensively by psychologists to study health and longevity, in relation to Conscientiousness and parental divorce or marriage.[3] Only few economists have worked with the data, focusing on family outcomes (marriage, divorce, fertility - see Becker et al., 1977; Michael, 1976; Tomes, 1981), retirement behavior (Hamermesh, 1984), and health (Savelyev, 2014; Savelyev and Tan, 2015). Earnings outcomes were studied by Leibowitz (1974), but she did not exploit the longitudinal data.

Drawing on the different waves of the survey, I construct earnings histories from age 18 to 75, as well as education and marriage profiles, for each participant. The age-by-age information stems from the feature that for many of the waves, respondents were asked about earnings in each of the 4 previous years separately. Earnings are imputed through linear interpolation for years without information. The earnings measures for all estimations are annual earnings after tuition in 2010 U.S. Dollars (CPI adjusted), truncated at the 97th percentile, before tax.[4] For inactive workers, as well as for the deceased, earnings are zero. For female participants, the Terman survey asked about their spouse's earned income. Family earnings can thus be constructed as the sum of own earnings and the husband's earnings, which are zero if the woman was not married.

The personality information in the Terman data stems from teachers and parents, who rated the participants on certain traits and behaviors in 1922, and from participants, who provided self-ratings on other items in 1940 (at around age 30). An exploratory factor analysis on all available items[5] reveals a structure that is remarkably similar to traits in the well-known Big Five taxonomy: **O**penness to Experience, **C**onscientiousness, **E**xtraversion, **A**greeableness, and **N**euroticism (OCEAN). Even though the Terman measures were taken about 70 years before the Big Five were codified (Goldberg, 1993), the factors correspond closely to these traits, measured for example by the NEO PI-R (Martin and Friedman, 2000).

Openness to Experience, the "tendency to be open to new aesthetic, cultural, or intellectual experiences" (American Psychological Association, 2007), was measured in 1922 by ratings from teachers and parents on descriptors such as "desire to know" or "originality." Extraversion was indicated by the subject's "fondness for large groups," "leadership," and "popularity with other children," also in 1922. The remaining traits are based on self-ratings in 1940. Conscientiousness describes an individuals' persistence, order, and need for achievement. In Terman, it is measured with "How persistent are you in the accomplishment of your ends?" or "In your work do you usually drive yourself steadily?". Agreeableness describes cooperation and a preference for harmonious relationships over antagonistic behavior. An example measure is "In general, how easy are you to get on with?". Neuroticism, the opposite of emotional stability, is based on questions such as "Are you moody?". These personality traits are summarized by factor scores (Jöreskog and Sörbom, 1979; Mulaik, 2010), and predicted with the Bartlett method (Bartlett, 1937). Each item is allowed to load on exactly one factor, and this dedicated factor structure guarantees identification of possibly correlated factors. In a few cases where not all personality items are observed, they are imputed with a multiple imputation routine exploiting the covariance with the other factors.

IQ was measured at study entry in 1922.[6] Scoring at 140 or higher, which corresponds to being in the top 1 of 200 children, was the criterion for being included in the study. Even though the Terman survey is selective in terms of IQ, it is not so for personality, as Martin and Friedman (2000) show. Generally, personality traits correlate only weakly

[2] Attrition is below 10%, and it is unrelated to income, education, demographic factors (Sears, 1984), or psychological measures (Friedman et al., 1993).

[3] Cf Friedman (2008); Friedman et al. (1993); Martin et al. (2005); Tucker et al. (1996).

[4] A Web Appendix to this paper, hosted at http://www.econ.ku.dk/gensowski/research/Terman/TermanApp.pdf, contains more detailed information on the data construction, estimation, and supplementary figures and tables. Section A describes the construction of the earnings profiles and tuition costs, and Section A.7 shows their distributions.

[5] The teacher and parent ratings are averaged within each item. In exploratory factor analysis, the researcher observes the covariance structure of the items, and determines the number of factors that capture most of the observed variation, as well as which items are associated with which factor. For the full list of items of all traits, see Section B.2 in the Web Appendix.

[6] The standard test was the Stanford–Binet IQ test that Terman himself had recently developed (Terman, 1916). Some of the participants took the closely related "Terman Group Test", specifically designed for screening these high achieving children (see Chapter I in Terman and Sears, 2002). Its scale was such that scores are comparable. In the subsequent analyses, I always allow for the possibility that there were differences between the two measures of IQ, by including an interaction with test type. The coefficients of the two tests are never statistically different from each other. The well-known Stanford–Binet IQ test has naturally undergone updates throughout its life time, notably to make it less verbally loaded, to measure domain-specific ability (such as verbal vs. quantitative), and to extend the age ranges to children younger than six and adults (overview in Becker, 2003). The latter is not of concern to the sample here, as the students were in the appropriate age range the test was designed for. The worry about a strong verbal content in an IQ test is that it puts children from non-native English households or from different cultural backgrounds at a disadvantage. In the selected sample at hand, this would imply that these usually disadvantaged children would have a higher IQ than their score lets us believe. In all analyses, this paper controls for parental immigrant status, and excludes non-Caucasian participants to address this potential bias. The original Stanford–Binet was only concerned with assessing general ability, which is conceptualized as the aggregate of domain-specific abilities. It is therefore the ultimate summary measure. Given that the current analysis mostly views IQ, within a rather restricted range, as a control variable, controlling for the general version seems appropriate, making the loss of specificity a small one.