Personality, Education, and Health-Related Outcomes of High-Ability Individuals*

Peter Savelyev† Vanderbilt University

Kegon Teng Kok Tan‡ University of Wisconsin–Madison

July 31, 2014

PRELIMINARY AND INCOMPLETE

*A version of this paper was presented to the 5th ASHEcon Conference in Los-Angeles, to the Empirical Micro Lunch at the University of Wisconsin at Madison, to the Vanderbilt Empirical Applied Micro Work-In-Progress Lunch, and to the Singapore Economic Review Conference. We thank participants of these meetings for useful suggestions and stimulating discussions. We are grateful to Erik Meijer for many detailed and helpful comments. Support from the Grey Funds at the Department of Economics, Vanderbilt University, is gratefully acknowledged by Peter Savelyev. Support of the Human Capital and Economic Opportunity Global Working Group, sponsored by the Institute for New Economic Thinking, is gratefully acknowledged by Kegon Tan. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the funders.

†Peter Savelyev, the corresponding author, is an Assistant Professor of Economics at Vanderbilt University and a Health Policy Associate of the Robert Wood Johnson Center for Health Policy at Meharry Medical College. Address: Department of Economics, 415 Calhoun Hall, Vanderbilt University, Nashville, TN 37235-0002. Email: peter.savelyev@vanderbilt.edu

‡Kegon Tan is a Ph.D. candidate at the University of Wisconsin–Madison, Department of Economics. Address: William H. Sewell Social Science Building, 1180 Observatory Drive, Madison, WI 53706-1393. Email: ttan8@wisc.edu

Abstract

We estimate the effects of five latent personality skills and college education on determinants of longevity, including health behaviors, lifestyles, lifetime earnings, as well as general and mental health. The latent personality skills are closely related to the well-established contemporary Big Five taxonomy of personality. We motivate this study by a theoretical framework, which shows multiple channels through which personality and education may affect health behaviors and lifestyles. We employ the Terman life-cycle data of children with high ability (1922–1991), which is uniquely suited to studying the developmental and behavioral origins of health and longevity. We uncover plausible mechanisms behind strong treatment effects of personality and college education on longevity documented in Savelyev (2014). We account for measurement error via factor-analytic methods and control for multiple hypotheses using a new version of the Holm-Bonferroni method with superior power proposed by Romano and Wolf (2005). We find strong effects of personality skills and education on health and health-related outcomes. The effects of education and the five personality skills differ by gender and outcome, demonstrating substantial heterogeneity in the role of multiple human skills in generating health. Variance explained by the five latent personality skills is comparable to the variance explained by all observable determinants taken together including education, IQ, early health, and family background.

Key words: post-compulsory education, cognitive skills, personality skills, Big Five personality taxonomy, health, health behaviors, health-related consumption, lifestyles, earnings, multiple-hypothesis testing, stepdown procedure, Terman data of children with high ability, gender difference

JEL codes: D91, I12, J24

Contents

List of Figures

List of Tables

1 Introduction

There is evidence documented in the literature that both cognitive and personality skills, as well as college education have an effect on longevity (e.g., Buckles et al., 2013; Savelyev, 2014). However, the mechanisms of such effects are still not well understood, especially those for personality skills. This paper offers a theoretical framework and empirical results that shed light on the mechanisms. We establish the effects of post-compulsory education, IQ, and personality skills on the following various outcomes that, according to the literature, are plausible determinants of longevity: health-related consumption, lifestyles, and earnings. Effects on these outcomes are of interest both in their own right and as part of longevity production.

Health economics papers examining the causal effect of education on health, health-related outcomes, or longevity largely use various natural experiments as a source of identification. A number of these papers are at odds with each other even though they use the same identification strategy. For instance, while some papers claim a causal effect of education on health or longevity (Grossman, 2004; Grossman and Kaestner, 1997; Lleras-Muney, 2005; van Kippersluis et al., 2011), some others find that there is hardly any effect (Albouy and Lequien, 2009; Clark and Royer, 2013; Mazumder, 2008). The contradiction is partly due to the serious limitations of natural experiments. Indeed, validity and monotonicity of instrumental variables are often strong assumptions.[1] Instruments are often weak. Loss of estimation precision relative to OLS is substantial. Also, effect is identified only for a specific sub-population that is induced by the instruments to change behavior, leading to an estimation of the effect that is not necessarily policy-relevant (Heckman and Vytlacil, 2007).

In this paper we take an approach that is an alternative to natural experiments, which we believe to be a useful source of additional information, given the limitations of natural experiments and the existing controversy. Our identification strategy is based on explicit factor modeling of latent skills that may contribute to the ability bias. Similar to Heckman et al. (2006), we assume that, conditional on cognitive skills, five latent person-

[1] As we know, the regression discontinuity estimator is a special case of IV.

ality skills, and an extensive set of theoretically relevant controls,[2] the choice of education is independent of health-related potential outcomes for those uneducated and educated. The conditional independence assumption invoked for the identification of causal effects is similar to the one used in the matching literature; but, unlike in matching, which only controls for observables, we additionally control for a set of comprehensive multi-dimensional latent personality skills using factor-analytic methods.[3] We acknowledge limitations of this approach.

Recent papers by Conti, Heckman, and Urzúa (2010a,b) present results regarding the relationship between education, personality, and a similar set of health outcomes using data from the British Cohort Study. The authors calculate the causal effects of education, cognition, and a one-dimensional representation of personality skills on health behaviors. We join Conti, Heckman, and Urzúa (2010a,b) in supporting the claim that personality skills play a key role in determining health-related outcomes, and complement their research by emphasizing the multi-dimensional aspect of personality and the heterogeneity of the effects of personality skills that is not captured by lower dimension representations. Further, we address the issue of multiple-hypothesis testing due to the large number of similar outcomes explored. We also analyze a dataset with richer background information and with a much longer follow-up (ages 86 in the Terman data vs. 42 in the British Cohort Study).

Literature in psychology reports multiple correlations between the Big Five personality skills and health-related outcomes.[4] In particular, Conscientiousness is strongly correlated with beneficial health behaviors and other health-related outcomes, while Neuroticism is strongly correlated with harmful health behaviors (Friedman, 2000; Goodwin and Friedman, 2006). Our work supplements this literature by providing estimates that can be interpreted as causal, by accounting for measurement error and multiple-hypothesis testing.

[2]Our uniquely detailed controls include parental occupation, employment, and education information; early parental and private tutoring; early health measures; health rating in childhood; early divorce or death of parents; ratings of family well-being; and social status (see Table 2).

[3]The Big Five is one of the most established taxonomies of personality that attempts to reduce multiple human skills to several latent factors (John and Srivastava, 1999).

[4]See, e.g, Friedman (2000, 2008); Friedman et al. (1994, 1995, 1993); Hampson and Friedman (2008); Martin et al. (2007, 2002).

We use the Terman life-cycle data of children with high ability. Boys and girls born around 1910 were selected from schools in California for their high IQ (above 140). The data prospectively covers the period from 1922 to 1991, and combines high-quality measures of IQ and personality obtained around age 12 with life-cycle measurements of health-related outcomes. The combination of early measures of psychological skills with a life-cycle follow-up is unique and ideally suited for studying developmental origins of health.

For each outcome of interest, we estimate a linear in parameters outcome equation that accounts for education, IQ, latent personality skills, and a set of observable controls together with a system of factor measurement equations that links latent skills to multiple noisy psychological measures. We account for the childhood personality skills of Conscientiousness, Openness, and Extraversion using a set of ratings given by teachers and parents; and we augment those with self-reported early adulthood Agreeableness and Neuroticism to complete the Big Five model of personality. While in the economics literature the problem of multiple-hypothesis testing is largely ignored, in this paper we adjust each single p-value to strongly account for family-wise error rate following Romano and Wolf (2005).

We find that the effects of personality and education differ considerably across genders. For males we find many effects of each skill on multiple health-related outcomes and health behaviors. Overall, at least for behaviors and measures that we observe, we can see that Conscientiousness and Education are good for heath through multiple health-related outcomes, while Neuroticism is bad. Extraversion shows mixed effects by increasing earnings and mental health, while also increasing heavy drinking. IQ for this high-ability group shows only small, mixed effects. The only strong potential link between Agreeableness and health is its negative effect on earnings. For females we observe smaller number of strong and statistically significant links between skills and health. The links through multiple channels for females are observed only for Education and Neuroticism, but the number of potential mediators is smaller than for males.

The rest of the paper is organized as follows: Section 2 describes the data , Section 3 elaborates on the Big Five taxonomy of personality, Section 4 introduces the theoretical

and statistical framework for our estimations, Section 5 presents the results, Section 6 provides a discussion of empirical findings, and Section 7 concludes.

2 Data Description

Research presented in this paper is based on the Terman Life-Cycle Data of Children with High Ability (Terman, 1986). This dataset prospectively follows a group of 1,583 high-IQ individuals from 1922 to 1991. It thus allows us to look at the effect of education and personality on essential health-related outcomes at multiple points in the life cycle. The availability of early childhood and early adulthood personality measures enables the construction of personality skills that are close to the contemporary and well-established Big Five taxonomy of personality (Martin and Friedman, 2000). The subjects were also surveyed for education data several times over the life cycle, which helps us construct reliable education data with minimal possible measurement error. The dataset includes detailed life histories, early childhood and adolescent health conditions, as well as information on parental and private tutoring and family background.

The Terman sample consists of 856 males and 672 females from public schools in California. The subjects were selected for an IQ of above 140, representing roughly the top 0.4% of the general population.[5] While the sample is homogenous in that the subjects are all highly intelligent, the personality skills show a wide variation. In fact, there is no evidence that the subjects differ significantly from the general population with regards to measures of personality (Friedman et al., 1993; Terman and Sears, 2002a). The possible exception is Openness, which is known to be linked to IQ, unlike other personality skills (Ackerman and Heggestad, 1997; Borghans et al., 2011; DeYoung et al., 2005). The Terman study has an attrition of less than 10%, which is low for a 70-year-long prospective study.

The wealth of information in the Terman data, like its low attrition, is remarkable. Some 4,500 measurements made in the period 1922–1991 describe the family back-

[5]To be more precise, 187 children had an IQ in the range 126–139, with most of them being in the range 135–139 (180 children).

ground, parental investment, personality, early health, and household economic status, among other important determinants of health behavior and education attainment of the subjects (Burks et al., 1930; Terman et al., 1925; Terman and Oden, 1959; Terman et al., 1947; Terman and Sears, 2002a,b; Terman et al., 2002).

2.1 Health Behaviors, Health Measures, and Lifestyle Characteristics

Table 1 presents health-related outcomes that we explore in this paper including health behaviors and their proxies, lifestyles, earnings, and general heath measures. Many of these outcomes were observed at multiple points of the life cycle.[6] We also use longitudinal data to construct aggregate measures of outcomes over the life cycle such as "ever drank heavily."

Education in this paper refers to whether or not the subject received a Bachelor's degree or above, an outcome that we shortly refer to as "college degree" or "college education." Although this was a particularly high-IQ sample, there was still a sizeable number of subjects (about 30% of the sample) who did not achieve a college degree.[7] The average IQ is approximately 149 for both genders. Subjects were tested using either the Stanford Binet Test or the Terman Group Test. The IQ variable is constructed by survey organizers from the two tests used and corrects for differences between them, as well as for the age of subjects at the time of testing. The rest of the variables are the background controls, which cover a wide range of underlying characteristics of the subjects, including early childhood health, key parental characteristics, parental investment in children, and cohort.

We restrict the regression sample the same way as in Savelyev (2014). The exclusions determined prior to estimation are as follows: subjects who were not born in the period 1904–1915; subjects who never participated or were lost or dropped out before

[6] We thank Miriam Gensowski for providing her calculations of earnings profiles. The life-cycle earnings measures that we use are net of tuition paid for schooling and taxes and are expressed in 2010 US dollars (see the Web Appendix of Gensowski (2013) for more details).

[7] Savelyev (2014) argues that for the Terman subjects' generation, the high school threshold was similar to today's college threshold in terms of the percentile of the population above this threshold. Back in 1920-1930s, high school was perceived as a relatively high education level and was competitive for many jobs.

Table 1: Description of Main Variables, Outcomes

	Females				Males			
Variable	mean	std. error	min	max	mean	std. error	min	max
Health Behaviors and proxies								
Heavy Drinking of Alcohol in 1940	.102	.014	0	1	.267	.018	0	1
Heavy Drinking of Alcohol in 1950	.038	.009	0	1	.118	.013	0	1
Heavy Drinking of Alcohol in 1960	.167	.018	0	1	.347	.021	0	1
Ever drank heavily, 1940–60	.205	.018	0	1	.394	.019	0	1
Ever smoked, 1991	.425	.036	0	1	.521	.037	0	1
Excercise, 1982	.173	.021	0	1	.176	.021	0	1
Abnormal BMI, 1940	.160	.017	0	1	.225	.017	0	1
Lifestyles								
Never married, 1922–86	.085	.012	0	1	.061	.009	0	1
Married once and still married, 1922–86	.417	.022	0	1	.576	.019	0	1
Ended up divorced, 1922–86	.121	.014	0	1	.064	.009	0	1
Ever divorced, 1922–86	.253	.018	0	1	.267	.017	0	1
Divorced at least twice, 1922–86	.059	.010	0	1	.063	.009	0	1
# of organizations in 1940	2.506	.083	0	4	2.435	.069	0	4
# of organizations in 1950	1.565	.078	0	10	2.714	.090	0	10
# of organizations in 1960	2.567	.124	0	10	3.423	.112	0	10
Ever a member of any organization, 1940–60	.900	.014	0	1	.937	.010	0	1
Earnings								
Earnings at age 40, 1945–57	11.54	.80	.00	116.71	62.23	1.84	.00	335.74
Earnings at age 50, 1955–67	16.08	1.01	.00	104.18	71.50	2.17	.00	294.10
Earnings at age 60, 1965–77	15.33	1.12	.00	138.82	62.00	2.58	.00	409.45
Lifetime earnings discounted at 3%, 1922–86	272.00	12.26	-18.24	1687.78	1115.88	26.16	.00	4189.50
Health Measures								
Mental difficulty in 1940	.328	.022	0	1	.278	.018	0	1
Mental difficulty in 1950	.333	.023	0	1	.272	.019	0	1
Mental difficulty in 1960	.344	.023	0	1	.294	.020	0	1
Ever had mental difficulty, 1940–60	.463	.022	0	1	.415	.019	0	1
General Health in 1940	-.016	.048	-4.24	1.11	.029	.043	-5.19	1.03
General Health in 1950	-.005	.046	-4.91	1.03	-.004	.042	-5.44	.92
General Health in 1960	.006	.048	-3.77	1.06	-.012	.045	-5.74	.97
Never poor or fair health, 1940–60	.873	.015	0	1	.926	.010	0	1
Base Estimation Sample		527				680		

Notes: [(a)]Calculations are based on the Terman data. Multi-period observations of heavy drinking, mental difficulty, and general health are summarized into binary indicators which equal 1 when negative outcomes are observed in any point in time of the life cycle, and 0 otherwise. Earnings are annual, in thousands of 2010 US dollars, net of tuition and taxes. General health variable is a standardized average of standardized general health and energy self-ratings.

1940; subjects who are missing both parent and teacher personality trait ratings in 1922; subjects who are high-school dropouts; subjects who died in service during World War II; subjects with serious diseases in their early life, such as chorea or Hodgkin's disease; subjects without education level information; and subjects who did not survive to age 30. Aside from subjects excluded due to missing data, these restrictions remove outliers[8] in the sample and help reduce reverse causality[9] between education and health.[10]

The base estimation sample contains 680 males and 529 females. All models are estimated using this base sample, but the actual estimation sample may be smaller and may vary from regression to regression due to missing information in left-hand side variables.

2.2 Main Regressors and Background Variables

The right-hand side variables are described in Table 2, which includes education, IQ, and background variables, but excludes personality measures, which are to be discussed in Section 3.

3 Understanding Personality Skills

The literature in personality psychology explores many ideas of what personality is and how to measure it. However, perhaps the most established contemporary categorization of personality is the Big Five taxonomy (John and Srivastava, 1999).[11] This taxonomy reduces the dimensionality of human personality to just five latent factors: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (see Table 3 for descriptions). This system of five personality skills is comprehensive enough to capture the multi-faceted nature of human personality, whilst remaining computationally tractable.

[8] For example, a few subjects who were high school dropouts despite extraordinary IQ.

[9] For example, subjects with serious early health problems that may have severely affected schooling choice.

[10] See Savelyev (2014) for more details.

[11] See also Digman (1990) and McCrae et al. (1986) for evidence regarding the comprehensive and rigorous nature of this taxonomy.

Table 2: Description of Main Variables[(a)], Background Controls

		Females				Males			
variable	year	mean	std. error	min	max	mean	std. error	min	max
Subject's Background									
IQ	1922	148.482	.446	135	201	149.288	.405	126	200
Bachelor's degree or above	1922–68	.686	.020	0	1	.734	.017	0	1
Extraordinary birth	1922	.629	.021	0	1	.571	.019	0	1
No breastfeeding	1922	.085	.012	0	1	.091	.011	0	1
Childhood health	1922	9.027	.083	4.5	13	8.526	.075	3	13
Childhood energy	1922	8.834	.078	4.5	13	8.219	.073	3	13
Age at 1922	1922	11.800	.122	7	18	11.838	.112	6	18
Cohort 1915–1918	1922	.172	.016	0	1	.237	.016	0	1
Cohort 1907–1910	1922	.361	.021	0	1	.296	.018	0	1
Participation in World War II	1945	-	-	-	-	.410	.019	0	1
Combatant in World War II	1945	-	-	-	-	.093	.011	0	1
Parental Background									
Mother dead	1922	.032	.008	0	1	.028	.006	0	1
Father dead	1922	.074	.011	0	1	.081	.010	0	1
Parents divorced	1922	.047	.009	0	1	.050	.008	0	1
Father's education	1922	.253	.019	0	1	.291	.017	0	1
Parental finances	1922	.384	.021	0	1	.371	.019	0	1
Parental social standing	1922	.153	.016	0	1	.253	.017	0	1
Mother working	1922	.132	.015	0	1	.126	.013	0	1
Father high-skilled	1922	.276	.019	0	1	.243	.016	0	1
Parent born abroad	1922	.267	.019	0	1	.304	.018	0	1
Parent born in Europe	1922	.202	.017	0	1	.218	.016	0	1
Duration of private tutoring (weeks)[(b)]	1922	.344	.026	.000	5.743	.105	.014	.000	4.247
Home investment (hours)[(b)]	1922	.409	.016	.000	1.839	.450	.014	.000	1.771
Estimation Sample			527				680		

Notes: [(a)]Calculations are based on the Terman data. [(b)]Durations are transformed using natural logarithm: $\ln(1 + duration)$.

Table 3: Description of Big Five Personality Skills[(a)]

Trait	Definition
1. Openness to Experience (Intellect)	The breadth, depth, originality, and complexity of individual's mental and experimental life
2. Conscientiousness	A propensity to follow socially prescribed norms for impulse control, to be task- and goal- directed, to be planfull, to delay gratification, and to follow norms and rules
3. Extraversion	An energetic approach to the social and material world, which includes traits such as sociability, activity, assertiveness, and positive emotionality
4. Agreeableness	A prosocial and communal orientation towards others (as opposed to antagonism), which includes traits such altruism, tender-mindedness, trust, and modesty
5. Neuroticism (Emotional Stability)	An emotional stability and even-temperedness as opposed to negative emotionality, such as feeling anxious, nervous, sad, and tense

Notes: [(a)]Description taken from John and Srivastava (1999).

Personality skills are not directly observed and are therefore modelled as latent factors. They are proxied by multiple psychological ratings, which are measures of specific behaviors that are deemed to be manifestations of the latent factors. In this paper, in line with previous research based on the same data (Friedman et al., 2010, 1995, 1993; Savelyev, 2014), we use a set of measures that are constructed as an average of parent and teacher ratings of child behavior and measures from early adulthood that are self-reported.[12] We present the list of measures of the five latent factors in Table 4.

These measures are summarized into five groups based on established psychometric techniques of exploratory and confirmatory factor analysis (see Heckman et al. (2013) and Savelyev (2014) for descriptions of the EFA and CFA approaches directed at economists, as well as classical references by Gorsuch (1983) and Bollen (1989)). This paper uses the childhood measures that were justified by Savelyev (2014) to best represent Conscientiousness, Openness and Extraversion[13] and further augments these with early adult personality skills represented by Agreeableness and Neuroticism to complete the Big Five system of personality.[14] Skills of Conscientiousness, Agreeableness and Neuroticism that are used in this paper strongly correlate with those used by Martin and Friedman (2000) based on the same data and are empirically shown to be closely related to their Big Five counterparts. Martin and Friedman (2000) also establish a strong link between a factor they call "Sociability" with Big Five Extraversion. Savelyev (2014) argues that the Extraversion factor as defined in this paper is close to Sociability and should therefore also be related to the Big Five Extraversion. Openness in this paper is theoretically related to Big Five Openness, but no empirical evidence is available to support this hypothesis yet. The constructed personality skills are similar in internal consistency reliability[15] to those used by Martin and Friedman (2000).

[12] As argued in Savelyev (2014), the average of ratings accounts for all available sources of information.

[13] See Savelyev (2014), Web Appendix, for a detailed analysis.

[14] Personality measures during childhood were not rich enough to account for Agreeableness and Neuroticism, so we use the early adulthood measures as the best available substitutes. EFA and CFA including Agreeableness and Neuroticism are available from authors upon request.

[15] Measured by Cronbach's Alpha (see Web Appendix A to Savelyev (2014)).

Table 4: Five Personality Factors and Their Measures

		Females				Males			
variable	year	mean	std. error	min	max	mean	std. error	min	max
Openness									
Knowledge	1922	10.091	.096	1	13	10.386	.077	3	13
Originality	1922	9.291	.102	1	13	9.563	.088	2	13
Intelligence	1922	10.677	.078	6	13	10.789	.069	1	13
Conscientiousness									
Prudence	1922	9.187	.104	1	13	8.554	.096	1	13
Conscientiousness	1922	10.253	.101	2	13	9.410	.101	1	13
Trustworthiness	1922	9.975	.106	2	13	9.750	.097	1	13
Extraversion									
Friendliness	1922	8.163	.111	1	13	7.582	.099	1	13
Leadership	1922	8.038	.098	2	13	7.481	.086	2	13
Popularity	1922	8.162	.101	2	13	7.488	.086	2	13
Agreeableness									
Easy to get along with	1940	7.147	.077	2	11	7.263	.068	2	11
Avoids arguments	1940	.694	.022	0	1	.525	.021	0	1
Critical	1940	.447	.025	0	1	.594	.022	0	1
Tactful	1940	.698	.023	0	1	.561	.022	0	1
Unfeeling	1940	.278	.022	0	1	.449	.022	0	1
Domineering	1940	.289	.023	0	1	.412	.022	0	1
Inflated opinion of self	1940	.239	.025	0	1	.360	.025	0	1
Neuroticism									
Miserable	1940	.249	.021	0	1	.215	.018	0	1
Touchy	1940	.355	.023	0	1	.373	.021	0	1
Periods of loneliness	1940	.273	.021	0	1	.250	.018	0	1
Lonely when with others	1940	.228	.020	0	1	.230	.018	0	1
Remorseful and regretful	1940	.228	.020	0	1	.194	.017	0	1
Lack self-confidence	1940	.470	.025	0	1	.242	.019	0	1
Worry about humiliating experiences	1940	.401	.023	0	1	.352	.020	0	1
Emotionally unstable	1940	.279	.022	0	1	.203	.017	0	1
Easily hurt	1940	.448	.024	0	1	.404	.022	0	1
Hard to be serene	1940	.101	.015	0	1	.094	.013	0	1
Moody	1940	6.215	.093	1	11	5.870	.080	1	10
Sensitive	1940	6.890	.076	3	11	6.614	.074	1	10
Estimation Sample			527				680		

Notes: Calculations are based on Terman data.

4 Methodology

4.1 Theoretical Framework

In line with Becker's (2007) approach, this project can be motivated by a generalization of a model suggested in Savelyev (2014) by explicitly modeling the key health-related economic choices: education, health-related consumption, and health investments.

Economic theories of demand for health usually do not investigate the role of personality (e.g., Grossman, 1972; Galama and Kippersluis, 2013). An exception is a recent paper-in-progress by Hai and Heckman (2014). Both this paper and the paper by Hai and Heckman take a similar approach of allowing a number of model parameters to depend on multi-dimensional personality (justified based on the literature), but models are different and have different focus. The model presented here contributes to the literature by summarizing multiple links between psychological skills and health-related economic decisions, namely health-related consumption, education, and health investments, thus linking psychology and economics. The paper by Hai and Heckman investigates the role of credit constraints and rational addiction among other aims. The main role of the model in this paper is a theoretical justification of causal links between skills and health-related outcomes that we estimate. We leave a more detailed analysis including generalization and calibration of this model to future research.

Consider a two-period model with time-separable utility.[16] A young adult makes decisions about college education, health investment, and consumption over the life cycle. There are only two periods of life that the young adult has ahead. She survives to the first period with certainty, but the probability S of surviving to the second period is below 1. In this simple model, IQ and personality are exogenous variables that can be influenced by parents, teachers, peers, specific interventions, or other factors of the environment prior to the education, health investment, and consumption decisions.[17]

[16] The model is easily generalizable to more periods, but a two-period model is sufficient to demonstrate the key features of the relevant economic decisions, which is the purpose of this paper.

[17] The model can be generalized in line with Becker and Mulligan (1997), who suggested that individuals may rationally invest in their imagination capital with the aim of reducing the discount on future utilities. In principle, individuals may likewise invest in their psychological skills, which may directly affect not only the discount rate but also the productivity of health investment, education cost, and wages. Via more

Let utility in both periods depend on health-related consumption C^H, health-neutral consumption C^N, and health stock H. In addition, due to the possibility of addiction to the health-related good, utility in the second period also depends on health-related consumption in the first period as in Becker (2007). The second period utility is discounted with discount factor B, which is assumed to depend on psychological skills $\mathbf{\Theta}$, and the survival function S, which is assumed to depend on health stock in the second period of life, H_2. The lifetime expected utility is given by

$$u_1(C_1^N, C_1^H, H_1) + B(\mathbf{\Theta}) \cdot S(H_2) \cdot u_2(C_2^N, C_2^H, H_2, C_1^H), \tag{1}$$

where $H_2 = f(I, D, \mathbf{\Theta}) + (1 - \delta(C_1^H))H_1$, in which f is the health production function, which depends on health investment I, education D, and personality skills $\mathbf{\Theta}$. The dependence on education is in line with the Grossman (1972) hypothesis, suggesting that education increases the efficiency of health production. The dependence on $\mathbf{\Theta}$ is in line with recent evidence from the literature, such as the productive role of Conscientiousness in treating diseases that require patients to follow complex treatment rules at home (Almlund et al., 2011). It is also well known that health-related consumption C_1^H affects how long people stay healthy and so determines health depreciation rate δ. For instance, smoking tobacco and heavy drinking of alcohol increase the deterioration of health, while good diet and physical exercise help to keep the health stock high.

Assume perfect capital and annuity markets. Let the cost of educational investment f depend on the chosen highest education level D, health in the first period H_1, and psychological skills $\mathbf{\Theta}$.[18] Let earnings depend on health and skills.[19] In the second pe-

effective health investment, skills may affect future health, longevity, and utility if alive.

[18]With education level D, both forgone earnings and the price of education become increasingly higher: compare costs of attending high school, college, and a professional school. Poor health is an obstacle for effective study. Among cognitive and personality skills, Cognition, Conscientiousness, and Openness are expected to lower the cost of education through lower tuition fees (e.g. higher probability of winning a scholarship), lower psychic costs, and greater time-efficiency in acquiring knowledge. Indeed, we can expect that Cognition helps to be effective at processing new information, Conscientiousness helps in staying organized and following rules, and Openness helps in creativity and by sustaining an interest in learning. We may also expect Extraversion to contribute to costs since studying implies forgone socializing. Neuroticism may increase psychic costs and reduce the efficiency in acquiring knowledge. It is unclear from theoretical considerations whether we should expect any effect of Agreeableness on college education.

[19]Clearly, greater health leads to greater productivity. Gensowski (2013) suggests that earnings are

riod let earnings also depend on education level. The individual maximizes the expected utility (1) subject to budget constraint

$$C_1^N + p^H C_1^H + g(D, H_1, \boldsymbol{\Theta}) + p^I I + \frac{S(H_2)}{1+r}(C_2^N + p^H C_2^H) = A + Y_1(H_1, \boldsymbol{\Theta}) + \frac{S(H_2)}{1+r} Y_2(D, H_2, \boldsymbol{\Theta}). \tag{2}$$

First order conditions (FOC) with respect to C_1^N, C_2^N and C_2^H are standard and are discussed in the Web Appendix. FOC with respect to health-related consumption, health investments, and education are more thought-provoking since they allow us to summarize theoretically expected inputs of psychological skills into marginal costs and benefits of the corresponding economic decisions. In line with the aim of the paper, we emphasise the FOC with respect to health-related consumption:

$$\underbrace{\frac{\partial u_1}{\partial C_1^H}}_{\text{cons. benefit}} \underbrace{-B(\boldsymbol{\Theta})S'(H_2)\delta'(C_1^H)H_1 u_2}_{\text{longevity benefit}} \underbrace{-B(\boldsymbol{\Theta})S(H_2)\frac{\partial u_2}{\partial H_2}\delta'(C_1^H)H_1}_{\text{health benefit}} + \underbrace{B(\boldsymbol{\Theta})S(H_2)\frac{\partial u_2}{\partial C_1^H}}_{\text{addiction benefit}} + \lambda \underbrace{\frac{S(H_2)}{1+r}\frac{\partial Y_2}{\partial H_2}\delta'(C_1^H)H_1}_{\text{health productivity}}$$
$$= \lambda\Big(\underbrace{p^H}_{\text{price}} + \underbrace{\frac{S'(H_2)\delta'(C_1^H)H_1}{1+r}(C_2^N + p^H C_2^H - Y_2(D, H_2, \boldsymbol{\Theta}))}_{\text{budget deficit}}\Big), \tag{3}$$

where $\lambda = \partial u_1 / \partial C_1^N$ is the shadow price of wealth. We can see that multiple marginal costs and marginal benefits contribute to equilibrium health-related consumption.

Generally, health-related consumption is a vector of multiple consumption types. Some of them, such as heavy drinking, smoking, or taking hard drugs have adverse effects on health depreciation, while others such as consuming healthy food or using gym services are beneficial. Consumption that is complementary with family stability (e.g., family trips) or with socialization (e.g., club or church memberships) are examples of both health-related consumption and beneficial addictions. For the ease of presentation

affected by Big Five skills and IQ, a result that we confirm in this paper

we treat health-related consumption as a composite consumption good that is a beneficial addiction in the sense of Becker's (2007) definition ($\partial u_2/\partial C_1^H > 0$), and has positive effects on health. Considering an adverse composite health-related consumption and generalizing to the case of multidimensional consumption is straightforward.

As it is clear from equation (3), marginal benefits are produced through enjoying health-related consumption ("consumption benefit"), higher probability to enjoy life in the second period ("longevity benefit"), greater expected utility in the second period due to better health ("health benefit"), greater expected utility in the second period due to beneficial addictions ("addiction benefit"), and greater earnings in the second period due to better health ("health productivity"). The marginal cost is the price of health-related consumption ("price") and additional spending (or revenues) due to a higher expected positive (or negative) budget deficit in the second period due to a higher probability of survival ("budget deficit"). We discuss other FOCs in the Web Appendix.

From this model we can see multiple links between psychological skills, education, health-related consumption, health investments, health, and longevity. In order to understand the multiple determinants of health-related consumption, it is productive to analyze major determinants of marginal costs and benefits shown in equation (3). Through $B(\mathbf{\Theta})$, skills affect longevity, health, and addiction marginal benefits. In addition, personality skills and education directly boost earnings, which increases u_1 through the wealth effect, thus contributing to the longevity benefit on all three margins. Moreover, complementarities may play a strong role (Becker, 2007). In this model, productive skills $\mathbf{\Theta}$ boost D and I, while D and I boost H_2, Y_2, S and u_2, thus further incentivizing healthier consumption. Similar complementarities are associated with an exogenous increase in D.

We therefore can conclude that mechanisms linking $\mathbf{\Theta}$, D, and C^H are numerous. While leaving detailed investigation of these mechanisms for future research, we concentrate in this paper on understanding a number of basic causal relationships implied by this model: the effect of education and psychological skills on health-related consumption and other health-related outcomes such as earnings.

4.2 Statistical Model

We estimate effects for education D, IQ (Θ^G), and five latent personality factors (Openness, Θ^O; Conscientiousness, Θ^C; Extraversion, Θ^E; Agreeableness, Θ^A; and Neuroticism, Θ^N) on each health-related outcome. Unlike in the theoretical model, vector $\boldsymbol{\Theta}$ in this section excludes IQ. The reason for this somewhat inconsistent notation is different approaches to estimating effects of skills and effects of IQ despite their similar role in theory. On one hand, IQ is a part of the set of psychological skills and so should be labeled as (Θ^G). On the other hand, given available measures in the Terman data, we cannot model IQ as a latent skill, but only as an observable variable.[20] Additionally, we want to separate the role of personality skills from the role of IQ, since IQ has become a traditional control, while personality is relatively new in economics literature.

We first identify all available health-related outcomes in the data, for which we expect a relationship with longevity based on theoretical considerations and literature results. We proceed with regression analysis, showing effects of education and latent psychological skills on various health-related outcomes, and adjust inference to account for multiple-hypothesis testing. To identify the effect of latent skills on an outcome, we need to estimate each outcome equation simultaneously with a system of measurement equations often referred to as a "measurement system."

4.2.1 Measurement System

The factors $i \in \mathcal{I} = \{O, C, E, A, N\}$ enter into a set of measurement equations, one for each psychological measure $M^j, j \in \mathcal{J} = \{1, \ldots, J\}$.[21] The factor model is thus defined

[20] There is only one reliable measure of IQ in the Terman data, which is not enough to identify a factor model involving the IQ as a latent factor.

[21] For example, "knowledge," "originality," "intelligence," "prudence," and so on (see Table 4).

by the following equations:

$$\begin{aligned} M_1 &= \alpha_1 + \boldsymbol{\beta}_1\boldsymbol{\Theta} + \gamma_1 A + \boldsymbol{\delta}_1\boldsymbol{X} + \eta_1 \\ &\vdots \\ M_j &= \alpha_j + \boldsymbol{\beta}_j\boldsymbol{\Theta} + \gamma_j A + \boldsymbol{\delta}_j\boldsymbol{X} + \eta_j \\ &\vdots \\ M_J &= \alpha_J + \boldsymbol{\beta}_J\boldsymbol{\Theta} + \gamma_J A + \boldsymbol{\delta}_J\boldsymbol{X} + \eta_J, \end{aligned} \tag{4}$$

where A refers to age in 1922 in order to control for differences in age at which personality was evaluated, and $\mathbf{X}$ represents background variables.[22] We use standard factor model identifying assumptions: $\Theta^i \perp\!\!\!\perp \eta_j$, for all $i \in \mathcal{I}$, and all $j \in \mathcal{J}$; $\eta_j \perp\!\!\!\perp \eta_{j'}$, for all j and j' such that $j \neq j'$. In addition, for all $i \in \mathcal{I}$ and $j \in \mathcal{J}$, $\mathbb{E}(\eta^j) = 0$ and $\mathbb{E}(\Theta^i) = 0$. Finally, we normalize the variance of each latent factor to one ($Var(\Theta^i) = 1, i \in \mathcal{I}$), and impose a set of exclusion restrictions for each measurement equation ($\beta^i_j = 0$ for a number of pairs (j, i)). The variance normalization is a standard technique that allows us to identify the model and interpret factor loadings as the effect of changing the factor by one standard deviation. The exclusion restrictions are grounded in our EFA and CFA analysis and are in line with the theory of the Big Five.[23] Since human skills including the Big Five correlate, we relax the assumption of orthogonality among skills: $\Theta^i \not\perp\!\!\!\perp \Theta^{i'}$ for $i \neq i'$. Identification of this standard factor model is discussed in detail in many papers including the classic Anderson and Rubin (1956), as well as more recent Heckman, Pinto, and Savelyev (2013), Heckman et al. (2014), and Williams (2011).

4.2.2 Linear Model for Health-Related Outcomes

We use a linear model to examine the effect of education and personality skills on health-related outcomes. Let H^k be the kth health-related outcome available in the Terman data,

[22] Controlling for A does add to the factor model, as discussed in Savelyev (2014), and linear approximation is adequate.

[23] See Web Appendix A to Savelyev (2014) for the EFA and CFA based on the Terman data uncovering exclusion restrictions for the childhood personality skills of Conscientiousness, Extraversion, and Openness. Supplementary EFA-CFA documentation specific to early adulthood personality of Agreeableness and Neuroticism is available from the authors upon request.

$k \in 1, ..., K$. We are interested in relationship

$$H^k = a^k D + \boldsymbol{b}^k \boldsymbol{\Theta} + c^k \Theta^G + \boldsymbol{d}^k \boldsymbol{X} + \epsilon^k, \tag{5}$$

where letters represent: D, the education indicator; $\boldsymbol{\Theta}$, the five personality skills; Θ^G, the IQ; $\boldsymbol{X}$, the control variables; and ϵ, the i.i.d. error term. For each k, we estimate equation (5) simultaneously with the measurement system (4), allowing us to identify the effect of latent factor $\boldsymbol{\Theta}$ whilst controlling for measurement error in measures, which is explicitly modelled in (4).

We further estimate two restricted models to understand the explanatory power of personality skills relative to that of other background controls and human capital measures traditionally employed by the literature. We estimate a model controlling only for personality skills:

$$H^k = \boldsymbol{b}_p^k \boldsymbol{\Theta}^u + \epsilon_p^k, \tag{6}$$

where $\boldsymbol{\Theta}^u$ are personality skills defined by a measurement system similar to (4), but without conditioning on $\boldsymbol{X}$. In this specification, neither D nor $\boldsymbol{X}$ enters any equation of the system, and so we explore the explanatory power of personality skills as we observe them.

We also estimate a model omitting the personality skills:

$$H^k = a_r^k \mathcal{D} + c_r^k \Theta^G + \boldsymbol{d}_r^k \boldsymbol{X} + \epsilon_r^k. \tag{7}$$

We compare the coefficient of determination (R^2) of models (5), (6), and (7) in Section 5 and argue that, at least for the Terman sample, personality alone explains about as much variation as Education, IQ, and background variables taken together.

4.2.3 Multiple-Hypothesis Testing Problem and the Stepdown Procedure

A major challenge in exploring treatment effects on multiple outcomes is accounting for false rejections due to the multiplicity of single hypotheses being tested (e.g., Westfall and Young, 1993). As the number of single hypotheses under consideration increases, the probability

that at least one of them is falsely rejected given that all of them are true, the (family-wise error rate) quickly increases. While this problem is well-recognized in genetic research, in which thousands of single hypotheses related to single genes are tested, in economics literature, this problem is largely neglected despite substantial probabilities of false rejection. Some researchers dealing with multiple-hypothesis testing provide arguments such as: "the number of single hypotheses in this study rejected at 5% level greatly exceeds 5%." While this is a valid argument to support the claim that a majority of rejected hypotheses are not rejected falsely, it provides no information on *which* specific inferences are to be trusted, creating an uncertainty about the validity of each single result of the study. Standard joint tests only provide limited help; by rejecting the hypothesis that all coefficients are jointly zero, all one knows is that at least one unknown effect is non-zero.

The problem with multiple-hypothesis testing in the literature is even worse than non-adjustment for FWE because of the related problem of selectively reporting statistically significant outcomes, which is also called "cherry-picking." The reader is not aware of the large number of unreported individual hypotheses, which authors consider and later selectively drop from consideration after finding out that they yield no statistically significant results.

In order to introduce some formal notation and definitions, consider a family of single tests. Let the chance of false rejection for every single test be $\alpha = P(H_1|H_0)$. The problem is that the chance of *at least one* false rejection in the family is substantially higher than α. Define family-wise error rate, FWE$=$ $P($Reject at least one $H_i|$ all H_i are true$)$. For instance, let α be 0.05 for each single test. Then, for a family of four independent tests, the FWE(4)=$1-(1-0.05)^4$=0.19. Likewise, FWE(7)=0.30; FWE(10)=0.40; FWE(60)=0.95; FWE(90)=0.99. In words, in a family of four single hypotheses at least one false rejection is not unlikely. In a family of 7–10 hypotheses such false rejection is likely. Finally, if there are 60 or more hypotheses, at least one false rejection happens almost for sure.

In line with the above analysis, given the multiplicity of outcomes explored in this paper, it would be overoptimistic to accept calculated single-hypothesis p-values at their face value. We need to account for the fact that testing a group of several hypotheses makes false rejection of at least one of them more likely. The question is how to

select the group of hypotheses. In principle, we could consider a group that contains all numerous single hypotheses that were ever tested in this paper, but this approach is overly-conservative, leading to the opposite problem: we risk adjusting inference in such a conservative way that many truly-rejected hypotheses will be labeled as falsely rejected.

Fortunately, we can make the procedure less conservative by using a-priori information and by asking a more precise question. Following Heckman, Moon, Pinto, Savelyev, and Yavitz (2010), we account for multiple-hypothesis testing within each group of single hypotheses that are clustered a-priori by type. For example, we wish to test whether education has a statistically significant negative effect on heavy alcohol consumption at each observed stage of the life cycle. Based on evidence from the literature we have prior knowledge that education should affect heavy drinking, but we are not sure whether this effect would persist over the entire life cycle. It is possible that random variation in sample size from follow-up to follow-up combined with multiple-hypothesis testing may lead to some false rejections. We therefore group all single hypotheses on the effect of education on heavy drinking at different years and perform the stepdown adjustment. For the single tests that survive the stepdown adjustment (so that the effect is still statistically significant after stepdown correction), we are sure that the result is not a false rejection due to the multiplicity of heavy-drinking measurements over the life cycle.

Since we are less sure about a-priori knowledge concerning certain other outcomes such as number of memberships, we supplement stepdown adjustment in groups by type of outcomes with a more conservative stepdown adjustment on the full set of aggregated health-related outcomes of various types. If a rejected individual test survives this adjustment, we conclude that the rejection did not occur falsely due to testing multiple single hypotheses on a set of aggregated health-related outcomes.

Let there be K individual hypotheses in a family. Then, adaptation of the general stepdown algorithm by Romano and Wolf (2005) to particular needs of this paper leads to the following procedure:

1. For each individual hypothesis in the family, obtain the true t-statistic and B bootstrap t-statistics. (Use absolute values of t-statistics since all tests are double-sided.)

2. Find the maximal t-statistic among K true t-statistics. Do the same for each pseudo-sample to get a bootstrap distribution of maximal t-statistics.

3. Use the distribution of maximal bootstrap t-statistics to test the hypothesis associated with the maximal true t-statistic. The p-value of this test is the stepdown-adjusted individual hypothesis p-value for that hypothesis.

3. If the test cannot be rejected at the chosen significance level, then stop the procedure and conclude that none of the remaining tests can be rejected either.

4. If the test can be rejected, then record the result and exclude the rejected hypothesis from the family for further steps. If only one hypothesis is left after the exclusion then test the hypothesis individually and stop the procedure. If multiple hypotheses are left then repeat the procedure starting from (2).

We discuss stepdown-adjusted results in Section 5.

Advantages of the stepdown procedure are the following. First, it strongly controls for the FWE. Strong control holds regardless of which subset of hypotheses happen to be true (any partial null), while weak control holds only if all hypotheses are true (the complete null) (Westfall and Young, 1993). Second, it tests for adjusted statistical significance of every single hypothesis, unlike standard joint tests. Finally, it is a more powerful method than the computationally simpler Bonferroni and Holm-Bonferroni methods. Gains in power come from accounting for statistical dependencies among individual test statistics using resampling (Romano and Wolf, 2005).

A big computational advantage of the stepdown procedure is the lack of a need to resample t-statistics again for the subsequent stages of stepdown. Despite this computational advantage, calculations for this paper still took us weeks because of the use of computationally-intensive five-factor models, multiplicity of outcomes modelled separately for each gender, and the need to calculate bootstrap distributions of all relevant statistics.

4.2.4 Assumptions for Claims of Causality

Similar to Heckman et al. (2006), we assume that conditional on detailed and theoretically relevant childhood and parental characteristics, the dependence across education choices and health-related outcomes comes from cognitive and personality skills. The richness of the Terman data and the comprehensive nature of the personality controls (as argued in the Big Five theory) give additional credibility to this assumption. The conditional independence assumption is the key assumption of the matching literature with the difference that we additionally control for latent ability via factor analysis. We acknowledge that it is impossible to control for all causes, but we hope that our relevant and elaborate set of controls makes the bias due to endogeneity of education negligible.

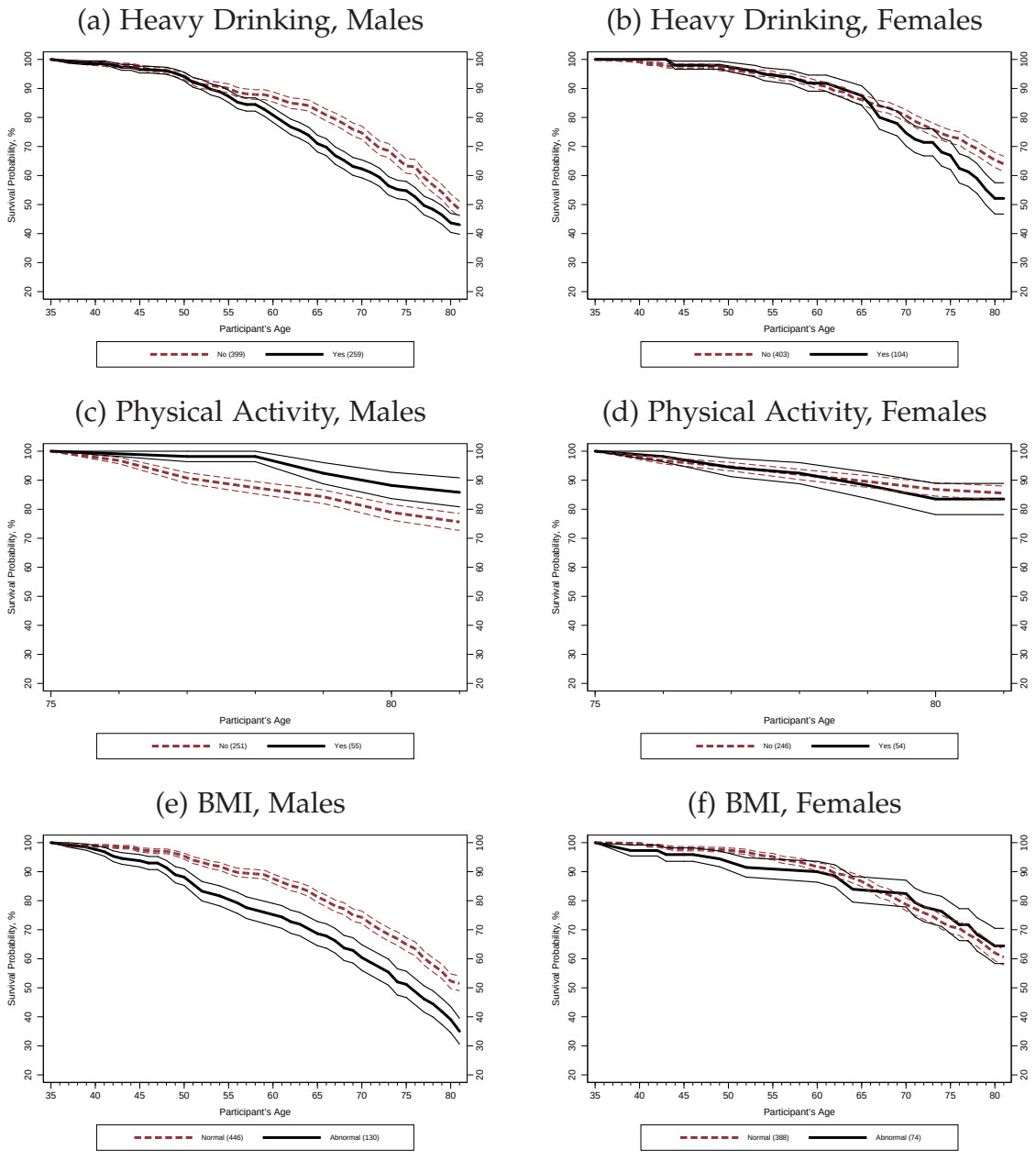
5 Empirical Results

5.1 Longevity and Health-Related Outcomes

We first perform a life-table descriptive analysis on a range of health-related adult outcomes. This exercise confirms that most of the adult outcomes thought to be important for predicting longevity are associated with longevity in the Terman sample, at least for males (see Figure 1 for health behaviors, Figure 2 for health measures, Figure 3 for lifestyle choices, and Figure 4 for earnings).[24] In particular, measures of general health, mental health, later life earnings, abnormal BMI, and heavy drinking exhibit strong correlations with survival probabilities for males, while the other health-related outcomes show smaller but still distinct differences in survival probabilities (divorce, group memberships). For females, differences in health behaviors, health measures, and lifestyle choices generally translate to a much smaller, if any, gradient in longevity as compared to males.

[24] We summarize the multi-period observations of heavy drinking, mental difficulty, and general health into binary indicators which equal 1 when negative outcomes are observed in any point in time of the life cycle, and 0 otherwise. We also present just a subset of marriage statuses and earnings measures in the interest of brevity.

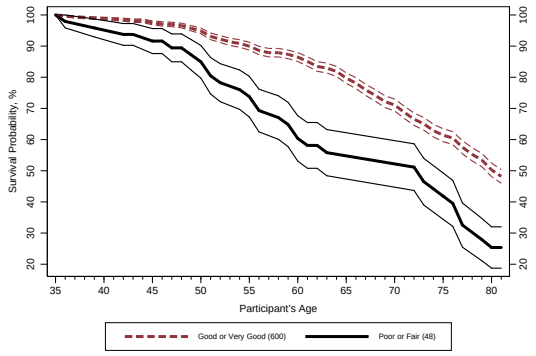
Figure 1: Survival by Health Behaviors and Their Proxies



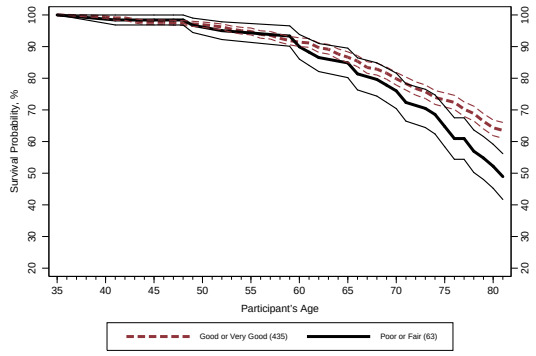
Notes: Heavy drinking in this graph is one if the subject has ever reported drinking heavily over the period of 1940–1960 and zero otherwise. Physical activity indicates whether or not the subject engaged in physical activity frequently in 1982. BMI indicates whether or not the subject had abnormal BMI in 1940, where abnormal means underweight or overweight. Overweight refers to subjects who had a BMI above 25. Underweight subjects had a BMI below 18.5, but the role of underweight BMI is negligible. Survival graphs are based on life-table calculations; standard errors above and below the estimate are represented by the thinner lines.

Figure 2: Survival by Health Measures

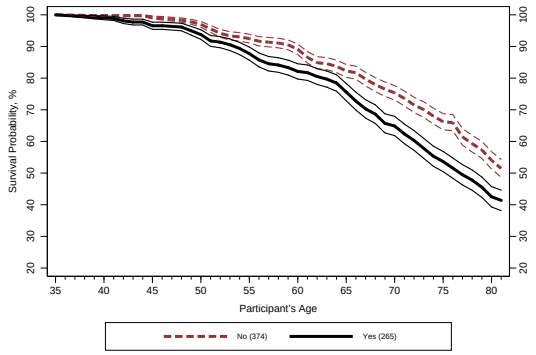
(a) General Health, Males



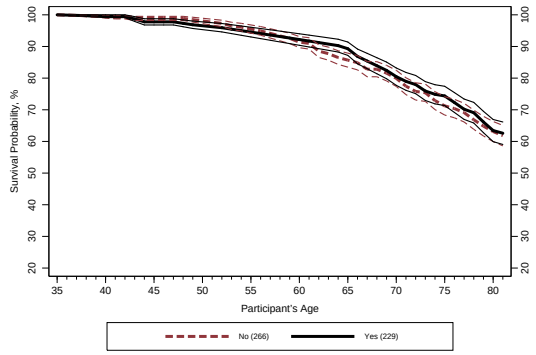
(b) General Health, Females



(c) Mental Difficulty, Males

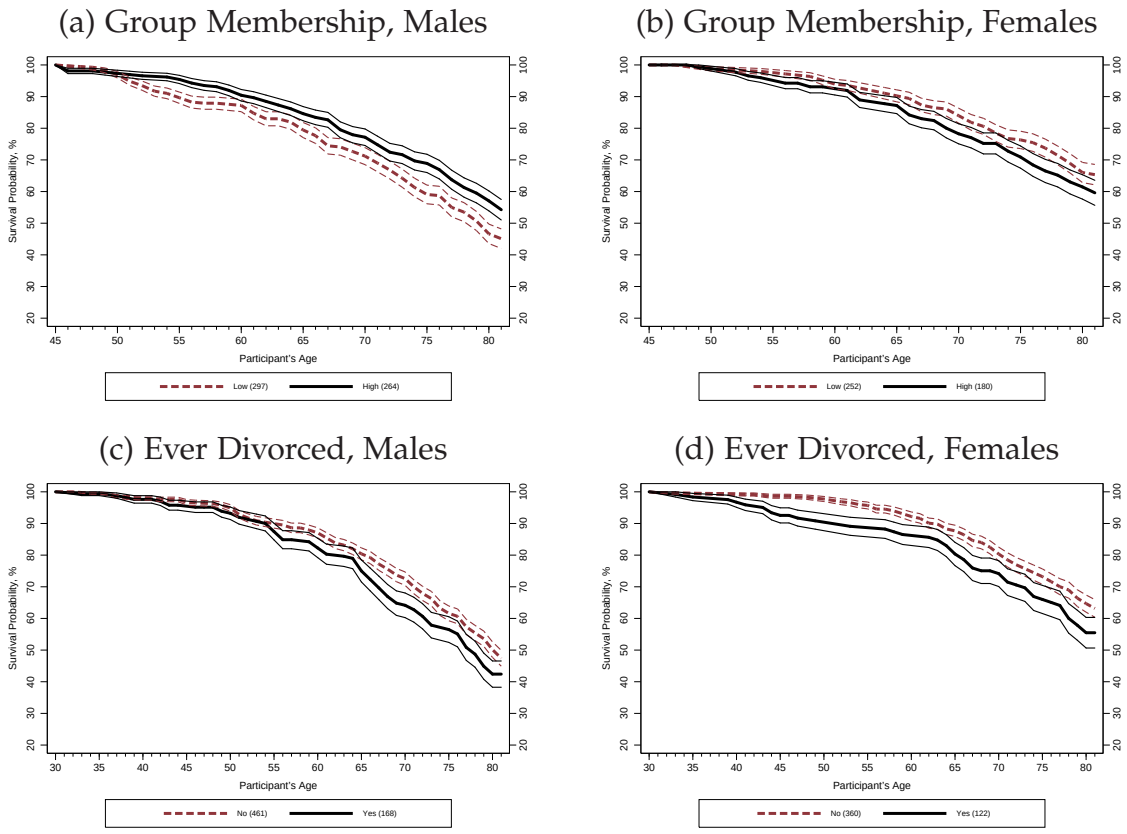


(d) Mental Difficulty, Females



Notes: General health is an index constructed from various self-reported health measures including "energy level", "vitality", and "physical health". It indicates whether the subject experienced poor or fair health over the years 1940–1960. Mental difficulty indicates whether or not the subject experienced any mental difficulty over the years 1950–1960. Survival graphs are based on life-table calculations; standard errors above and below the estimate are represented by the thinner lines.

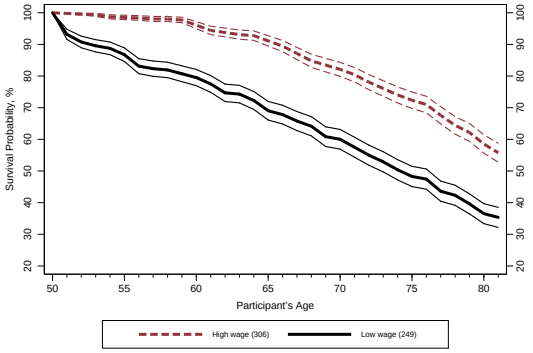
Figure 3: Survival by Lifestyle Choices - Social and Family



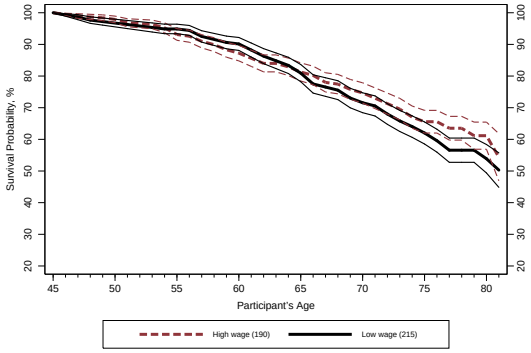
Notes: Lifestyle choices refer to group membership in 1950 and marriage status. 'High' membership refers to subjects having a greater number of organization memberships than the median. 'Low' membership refers to subjects at or below the median number of organization memberships. "Ever divorced" indicates whether the subject was divorced at least once. Survival graphs are based on life-table calculations; standard errors above and below are represented by the thinner lines.

Figure 4: Survival by Age 50 Earnings

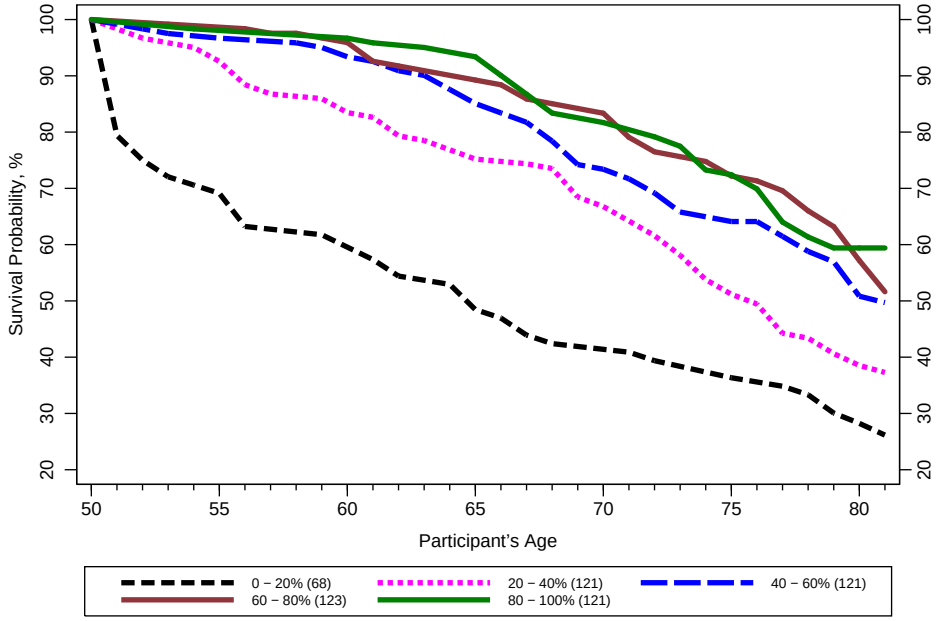
(a) Above Median, Males



(b) Above Median, Females



(c) Quintiles, Males



Notes: "High wage" refers to earnings above the median, "low wage" refers to earnings at or below the median. For females, the median wage is zero. Survival graphs are based on life-table calculations; standard errors above and below each estimate are represented by the thinner lines.

5.2 Personality, Education, and Health-Related Outcomes

We present the summary of our main results for health-related outcomes in Tables 5 and 6. Each cell in the table shows only the regression coefficient and asterisks denoting the stepdown-adjusted statistical significance level. (More detailed results presenting both adjusted and unadjusted p-values are in Tables 10–16 of the Appendix.) The p-values are adjusted for multiple hypothesis testing within blocks of outcomes of the same type, for example, to all available heavy alcohol drinking-related outcomes across the life cycle, or to all available marriage-related outcomes.[25] As argued above, for a number of skills and outcomes we have prior evidence that we should expect a treatment effect of this particular type, such as the effect of education on heavy drinking, but we are less sure about the behavior of this effect over time, which motivates stepdown adjustments within blocks but not across blocks.[26] Coefficients from the same block are marked by bold frames in Tables 5 and 6. We report both statistically significant coefficients using standard thresholds ($p < 0.01$, 0.05, and 0.10), and borderline statistically significant coefficients ($p < 0.15$). Coefficients with p-values above 0.15 are not shown since we can hardly distinguish them from zero, but they are available in Tables 10–16 of the Appendix. The results are color-coded so that green (or light grey in print) refers to effects that are considered in the literature to be beneficial for longevity (such as a decrease in heavy drinking or an increase in physical activity), and red (or dark grey in print) refers to adverse effects.

One quick way to analyze these summary tables is to look at paterns of green and red only, and notice that some skills show multiple effects that are beneficial for health, while some others are of the opposite sort. We discuss these patterns below.

5.2.1 Conscientiousness, Neuroticism, IQ, and Education

Our results show that, for males, Conscientiousness and education act on health-related outcomes beneficially (look at patterns of green (light grey) associated with them), while

[25] Physical exercise, BMI, and smoking are exceptions as we know them at only one specific age. We do not adjust them in Tables 5 and 6, but we do adjust them in a more conservative test described below.

[26] Our confidence in this prior evidence differs from outcome to outcome. Below we provide a more conservative test that essentially works across blocks.

Table 5: Summary of Effects on Health-Related Outcomes, Males

	C	O	E	A	N	IQ	Education
A. Health behaviors and their proxies							
1940–1960 Ever Drank Heavily	-.055 **		.061 **				-.109 **
1940 Heavy Drinking	-.046 *		.044			.057 **	-.086
1950 Heavy Drinking			.040 **		.039 *		-.090 **
1960 Heavy Drinking	-.072 **	.056	.044 *				-.077
1940 Overweight				-.034		-.023	
1982 Physical Activity, Freq.		-.044 *			-.066 **		.108 *
1991 Ever Smoked	-.107 **						
1940–1960 Any Organization							.084 ***
1940 Number of Organizations						-.175 *	.245
1950 Number of Organizations				.258 *			1.172 ***
1960 Number of Organizations						.327 **	1.501 ***
Never Married	.023				.024		
Married Once and Still Married	.056 *						.120 **
Ended up Divorced	-.023 *	.050 ***			.024		
Ever Divorced	-.055 *						-.137 **
Divorced at least Twice	-.044 **	.031 *			.025		
B. Earnings							
Lifetime earnings, 3%			79.908 **	-94.713 **		44.431	209.191 ***
Earnings at age 40				-6.556 ***		3.280	14.585 ***
Earnings at age 50			4.122	-6.787 **	-6.553 **	4.758 *	19.788 ***
Earnings at age 60			5.814 *		-7.466 **		30.530 ***
C. Mental Health (MH)							
Ever Poor/Fair MH	-.071 ***	.085 ***	-.051 *		.134 ***		
1940 Mental Difficulty	-.078 ***	.086 ***	-.077 ***		.120 ***		
1950 Mental Difficulty	-.040 *				.111 ***		
1960 Mental Difficulty	-.080 ***	.091 ***	-.101 ***		.120 ***		
D. General Health (GH)							
Never Poor/Fair GH		-.032 *			-.021		
1940 General Health					-.279 ***		
1950 General Health	.135 **	-.152 **	.096		-.242 ***		
1960 General Health					-.211 ***		

Notes: Calculations are based on the Terman data. Letters denote: C, Conscientiousness; O, Openness; E, Extraversion; A, Agreeableness; N, Neuroticism. Coefficients are reported with accompanying statistical significance represented by stars, where $***, **, *$ indicates $p < 0.01, 0.05, 0.10$ respectively. A coefficient with no star refers to $p < 0.15$, while a blank cell refers to a coefficient with p-value above 0.15. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005). Coefficients shaded green (light grey in print) and red (dark grey in print) denote beneficial and adverse implications for health. See Tables 10–16 for a full set of results used for the summary.

Table 6: Summary of Effects on Health-Related Outcomes, Females

	C	O	E	A	N	IQ	Education
A. Health behaviors and their proxies							
1940–1960 Ever Drank Heavily		-.073 **	.054 *				
1940 Heavy Drinking				-.041 *			
1950 Heavy Drinking							
1960 Heavy Drinking		-.060 *	.049				
1940 Overweight					-.037 *		-.074 *
1982 Physical Activity, Freq.							
1991 Ever Smoked							
1940–1960 Any Organization							.066 **
1940 Number of Organizations							.789 ***
1950 Number of Organizations							.877 ***
1960 Number of Organizations						-.352 **	1.213 ***
Never Married							.074 ***
Married Once and Still Married							.129 *
Ended up Divorced							
Ever Divorced							-.111 **
Divorced at least Twice							-.054 *
B. Earnings							
Lifetime earnings, 3%							
Earnings at age 40							3.946 *
Earnings at age 50							
Earnings at age 60					-4.650		
C. Mental Health (MH)							
Ever Poor/Fair MH					.152 ***		
1940 Mental Difficulty					.137 ***		
1950 Mental Difficulty					.134 ***		
1960 Mental Difficulty					.123 ***		
D. General Health (GH)							
Never Poor/Fair GH					-.044 ***		.116 ***
1940 General Health				-.133 *	-.318 ***		.283 **
1950 General Health				-.094	-.267 ***		.172
1960 General Health					-.241 ***		

Notes: Calculations are based on the Terman data. Letters denote: C, Conscientiousness; O, Openness; E, Extraversion; A, Agreeableness; N, Neuroticism. Coefficients are reported with accompanying statistical significance represented by stars, where $***, **, *$ indicates $p < 0.01, 0.05, 0.10$ respectively. A coefficient with no star refers to $p < 0.15$, while a blank cell refers to a coefficient with p-value above 0.15. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005). Coefficients shaded green (light grey in print) and red (dark grey in print) denote beneficial and adverse implications for health. See Tables 10–16 for a full set of results used for the summary.

Neuroticism is disadvantageous (see Table 5 for patterns of red or dark gray). Both Conscientiousness and education reduce heavy drinking over the life cycle and protect against divorce. Education also increases earnings over the life cycle. Childhood Conscientiousness also reduces mental difficulty and increases general health at age 50.[27] Education enhances physical activity, while Openness and Neuroticism decrease it. Neuroticism also has large and statistically significant negative effects on general and mental health, as well as negative effects on earnings at ages 50 and 60.

IQ of males increases heavy drinking in early adulthood (year 1940), has a borderline statistically significant negative effect on being overweight, has mixed effects on social participation, and a borderline statistically significant positive effect on earnings.[28]

For females, we reject only a few hypotheses, which, coupled with little correlation between behavior and longevity of females discussed above, is in line with no effect of education and skills on longevity for females found in Savelyev (2014) based on the same data.[29] We can see that education consistently encourages group membership among women over the life cycle, improves their general health at least in young adulthood, has an effect on earnings at age 40, and decreases the likelihood of divorce. Both education and Neuroticism reduce the incidence of an overweight BMI. In this case, Neuroticism, an otherwise unproductive trait, has an effect that is potentially good for health. Interestingly, we see no such effect for males. Understanding the psychological and biological mechanisms behind this unexpected effect as well as the gender difference is beyond the scope of this paper. To provide one possible explanation, a neurotic woman may worry more about her appearance, leading to a reduced likelihood of being overweight.

5.2.2 Openness, Agreeableness, and Extraversion

For males, the effects of Extraversion are mixed with regard to potential health-enhancing effects. Extraversion encourages heavy drinking over the life cycle (probably through in-

[27] Gensowski (2013) finds the effect of adult Conscientiousness on earnings. While we confirm this result (not documented in this paper), we find no such effect of childhood Conscientiousness.

[28] The effect on earnings is in line with Gensowski (2013).

[29] Effects on potential mediators that are weak and not numerous may lead to so small a causal effect on longevity that we cannot distinguish it from zero.

creased participation in social gatherings) but is at the same time beneficial for earnings and mental health (probably through better communication skills, greater networks, and emotional support from others).

Agreeableness does not have statistically significant effects on most health-related outcomes. The exceptions include earnings, where Agreeableness has a large negative effect for males, and number of organizations in 1950, with a small positive effect. A possible reason of the negative effect on earnings is that Agreeable people are less likely to be promoted due to their unwillingness to criticize others and make tough decisions, which are valuable skills in management. Further, Agreeable persons may be less willing to change employers and locations for career development (which usually leads to a wage boost), due to placing a greater importance on the interests of other members of their family, who may prefer to stay in the current location, and on the interests of their current employer.

Openness to experiences increases mental difficulty and divorce rates, and has some negative effect on general health. Thus, even though Openness is productive for creativity, it comes at some health cost. Note that since IQ correlates with Openness, Openness results should be interpreted with caution as this sample is selected on high IQ.

Overall, we find no persistent, strong, and precisely determined effects of Openness, Agreeableness, and Extraversion on health-related outcomes for females. Similar to males, there is some positive effect of Extraversion on heavy drinking, but there is also some negative effect of Openness on the same outcome. Agreeableness has some negative effect on heavy drinking and some negative effect on general health in young adulthood, with both effects being borderline statistically significant.

5.3 Mid-life and Lifetime Outcomes

Step-down adjustment within groups by age was justified by an a-priori expectation that for many groups we expect the effect for at least some ages. However, our confidence in this a-priori knowledge differs from outcome to outcome; so, in addition to our main approach described above, we provide a more conservative approach, in which we adjust

single hypotheses in a family of tests involving outcomes of different type, thus allowing for the chance that some of them could be rejected only due to multiplicity of different health-related outcomes. We use two alternative groups of outcomes: (1) outcomes aggregated over the life cycle whenever information for such aggregation is available; (2) outcomes in year 1960, around age 50, or as close to 1960 as possible given available outcomes (see Tables 7 and 8).[30] We can see that most key results discussed above survive even this conservative adjustment.

Interpretation of results of this more conservative approach should be done with caution, taking into account all available prior information. For instance, while in panel (B) of Table 7 the effect of Conscientiousness on heavy drinking is statistically significant at 5% level, in panel (A) of the same figure p is below 0.15, a borderline statistically significant result. We argue that despite the result in panel (A), false rejection for this specific outcome is unlikely. First, we have multiple sources of strong prior evidence showing the negative effect (see Bogg and Roberts (2004) for a survey). Second, there are several plausible causal mechanisms behind the effect.[31] Third, we use a conservative two-sided p-value, despite the strong prior evidence of a negative effect leading to unadjusted p-values that are twice as large as one-sided ones. In other specific cases, such as the loss of a statistically significant negative effect of Openness on physical activity, it is better to adopt the conservative approach and not reject the hypothesis given that prior empirical evidence based on different data in favour of rejecting the hypothesis is absent. Mechanisms behind such possible effects are unclear, and there is no reason to use a less conservative one-sided test in this case. As before, we see few statistically significant results for females (see Figure 8).

[30] Using the latest outcome available whenever possible, such as 1960 chosen from the period 1940–60, should increase the power of estimates on health stock such as general and mental health due to longer period of effect accumulation.

[31] Mechanisms are easy to understand by revisiting the theoretical model, especially the FOC with respect to health-related consumption such as heavy drinking (see Equation [3]). Conscientiousness is expected to increase the discount factor B, thus affecting longevity and morbidity marginal benefits, and the marginal cost of addiction. Additionally, as discussed above, Conscientiousness makes heath investments more efficient, thus contributing to health production and income effect. Greater health in the second part of life contributes to addiction, health, and health productivity marginal benefits, as well as to the marginal cost related to budget deficit. Income effect due to higher wage and more efficient health production contributes to utility in the second period, thus increasing the longevity marginal benefit. As discussed, complementarities additionally strengthen the causal effect.

Table 7: Lifetime and 1960 Outcomes and Proxies, Males

	C	O	E	A	N	IQ	Education
A. Life-cycle Outcomes and Proxies							
1940–1960 Ever Drank Heavily	-.055						-.109 *
1940 Overweight							
1982 Physical Activity, Freq.					-.066 *		
1991 Ever Smoked	-.107 **						
1940–1960 Any Organization							
Ever Divorced	-.055						-.137 ***
Lifetime earnings, 3%			79.908 *	-94.713 *			209.191 ***
Ever Poor/Fair MH	-.071 *	.085 ***			.134 ***		
Never Poor/Fair GH							
B. 1960 Outcomes and Proxies							
Drank Heavily	-.072 **						
1940 Overweight							
1982 Physical Activity, Freq.					-.066 *		
1991 Ever Smoked	-.107 *						
# of Organizations						.327 *	1.501 ***
Ever Divorced	-.055 *						-.137 **
Age 50 earnings				-6.787 *	-6.553 *		19.788 ***
Mental Difficulty	-.080 **	.091 ***	-.101 ***		.120 ***		
General Health					-.211 ***		

Notes: Calculations are based on the Terman data. Letters denote: C, Conscientiousness; O, Openness; E, Extraversion; A, Agreeableness; N, Neuroticism. Coefficients are reported with accompanying statistical significance represented by stars, where $***, **, *$ indicates $p < 0.01, 0.05, 0.10$ respectively. A coefficient with no star refers to $p < 0.15$, while a blank cell refers to a coefficient with p-value above 0.15. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005). Coefficients shaded green (light grey in print) and red (dark grey in print) denote beneficial and adverse implications for health.

Table 8: Lifetime and 1960 Outcomes and Proxies, Females

	C	O	E	A	N	IQ	Education
A. Life-cycle Outcomes and Proxies							
1940–1960 Ever Drank Heavily							
1940 Overweight							
1982 Physical Activity, Freq.							
1991 Ever Smoked							
1940–1960 Any Organization							
Ever Divorced							-.111
Lifetime earnings, 3%							
Ever Poor/Fair MH					.152 ***		
Never Poor/Fair GH					-.044		.116 **
B. 1960 Outcomes and Proxies							
Drank Heavily							
1940 Overweight							
1982 Physical Activity, Freq.							
1991 Ever Smoked							
# of Organizations						-.352 *	1.213 ***
Ever Divorced							-.111
Age 50 earnings							
Mental Health					.123 ***		
General Health					-.241 ***		

Notes: Calculations are based on the Terman data. Letters denote: C, Conscientiousness; O, Openness; E, Extraversion; A, Agreeableness; N, Neuroticism. Coefficients are reported with accompanying statistical significance represented by stars, where $***, **, *$ indicates $p < 0.01, 0.05, 0.10$ respectively. A coefficient with no star refers to $p < 0.15$, while a blank cell refers to a coefficient with p-value above 0.15. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005). Coefficients shaded green (light grey in print) and red (dark grey in print) denote beneficial and adverse implications for health.

5.4 Personality vs. Traditional Controls

The importance of personality skills is comparable to the combined role of education, IQ, and detailed background controls for many of our health-related outcomes. Figure 5 presents the R^2 statistic for three models: the full model, the model with only personality skills, and the model omitting personality skills.[32] The results suggest that omitting personality skills leads to a dramatic reduction in R^2 for all health-related outcomes, particularly mental health. This strong result establishes personality skills as an important aspect of human capital that should receive more attention from economists. We acknowledge that in a more heterogeneous sample traditional controls are expected to explain more variance, especially due to variation in IQ.

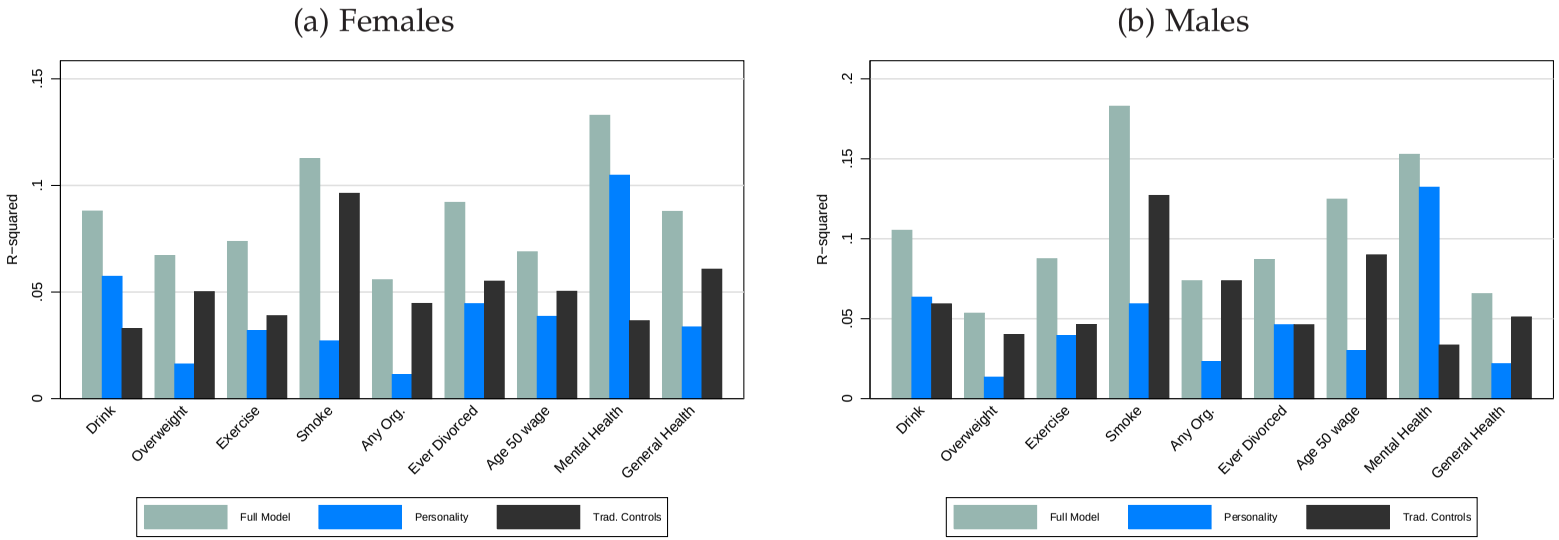
6 Discussion

We contribute to the existing research by jointly estimating the effects of education and five personality skills conditional on education, as well as by accounting for multiple-hypothesis testing, which is uncommon in the literature. This is an improvement over a treatment of personality as a low-dimensional object often reduced to just a single-dimensional variable, and it adds interpretability and refinement to the role of personality as a determinant of health-related outcomes. The unique Terman data allows us to establish links between childhood measures of psychological skills and multiple health-related outcomes over the life cycle. Our paper thus provides new evidence for education and personality skills as major determinants of health-related outcomes, and also raises awareness for a wider set of health-related outcomes than commonly considered.

Education and Health Our paper uses a methodology that serves as an alternative to natural experiments in order to provide evidence that education has causal effects on heavy drinking, earnings, divorce, and physical activity. Results are in line with a number of effects documented in the literature: effects on reducing heavy drinking

[32]The sum of the R^2 statistic across model 2 and model 3 may fail to coincide exactly with the R^2 statistic of model 1 due to correlations between personality skills and traditional economic controls.

Figure 5: Coefficient of Determination (R^2) Comparison



Notes: Calculations are based on the Terman data. For each health-related outcome, R^2 is reported for the full model, the model based on only personality skills, and the model omitting personality skills. Due to correlations between latent personality skills and observable regressors, the R^2 for the full model can be somewhat smaller than the sum of R^2 for the partial models.

(Conti and Hansman, 2013; Crum et al., 1993; Droomers et al., 1999), increasing earnings (Card, 1999), lowering divorce rates (Stevenson and Wolfers, 2007), and encouraging physical activity (Conti and Hansman, 2013; Conti et al., 2010b). Results are also in line with the theory presented in Section 4.1, which suggests multiple channels through which greater education leads to healthier consumption and superior health. Indeed, if we raise college education exogenously, we expect to see an increase in $S(H_2)$[33] and Y_2 (see Table 16), therefore increasing the marginal benefits of health-related consumption through longevity, morbidity, addiction, and health productivity.[34] It is beyond the scope of this paper to break down precisely the relative importance of the various channels.

Personality and Health Many of the effects of Conscientiousness and Neuroticism on health-related outcomes that we estimate are large and statistically significant. Estimated coefficients reflect a substantial percentage of sample means for most outcomes, and a few of the less-precisely estimated coefficients are still sizeable.

Our results therefore confirm the positive effects of Conscientiousness on health, while for Neuroticism we add to a growing body of evidence that it is a major determinant of health-related outcomes (Lahey, 2009). Our results are also consistent with the literature with regard to the effect of personality skills (see Bogg and Roberts (2004); Droomers et al. (1999); Friedman (2000); Friedman et al. (1993); Lahey (2009)). For example, the negative effects of Agreeableness on earnings and the positive effects of Extraversion on drinking alcohol are widely recognized patterns (Cookson, 1994; Flory et al., 2002; Heineck and Anger, 2010; Judge and Livingston, 2011; Mueller and Plug, 2006). We confirm these patterns conditional on a substantial set of controls, IQ, and other personality skills, which makes causal interpretation of the estimated effects more plausible under conditional independence assumption.

Our results also have a number of implications for our theoretical model. First, the estimates with regard to mental and general health can be viewed as estimates of $\frac{\partial H_2}{\partial \Theta}$ from our economic model. We find that Conscientiousness and Extraversion affect health stock positively while Openness and Neuroticism have negative effects. Sec-

[33] See, e.g., Buckles et al. (2013) and Savelyev (2014).

[34] The sign of the effect on the marginal budget deficit is not clear.

ond, similar to the case of education discussed above, our findings suggest that skills change the marginal effects of health-related consumption through at least five margins ("longevity," "health," "addiction," "health productivity," and "budget deficit"). The effects of skills on earnings that we find speak directly to the "budget deficit" channel, and indirectly to the longevity benefit through the wealth effect on utility, in addition to complementarity channels such as complementarity between wealth and education.

Data Limitations and External Validity The results in this paper are based on a historical sample with exceptional IQ. We have access to early measures of psychological skills and high-quality life-cycle data at the expense of dealing with both an unusual and deceased cohort. We follow Savelyev (2014) in our approach to external validity and data limitations.

Effects of education may differ with the level of IQ, and so it is useful to know such effects in the limiting case of very high IQ. Another benefit of selection on high IQ is that it reduces the potential of IQ to confound the effects of education on health: all subjects in the sample were smart enough to finish college.

We do not claim applicability of results to the general population, but results may be applicable to people who are smart, although not necessarily as smart as the Terman subjects. Indeed, if health choices are not specific to extraordinarily high IQs, we can expect similar results to hold for less-exceptional populations.

Application to more recent cohorts presents another challenge. Social norms toward many of these health behaviors have changed dramatically over time, which may affect the magnitude of the effects. Another big factor is the amount of available information on how these health behaviors affect health and longevity discussed in detail in Savelyev (2014). We can now do more good to our health through the informed choices we make. To the extent that Conscientiousness and education act as skills that motivate us to inform ourselves and adopt new technologies, this could potentially increase their importance in determining health-related outcomes compared to results based on the Terman sample.

Implications for Mental Health To the extent that we can generalize our results to a modern population and to a wider population, our findings may be of interest to policy makers. Our findings with regard to mental health are of interest given the growing awareness of mental well-being as a key component of health. Neuroticism in both genders plays the largest role in determining later life mental health, while for males Conscientiousness, Openness, and Extraversion are also important. In this regard, our results suggest that childhood personality skills can act as a direct and early life mechanism for improving mental health and, through mental health, all other essential outcomes mediated by mental health.

Stepdown Adjustment In all of our analyses, we employ the stepdown procedure suggested by Romano and Wolf (2005). This adjustment decreases the number of rejected single hypotheses. We show an example of the difference that the stepdown procedure makes in Table 9, where items shaded in blue (light grey in print) refer to results that would have been statistically significant without the stepdown adjustment over the set of life-cycle outcomes but do not survive the adjustment. The survival rate of unadjusted statistically significant estimates after the stepdown adjustment was about 40% and 60% for males and females respectively. While most of the effects that did not survive the adjustment may look like sensible results, we stress the importance of conservative inference based on stepdown-adjusted p-values when analyzing multiple single hypotheses. Single hypotheses that did not survive the adjustment are suspects for false rejection and should be interpreted with caution.

Table 9: Lifetime Outcomes and Proxies, Stepdown Comparison

	C	O	E	A	N	IQ	Edu
A. Males							
1940–1960 Ever Drank Heavily	-.055 **		.061 ***	-.045 *			-.109 **
1940 Overweight							
1982 Physical Activity, Freq.		-.044 *			-.066 **		.108 *
1991 Ever Smoked	-.107 **						
1940–1960 Any Organization							.084 ***
Ever Divorced	-.055 **			-.045 *			-.137 ***
Lifetime earnings, 3%		-75.290 *	79.908 **	-94.713 ***		44.431 *	209.191 ***
Ever Poor/Fair MH	-.071 ***	.085 ***	-.051 **		.134 ***		
Never Poor/Fair GH		-.032 **					.046 *
B. Females							
1940–1960 Ever Drank Heavily		-.073 ***	.054 **				
1940 Overweight							
1982 Physical Activity, Freq.							
1991 Ever Smoked							
1940–1960 Any Organization							.066 **
Ever Divorced					.045 *		-.111 **
Lifetime earnings, 3%		31.623 *					44.451 *
Ever Poor/Fair MH					.152 ***		
Never Poor/Fair GH					-.044 **		.116 ***

Notes: Calculations are based on the Terman data. Letters denote: C, Conscientiousness; O, Openness; E, Extraversion; A, Agreeableness; N, Neuroticism. Coefficients are reported with accompanying statistical significance represented by stars, where $***, **, *$ indicates $p < 0.01, 0.05, 0.10$ respectively. A coefficient with no star refers to $p < 0.15$, while a blank cell refers to a coefficient with p-value above 0.15. p-values are calculated using bootstrap techniques and are *unadjusted*. Coefficients shaded in blue (light grey in print) did not survive the stepdown adjustment whereas coefficients shaded in lilac (dark grey in print) survived.

7 Conclusions

The importance of personality skills in the analysis of health is gaining recognition among economists. We contribute to this emerging literature by investigating the role of the Big Five personality skills on health-related outcomes based on unique life-cycle prospective data with personality and IQ measured early in life and find that their effects are substantial. For males, we find that Conscientiousness benefits the health-related outcomes explored in this paper on a statistically significant level. We report the negative effects of Openness, Agreeableness and Neuroticism on health, and the mixed effects of Extraversion. For females, fewer results can be distinguished from zero, as expected based on related longevity research (Savelyev, 2014).

We also find that education has a statistically significant effect on several important health-related outcomes including alcohol consumption and earnings. This adds new evidence from the Terman data to the literature regarding the causal effect of education on health. We also find that the role of personality skills in explaining health outcomes is comparable to that of education, IQ, and background controls combined, at least for a sample of high-IQ people.[35] Under assumptions of the statistical model, our estimates of the effect of education choice and psychological skills on health-related outcomes can be interpreted as causal effects.

The findings with regard to personality skills open up a new dimension for economists to consider. If childhood personality skills are malleable, and socially-acceptable interventions are possible (through better schooling and parenting, for example), then we can improve health outcomes by specifically encouraging Conscientiousness and Emotional Stability (the inverse of Neuroticism). The estimated effects of Openness, Agreeableness, and Extraversion on health-health related outcomes suggest that we are less sure about these skills as potentially valuable policy targets because of mixed or small effects.

[35] IQ is expected to be more predictive for general population.

References

Ackerman, P. L. and E. D. Heggestad (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin 121*(2), 219–245.

Albouy, V. and L. Lequien (2009). Does compulsory education lower mortality? *Journal of Health Economics 28*, 155–168.

Almlund, M., A. Duckworth, J. J. Heckman, and T. Kautz (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, and L. Wößmann (Eds.), *Handbook of the Economics of Education*, Volume 4, pp. 1–181. Amsterdam: Elsevier.

Anderson, T. W. and H. Rubin (1956). Statistical inference in factor analysis. In J. Neyman (Ed.), *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 5*, pp. 111–150. Berkeley: University of California Press.

Becker, G. S. (2007, July). Health as human capital: synthesis and extensions. *Oxford Economic Papers 59*(3), 379–410.

Becker, G. S. and C. B. Mulligan (1997, August). The endogenous determination of time preference. *Quarterly Journal of Economics 112*(3), 729–758.

Bogg, T. and B. W. Roberts (2004, November). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin 130*(6), 887–919.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.

Borghans, L., B. H. H. Golsteyn, J. J. Heckman, and J. E. Humphries (2011). IQ, achievement, and personality. Unpublished manuscript, University of Maastricht and University of Chicago (revised from the 2009 version).

Buckles, K., A. Hagemann, O. Malamud, M. S. Morrill, and A. K. Wozniak (2013, July). The effect of college education on health. NBER Working Paper NO. 19222.

Burks, B. S., D. W. Jensen, L. M. Terman, A. M. Leahy, H. Marshall, and M. H. Oden (1930). *Genetic Studies of Genius: The Promise of Youth: Follow-Up Studies of a Thousand Gifted Children*, Volume 3. Stanford University, CA: Stanford University Press.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 5, pp. 1801–1863. New York: North-Holland.

Clark, D. and H. Royer (2013). The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review 103*(6), 2087–2120.

Conti, G. and C. Hansman (2013). Personality and the education-health gradient: A note on "Understanding differences in health behaviors by education". *Journal of Health Economics 32*, 480–485.

Conti, G., J. J. Heckman, and S. Urzúa (2010a). The early origins of the education-health gradient. Under revision, *Perspectives on Psychological Science*.

Conti, G., J. J. Heckman, and S. Urzúa (2010b, May). The education-health gradient. *American Economic Review: Papers and Proceedings 100*(2), 1–5.

Cookson, H. (1994). Personality variables associated with alcohol use in young offenders. *Personality and Individual Differences 16*, 179–182.

Crum, R. M., J. E. Helzer, and J. C. Anthony (1993, June). Levels of education and alcohol abuse and dependence in adulthood: A further inquiry. *Journal of Public Health 83*(6), 830–838.

DeYoung, C. G., J. B. Peterson, and D. M. Higgins (2005). Sources of Openness/Intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality and Individual Differences 73*(4), 825–858.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology 41*, 417–440.

Droomers, M., C. T. M. Schrijvers, K. Stronks, D. van de Mheen, and J. P. Mackenbach (1999). Educational differences in excessive alcohol consumption: The role of psychosocial and material stressors. *American Journal of Preventive Medicine 29*, 1–10.

Flory, K., D. R. Lynam, R. Milich, C. G. Leukefeld, and R. Clayton (2002). The relations among personality, symptoms of alcohol and marijuana abuse, and symptoms of comorbid psychopathology: Results from a community sample. *Experimental and Clinical Psychopharmocology 10*, 425–34.

Friedman, H. S. (2000). Long-term relations of personality and health: Dynamism, mechanisms, tropisms. *Journal of Personality 68*(6), 1089–1107.

Friedman, H. S. (2008, July). The multiple linkages of personality and disease. *Brain, Behavior, and Immunity 22*(5), 668–675.

Friedman, H. S., P. H. Hawley, and J. S. Tucker (1994). Personality, health, and longevity. *Current Directions in Psychological Science 3*(2), 37–41.

Friedman, H. S., M. L. Kern, and C. A. Reynolds (2010). Personality and health, subjective well-being, and longevity. *Journal of Personality 78*(1), 179–216.

Friedman, H. S., J. S. Tucker, J. E. Schwartz, L. R. Martin, C. Tomlinson-Keasey, D. L. Wingard, and M. H. Criqui (1995). Childhood conscientiousness and longevity: Health behaviors and cause of death. *Journal of Personality and Social Psychology 68*(4), 696–703.

Friedman, H. S., J. S. Tucker, C. Tomlinson-Keasey, J. E. Schwartz, D. L. Wingard, and M. H. Criqui (1993). Does childhood personality predict longevity? *Journal of Personality and Social Psychology 65*(1), 176–185.

Galama, T. and H. v. Kippersluis (2013). Joint investment in education and health: A unified theory of human capital formation. Unpublished manuscript, USC Dornsife Center for Economic and Social Research.

Gensowski (2013). The effects of education, personality, and IQ on earnings of high-ability individuals. Unpublished manuscript, University of Chicago, Department of Economics.

Goodwin, R. D. and H. S. Friedman (2006). Health status and the five-factor personality traits in a nationally representative sample. *Journal of Health Psychology 11(5)*, 643–654.

Gorsuch, R. (1983). *Factor Analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Grossman, M. (1972, March-April). On the concept of health capital and the demand for health. *Journal of Political Economy 80*(2), 223–255.

Grossman, M. (2004). The demand for health, 30 years later: A very personal retrospective and prospective reflection. *Journal of Health Economics 23*(4), 629–636.

Grossman, M. and R. Kaestner (1997). Effects of education on health. In J. R. Behrman and N. Stacey (Eds.), *The Social Benefits of Education*, pp. 69–124. Ann Arbor, MI: University of Michigan Press.

Hai, R. and J. J. Heckman (2014). A dynamic model of health, education and wealth with credit constraints and rational addiction. Unpublished manuscript, University of Chicago, Department of Economics.

Hampson, S. and H. S. Friedman (2008). Personality and health: A life span perspective. In O. P. John, R. W. Robins, and L. Pervin (Eds.), *The Handbook of Personality*. New York, NY: Guilford Press.

Heckman, J. J., J. E. Humphries, S. Urzúa, and G. Veramendi (2014). Education, health, and wages. NBER Working Paper.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz (2010, August). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics 1*(1), 1–46.

Heckman, J. J., R. Pinto, and P. A. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103*(6), 1–35.

Heckman, J. J., J. Stixrud, and S. Urzúa (2006, July). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*(3), 411–482.

Heckman, J. J. and E. J. Vytlacil (2007). Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative economic estimators to evaluate social programs and to forecast their effects in new environments. In

J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6B, Chapter 71, pp. 4875–5143. Amsterdam: Elsevier.

Heineck, G. and S. Anger (2010). The returns to cognitive abilities and personality traits in Germany. *Labour Economics 17*, 535–546.

John, O. P. and S. Srivastava (1999). The Big Five trait taxonomy: History, measurement and theoretical perspectives. In L. A. Pervin and O. P. John (Eds.), *Handbook of Personality: Theory and Research*, Chapter 4, pp. 102–138. New York: The Guilford Press.

Judge, T. A. and B. A. Livingston (2011). Do nice guys-and gals-really finish last? the joint effects of sex and agreeableness on income. *Journal of Personality and Social Psychology 102*, 390–407.

Lahey, B. B. (2009). Public health significance of neuroticism. *American Psychological Association 64*, 241–256.

Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *Review of Economic Studies 72*(1), 189–221.

Martin, L. R. and H. S. Friedman (2000). Comparing personality scales across time: An illustrative study of validity and consistency in life-span archival data. *Journal of Personality 68*(1), 85–110.

Martin, L. R., H. S. Friedman, and J. E. Schwartz (2007). Personality and mortality risk across the life span: The importance of conscientiousness as a biopsychosocial attribute. *Health Psychology 26*(4), 428–436.

Martin, L. R., H. S. Friedman, J. S. Tucker, C. Tomlinson-Keasey, M. H. Criqui, and J. E. Schwartz (2002). A life course perspective on childhood cheerfulness and its relation to mortality risk. *Personality and Social Psychology Bulletin 28*(9), 1155–1165.

Mazumder, B. (2008). Does education improve health? A reexamination of the evidence from compulsory schooling laws. *Economic Perspectives 32*(2), 2–16.

McCrae, R. R., P. T. J. Costa, and C. M. Busch (1986). Evaluating comprehensiveness in personality heritability California q-set and the five-factor model. *Journal of Personality 54*.

Mueller, G. and E. Plug (2006, October). Estimating the effect of personality on male and female earnings. *Industrial and Labor Relations Review 60*(1), 3–22.

Romano, J. P. and M. Wolf (2005, March). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association 100*(469), 94–108.

Savelyev, P. A. (2014). Psychological skills, education, and longevity of high-ability individuals. Unpublished manuscript, Vanderbilt University, Department of Economics.

Stevenson, B. and J. Wolfers (2007). Marriage and divorce: Changes and their driving forces. *Journal of Economic Perspectives 21*, 27–52.

Terman, L. M. (1986). Terman Life-Cycle Study of Children with High Ability by Terman L. M. et al., 1922-1986 [computer file]. 2nd ICPSR release. Palo Alto, CA: Robert R. Sears [producer], 1986. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1989. doi:10.3886/ICPSR08092.

Terman, L. M., B. T. Baldwin, E. Bronson, J. C. DeVoss, F. Fuller, T. Lee Kelley, M. Lima, H. Marshall, A. H. Moore, A. S. Raubenheimer, G. M. Ruch, R. L. Willoughby, J. Benson Wyman, and D. Hazeltine Yates (1925). *Genetic Studies of Genius: Mental and Physical Traits of a Thousand Gifted Children*, Volume 1. Stanford University, CA: Stanford University Press.

Terman, L. M. and M. H. Oden (1959). *Genetic Studies of Genius: Thirty-Five Years' Follow-Up of the Superior Child*, Volume 5. Stanford, CA: Stanford University Press.

Terman, L. M., M. H. Oden, N. Bayley, H. Marshall, Q. McNemar, and E. B. Sullivan (1947). *Genetic Studies of Genius: The Gifted Child Grows Up: Twenty-Five Years' Follow-Up of a Superior Group*, Volume 4. Stanford University, CA: Stanford University Press.

Terman, L. M. and R. R. Sears (2002a). *The Terman Life-Cycle Study of Children with High Ability, 1922–1986*, Volume 1, 1922–1928. Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

Terman, L. M. and R. R. Sears (2002b). *The Terman Life-Cycle Study of Children with High Ability, 1922–1986*, Volume 3, 1950–1986. Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

Terman, L. M., R. R. Sears, L. J. Cronbach, and P. S. Sears (2002). *The Terman Life-Cycle Study of Children with High Ability, 1922–1986*, Volume 2, 1936–1945. Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

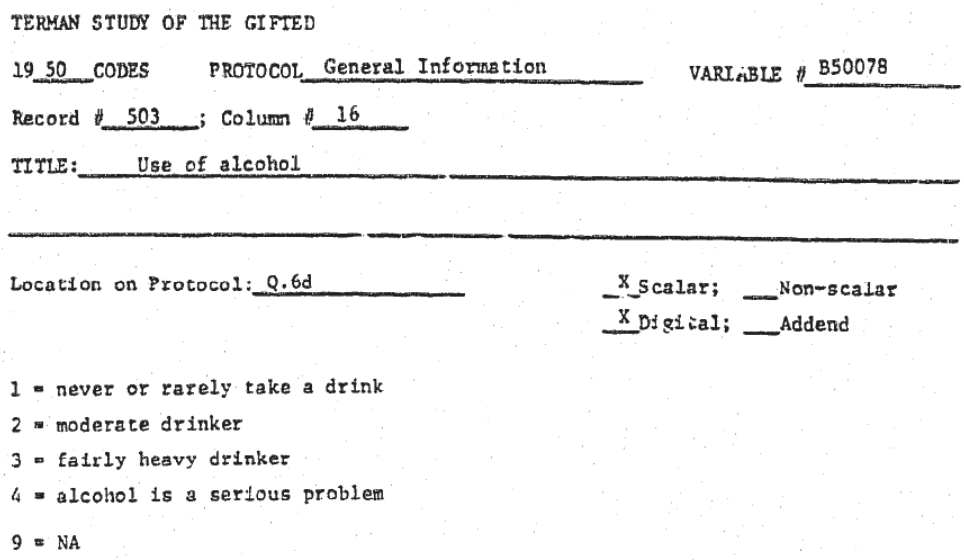
van Kippersluis, H., O. O'Donnell, and E. van Doorslaer (2011). Long-run returns to education: Does schooling lead to an extended old age? *The Journal of Human Resources 46*, 695–721.

Westfall, P. H. and S. S. Young (1993). *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. John Wiley and Sons.

Williams, B. (2011, January). Notes on identification of factor models. Unpublished manuscript, University of Chicago, Department of Economics.

A Appendix

Figure 6: Survey Questions for Heavy Drinking



TERMAN STUDY OF THE GIFTED

19 50 CODES PROTOCOL General Information VARIABLE # B50078

Record # 503 ; Column # 16

TITLE: Use of alcohol

Location on Protocol: Q.6d

X Scalar; ___Non-scalar

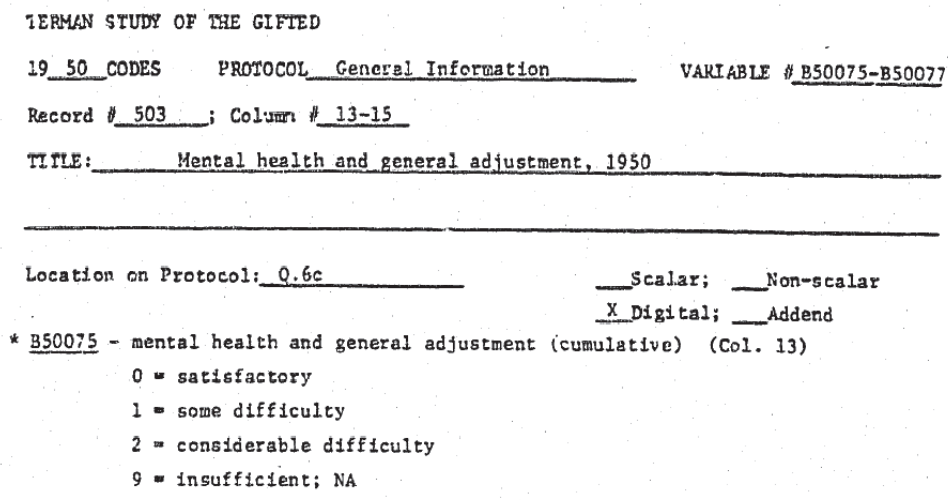
X Digital; ___Addend

1 = never or rarely take a drink
2 = moderate drinker
3 = fairly heavy drinker
4 = alcohol is a serious problem

9 = NA

Notes: Taken from Terman (1986). Ratings 3 and 4 were considered to be indicative of heavy drinking.

Figure 7: Survey Questions for Mental Health



TERMAN STUDY OF THE GIFTED

19 50 CODES PROTOCOL General Information VARIABLE # B50075-B50077

Record # 503 ; Column # 13-15

TITLE: Mental health and general adjustment, 1950

Location on Protocol: Q.6c

___Scalar; ___Non-scalar

X Digital; ___Addend

* B50075 - mental health and general adjustment (cumulative) (Col. 13)
0 = satisfactory
1 = some difficulty
2 = considerable difficulty
9 = insufficient; NA

Source: Taken from Terman (1986).

Table 10: Heavy Drinking

		Females					Males				
		mean	effect	std. error	p-value	stepdown adjusted p-value	mean	effect	std. error	p-value	stepdown adjusted p-value
Ever Drank Heavily in 1940–1960	Conscientiousness	.205	-.011	(.028)	.695	.886	.394	-.055	(.023)	***.018***	***.036***
	Openness		-.073	(.028)	***.011***	***.036***		.035	(.025)	.167	.362
	Extraversion		.054	(.024)	***.028***	***.091***		.061	(.024)	***.009***	***.031***
	Agreeableness		-.017	(.024)	.470	.597		-.045	(.027)	.103	.158
	Neuroticism		.020	(.023)	.403	.403		.036	(.023)	.118	.248
	Cognition		.033	(.021)	.106	.246		.024	(.021)	.234	.411
	Bachelor's degree or above		-.019	(.039)	.627	.627		-.109	(.046)	***.013***	***.038***
	Joint				***.002***	***.008***				***.000***	***.001***
Drank Heavily in 1940	Conscientiousness	.102	.005	(.022)	.822	.822	.267	-.046	(.022)	***.033***	***.065***
	Openness		-.040	(.025)	.112	.201		.008	(.024)	.728	.926
	Extraversion		.008	(.018)	.678	.678		.044	(.024)	***.071***	.127
	Agreeableness		-.041	(.019)	***.040***	***.072***		-.048	(.025)	***.063***	.170
	Neuroticism		.020	(.017)	.239	.331		.033	(.022)	.132	.252
	Cognition		.029	(.016)	***.059***	.222		.057	(.022)	***.011***	***.031***
	Bachelor's degree or above		-.042	(.031)	.172	.426		-.086	(.046)	***.063***	.118
	Joint				***.027***	***.053***				***.001***	***.002***
Drank Heavily in 1950	Conscientiousness	.038	-.009	(.014)	.523	.892	.118	-.012	(.017)	.518	.518
	Openness		.008	(.014)	.559	.559		.003	(.019)	.893	.893
	Extraversion		.014	(.011)	.206	.320		.040	(.016)	***.011***	***.040***
	Agreeableness		-.016	(.010)	***.055***	.180		-.031	(.017)	***.083***	.172
	Neuroticism		.018	(.011)	***.057***	.244		.039	(.017)	***.016***	***.067***
	Cognition		.013	(.010)	.206	.352		.009	(.015)	.554	.554
	Bachelor's degree or above		-.014	(.021)	.540	.746		-.090	(.034)	***.013***	***.029***
	Joint				.176	.176				***.002***	***.004***
Drank Heavily in 1960	Conscientiousness	.167	-.012	(.027)	.659	.919	.347	-.072	(.026)	***.004***	***.018***
	Openness		-.060	(.028)	***.030***	***.089***		.056	(.028)	***.038***	.147
	Extraversion		.049	(.024)	***.038***	.119		.044	(.026)	***.080***	***.080***
	Agreeableness		.013	(.023)	.523	.523		-.020	(.029)	.475	.475
	Neuroticism		.028	(.023)	.161	.383		.024	(.025)	.317	.317
	Cognition		.019	(.020)	.358	.358		.030	(.024)	.205	.435
	Bachelor's degree or above		-.036	(.040)	.373	.663		-.077	(.051)	.138	.138
	Joint				***.023***	***.068***				***.007***	***.007***

Notes: Calculations are based on the Terman data. Statistically significant p-values at the 10% level are bolded. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005).

Table 11: Mental Health

		Females					Males				
		mean	effect	std. error	p-value	stepdown adjusted p-value	mean	effect	std. error	p-value	stepdown adjusted p-value
Ever Had Mental Difficulty in 1940–1960	Conscientiousness	.463	.026	(.029)	.345	.619	.415	-.071	(.025)	***.002***	***.004***
	Openness		-.007	(.034)	.833	.959		.085	(.027)	***.000***	***.004***
	Extraversion		.018	(.029)	.525	.846		-.051	(.024)	***.034***	***.066***
	Agreeableness		-.022	(.030)	.478	.698		-.010	(.028)	.704	.882
	Neuroticism		.152	(.026)	***.000***	***.000***		.134	(.022)	***.000***	***.000***
	Cognition		.000	(.023)	.987	.987		.007	(.022)	.774	.936
	Bachelor's degree or above		-.009	(.049)	.876	.961		-.026	(.044)	.526	.740
	Joint				***.000***	***.000***				***.000***	***.000***
Mental Difficulty in 1940	Conscientiousness	.328	.015	(.029)	.619	.619	.278	-.078	(.024)	***.002***	***.006***
	Openness		-.006	(.032)	.863	.863		.086	(.024)	***.000***	***.002***
	Extraversion		.022	(.027)	.381	.756		-.077	(.023)	***.004***	***.006***
	Agreeableness		-.028	(.028)	.298	.559		.002	(.025)	.918	.918
	Neuroticism		.137	(.025)	***.000***	***.000***		.120	(.020)	***.000***	***.000***
	Cognition		.000	(.023)	.989	1.000		-.013	(.020)	.536	.870
	Bachelor's degree or above		-.040	(.048)	.400	.683		-.052	(.042)	.204	.474
	Joint				***.000***	***.000***				***.000***	***.000***
Mental Difficulty in 1950	Conscientiousness	.333	.042	(.031)	.167	.409	.272	-.040	(.024)	***.088***	***.088***
	Openness		.015	(.033)	.612	.944		.034	(.026)	.190	.190
	Extraversion		.015	(.029)	.610	.839		.008	(.025)	.762	.762
	Agreeableness		.003	(.029)	.895	.895		-.030	(.028)	.280	.536
	Neuroticism		.134	(.027)	***.000***	***.000***		.111	(.021)	***.000***	***.000***
	Cognition		.010	(.024)	.655	.916		.001	(.019)	.946	.946
	Bachelor's degree or above		.001	(.049)	.970	.970		.033	(.043)	.460	.768
	Joint				***.000***	***.000***				***.000***	***.000***
Mental Difficulty in 1960	Conscientiousness	.344	.024	(.030)	.422	.619	.294	-.080	(.027)	***.006***	***.008***
	Openness		-.009	(.034)	.777	.976		.091	(.027)	***.004***	***.008***
	Extraversion		.012	(.028)	.681	.681		-.101	(.024)	***.000***	***.000***
	Agreeableness		-.044	(.031)	.141	.355		.030	(.028)	.288	.598
	Neuroticism		.123	(.027)	***.000***	***.000***		.120	(.022)	***.000***	***.000***
	Cognition		-.017	(.024)	.490	.829		.009	(.022)	.640	.942
	Bachelor's degree or above		-.042	(.050)	.403	.734		.003	(.044)	.948	.948
	Joint				***.000***	***.000***				***.000***	***.000***

Notes: Calculations are based on the Terman data. Statistically significant p-values at the 10% level are bolded. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005).

Table 12: General Health

		Females					Males				
		mean	effect	std. error	p-value	stepdown adjusted p-value	mean	effect	std. error	p-value	stepdown adjusted p-value
Never Poor or Fair Health in 1940-1960	Conscientiousness	.873	.024	(.023)	.287	.679	.926	.012	(.015)	.428	.659
	Openness		-.012	(.023)	.621	.925		-.032	(.014)	***.028***	***.070***
	Extraversion		.009	(.020)	.640	.934		.019	(.014)	.169	.283
	Agreeableness		-.016	(.015)	.259	.259		.005	(.013)	.659	.880
	Neuroticism		-.044	(.018)	***.009***	***.009***		-.021	(.015)	.157	.157
	Cognition		-.017	(.019)	.394	.754		.003	(.011)	.745	.992
	Bachelor's degree or above		.116	(.035)	***.000***	***.002***		.046	(.026)	***.078***	.237
	Joint				***.002***	***.003***				***.096***	***.096***
General Health in 1940	Conscientiousness	-.016	-.029	(.066)	.657	.880	.029	.043	(.057)	.446	.446
	Openness		-.021	(.074)	.805	.946		-.081	(.058)	.161	.277
	Extraversion		-.096	(.064)	.124	.360		.103	(.051)	***.058***	.155
	Agreeableness		-.133	(.056)	***.019***	***.071***		-.029	(.059)	.671	.944
	Neuroticism		-.318	(.051)	***.000***	***.000***		-.279	(.056)	***.000***	***.000***
	Cognition		.009	(.052)	.880	.880		-.007	(.044)	.890	.890
	Bachelor's degree or above		.283	(.106)	***.004***	***.019***		.164	(.108)	.114	.223
	Joint				***.000***	***.000***				***.000***	***.000***
General Health in 1950	Conscientiousness	-.005	-.011	(.067)	.872	.872	-.004	.135	(.052)	***.008***	***.036***
	Openness		.018	(.072)	.769	.769		-.152	(.056)	***.014***	***.032***
	Extraversion		.029	(.065)	.668	.863		.096	(.052)	***.060***	.147
	Agreeableness		-.094	(.052)	***.056***	.148		.003	(.055)	.956	.956
	Neuroticism		-.267	(.055)	***.000***	***.000***		-.242	(.056)	***.000***	***.000***
	Cognition		-.025	(.055)	.610	.944		-.020	(.043)	.671	.972
	Bachelor's degree or above		.172	(.100)	***.094***	.154		.143	(.105)	.173	.173
	Joint				***.000***	***.000***				***.000***	***.000***
General Health in 1960	Conscientiousness	.006	.048	(.081)	.522	.880	-.012	.056	(.058)	.351	.697
	Openness		-.070	(.080)	.405	.805		-.040	(.065)	.560	.560
	Extraversion		.017	(.071)	.835	.835		.068	(.071)	.329	.329
	Agreeableness		-.087	(.060)	.148	.244		.094	(.062)	***.090***	.355
	Neuroticism		-.241	(.060)	***.002***	***.002***		-.211	(.058)	***.000***	***.000***
	Cognition		.021	(.054)	.704	.895		-.011	(.044)	.841	.978
	Bachelor's degree or above		.060	(.109)	.582	.582		.197	(.120)	***.096***	.243
	Joint				***.003***	***.003***				***.001***	***.002***

Notes: Calculations are based on the Terman data. Statistically significant p-values at the 10% level are bolded. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005).

Table 13: Physical Activity and BMI

		Females					Males				
		mean	effect	std. error	p-value	stepdown adjusted p-value	mean	effect	std. error	p-value	stepdown adjusted p-value
Exercise	Conscientiousness	.173	.027	(.027)	.316	-	.176	.002	(.024)	.926	-
	Openness		-.048	(.031)	.118	-		-.044	(.026)	***.089***	-
	Extraversion		-.039	(.029)	.185	-		.000	(.027)	.988	-
	Agreeableness		-.016	(.028)	.583	-		.011	(.030)	.709	-
	Neuroticism		-.025	(.028)	.364	-		-.066	(.026)	***.012***	-
	Cognition		.026	(.023)	.252	-		-.007	(.020)	.720	-
	Bachelor's degree or above		.010	(.046)	.829	-		.108	(.058)	***.062***	-
	Joint				.402					***.059***	
BMI	Conscientiousness	.160	.009	(.024)	.718	-	.225	-.007	(.022)	.807	-
	Openness		.010	(.025)	.699	-		-.014	(.023)	.456	-
	Extraversion		-.012	(.023)	.595	-		-.001	(.022)	.783	-
	Agreeableness		-.010	(.025)	.688	-		-.037	(.023)	.120	-
	Neuroticism		-.037	(.021)	***.071***	-		.021	(.021)	.842	-
	Cognition		-.006	(.017)	.730	-		-.019	(.018)	.159	-
	Bachelor's degree or above		-.074	(.042)	***.075***	-		.012	(.044)	.661	-
	Joint				.486					.571	
Ever Smoked	Conscientiousness	.425	-.052	(.050)	.294	-	.521	-.107	(.045)	***.017***	-
	Openness		-.019	(.054)	.731	-		.016	(.055)	.773	-
	Extraversion		.013	(.045)	.770	-		.067	(.046)	.145	-
	Agreeableness		-.003	(.050)	.951	-		.057	(.048)	.228	-
	Neuroticism		.023	(.045)	.606	-		.029	(.048)	.549	-
	Cognition		.002	(.033)	.940	-		.017	(.039)	.672	-
	Bachelor's degree or above		.089	(.081)	.273	-		.165	(.119)	.163	-
	Joint				.772					.162	

Notes: Calculations are based on the Terman data. Statistically significant p-values at the 10% level are bolded. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005).

Table 14: Group Membership

		Females					Males				
		mean	effect	std. error	p-value	stepdown adjusted p-value	mean	effect	std. error	p-value	stepdown adjusted p-value
Any Organization in 1940–1960	Conscientiousness	.900	.027	(.018)	.123	.319	.937	.000	(.013)	.969	.969
	Openness		.000	(.019)	.978	.978		.004	(.014)	.732	.924
	Extraversion		.024	(.020)	.225	.511		.020	(.012)	***.096***	.340
	Agreeableness		.003	(.023)	.901	.985		.009	(.015)	.534	.808
	Neuroticism		.003	(.017)	.852	.852		-.011	(.013)	.377	.621
	Cognition		.006	(.013)	.581	.581		-.017	(.011)	***.002***	***.004***
	Bachelor's degree or above		.066	(.033)	***.044***	***.044***		.084	(.027)	.107	.203
	Joint				.400	.400				***.007***	***.022***
# of Organizations in 1940	Conscientiousness	2.506	.162	(.115)	.156	.273	2.435	.086	(.092)	.349	.786
	Openness		-.138	(.124)	.244	.526		-.107	(.102)	.281	.721
	Extraversion		.132	(.111)	.233	.407		.073	(.085)	.399	.619
	Agreeableness		.056	(.096)	.559	.916		.034	(.095)	.728	.728
	Neuroticism		.101	(.085)	.238	.604		-.072	(.078)	.355	.732
	Cognition		-.093	(.085)	.278	.458		-.175	(.078)	.111	.111
	Bachelor's degree or above		.789	(.175)	***.000***	***.000***		.245	(.154)	***.028***	***.081***
	Joint				***.001***	***.003***				.114	.114
# of Organizations in 1950	Conscientiousness	1.565	.186	(.098)	***.062***	.198	2.714	.007	(.120)	.972	1.000
	Openness		.046	(.124)	.707	.919		-.034	(.132)	.765	.765
	Extraversion		.169	(.104)	***.084***	.333		-.001	(.132)	.993	.993
	Agreeableness		.007	(.112)	.958	.958		.258	(.114)	***.015***	***.087***
	Neuroticism		-.104	(.092)	.244	.535		-.008	(.107)	.926	.926
	Cognition		-.120	(.079)	.143	.352		.012	(.104)	***.000***	***.000***
	Bachelor's degree or above		.877	(.153)	***.000***	***.000***		1.172	(.198)	.919	.919
	Joint				***.000***	***.000***				***.000***	***.000***
# of Organizations in 1960	Conscientiousness	2.567	.072	(.154)	.676	.676	3.423	-.033	(.149)	.810	.993
	Openness		.248	(.162)	.117	.394		.142	(.164)	.373	.741
	Extraversion		-.139	(.157)	.346	.346		-.182	(.142)	.203	.505
	Agreeableness		-.135	(.170)	.383	.877		.255	(.144)	***.076***	.190
	Neuroticism		.105	(.137)	.454	.685		-.214	(.122)	***.085***	.257
	Cognition		-.352	(.128)	***.007***	***.022***		.327	(.121)	***.000***	***.000***
	Bachelor's degree or above		1.213	(.256)	***.000***	***.000***		1.501	(.244)	***.002***	***.031***
	Joint				***.000***	***.000***				***.000***	***.000***

Notes: Calculations are based on the Terman data. Statistically significant p-values at the 10% level are bolded. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005).

Table 15: Marriage Status

		Females					Males				
		mean	effect	std. error	p-value	stepdown adjusted p-value	mean	effect	std. error	p-value	stepdown adjusted p-value
Never Married	Conscientiousness	.085	.017	(.016)	.281	.492	.061	.023	(.012)	***.048***	.106
	Openness		.012	(.021)	.559	.559		.012	(.011)	.263	.263
	Extraversion		-.001	(.017)	.937	.937		-.028	(.014)	***.036***	.158
	Agreeableness		.004	(.019)	.805	.805		-.017	(.013)	.189	.556
	Neuroticism		.005	(.017)	.784	.940		.024	(.011)	***.024***	.120
	Cognition		-.023	(.012)	***.044***	.209		-.014	(.009)	.120	.341
	Bachelor's degree or above		.074	(.022)	***.002***	***.007***		-.006	(.022)	.797	.797
	Joint				***.006***	***.024***				***.006***	***.019***
Married Once and Still Married	Conscientiousness	.456	-.003	(.029)	.912	.912	.614	.056	(.025)	***.022***	***.072***
	Openness		-.024	(.035)	.494	.717		-.048	(.028)	***.091***	.192
	Extraversion		-.008	(.031)	.770	.963		.011	(.026)	.687	.871
	Agreeableness		.029	(.032)	.297	.691		.031	(.029)	.301	.621
	Neuroticism		-.006	(.030)	.833	.833		-.010	(.023)	.694	.694
	Cognition		-.007	(.026)	.775	.775		.012	(.021)	.627	.866
	Bachelor's degree or above		.129	(.051)	***.012***	***.053***		.120	(.048)	***.010***	***.050***
	Joint				.235	.235				***.014***	***.027***
Ended up Divorced	Conscientiousness	.133	-.041	(.022)	***.063***	.251	.068	-.023	(.014)	***.084***	***.084***
	Openness		.039	(.027)	.146	.466		.050	(.015)	***.002***	***.007***
	Extraversion		.023	(.019)	.244	.596		-.027	(.014)	***.045***	.158
	Agreeableness		-.013	(.020)	.506	.726		.000	(.015)	.971	.971
	Neuroticism		.011	(.021)	.617	.933		.024	(.012)	***.053***	.127
	Cognition		.008	(.016)	.594	.937		-.018	(.010)	***.074***	.264
	Bachelor's degree or above		-.039	(.035)	.262	.262		-.007	(.025)	.773	.969
	Joint				***.089***	.178				***.006***	***.022***
Ever Divorced	Conscientiousness	.253	-.032	(.027)	.227	.524	.267	-.055	(.022)	***.012***	***.055***
	Openness		.036	(.032)	.248	.575		.039	(.024)	.105	.189
	Extraversion		.027	(.026)	.278	.629		-.020	(.023)	.392	.698
	Agreeableness		-.038	(.026)	***.077***	.323		-.045	(.026)	***.072***	.307
	Neuroticism		.045	(.027)	***.079***	.237		.026	(.021)	.230	.331
	Cognition		.022	(.022)	.318	.705		-.012	(.019)	.510	.878
	Bachelor's degree or above		-.111	(.046)	***.019***	***.046***		-.137	(.044)	***.002***	***.012***
	Joint				***.000***	***.001***				***.000***	***.001***
Divorced at least Twice	Conscientiousness	.064	-.034	(.019)	***.070***	.253	.067	-.044	(.015)	***.002***	***.022***
	Openness		.022	(.019)	.213	.619		.031	(.014)	***.026***	***.094***
	Extraversion		.026	(.015)	***.077***	.295		-.004	(.012)	.773	.773
	Agreeableness		.034	(.016)	***.035***	.107		.002	(.013)	.854	.978
	Neuroticism		.028	(.016)	***.086***	.244		.025	(.012)	***.043***	.129
	Cognition		.005	(.011)	.640	.870		-.002	(.010)	.785	.785
	Bachelor's degree or above		-.054	(.026)	***.056***	***.086***		-.042	(.027)	.117	.331
	Joint				***.008***	***.023***				***.024***	***.024***

Notes: Calculations are based on the Terman data. Statistically significant p-values at the 10% level are bolded. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005).

Table 16: Life-cycle Earnings

		Females					Males				
		mean	effect	std. error	p-value	stepdown adjusted p-value	mean	effect	std. error	p-value	stepdown adjusted p-value
Lifetime Wages	Conscientiousness	271.997	-7.761	(16.370)	.824	.937	1115.885	34.003	(34.462)	.299	.393
Discounted 3%	Openness		31.623	(18.171)	.433	.500		-75.290	(43.168)	***.090***	.192
	Extraversion		-16.679	(16.495)	.622	.805		79.908	(32.102)	***.014***	***.026***
	Agreeableness		-23.526	(16.693)	.120	.397		-94.713	(36.553)	***.004***	***.018***
	Neuroticism		-24.824	(15.653)	***.084***	.263		-48.963	(32.901)	.130	.178
	Cognition		-1.203	(11.061)	.889	.889		44.431	(24.693)	***.074***	.120
	Bachelor's degree or above		44.451	(26.910)	.164	.300		209.191	(63.163)	***.000***	***.000***
	Joint				.230	.230				***.000***	***.000***
Wages at age 40	Conscientiousness	11.541	-.677	(1.090)	.538	.941	62.231	.801	(2.489)	.725	.725
	Openness		.533	(1.178)	.637	.637		-3.908	(3.389)	.253	.429
	Extraversion		-.967	(1.013)	.347	.504		3.843	(2.326)	.102	.176
	Agreeableness		-.842	(1.148)	.475	.475		-6.556	(2.346)	***.004***	***.008***
	Neuroticism		.104	(.961)	.912	.912		-1.940	(2.187)	.325	.325
	Cognition		3.946	(1.686)	.210	.576		3.280	(1.723)	***.056***	.118
	Bachelor's degree or above		-.887	(.707)	***.008***	***.084***		14.585	(4.875)	***.000***	***.000***
	Joint				.390	.390				***.000***	***.000***
Wages at age 50	Conscientiousness	16.083	.391	(1.368)	.767	.767	71.503	4.606	(2.739)	***.088***	.198
	Openness		1.904	(1.468)	.183	.305		-5.601	(3.124)	***.072***	.210
	Extraversion		-.687	(1.539)	.628	.628		4.122	(2.564)	.120	.120
	Agreeableness		-1.716	(1.330)	.176	.380		-6.787	(2.725)	***.012***	***.018***
	Neuroticism		-1.549	(1.307)	.221	.357		-6.553	(2.622)	***.012***	***.032***
	Cognition		3.303	(2.278)	.786	.952		4.758	(2.228)	***.030***	***.094***
	Bachelor's degree or above		.298	(1.119)	.130	.231		19.788	(4.990)	***.000***	***.000***
	Joint				.296	.296				***.000***	***.000***
Wages at age 60	Conscientiousness	15.327	-1.028	(1.440)	.464	.947	62.001	4.404	(3.250)	.182	.305
	Openness		3.903	(1.696)	***.021***	.279		-3.562	(3.365)	.291	.291
	Extraversion		-1.533	(1.509)	.328	.845		5.814	(2.805)	***.040***	***.090***
	Agreeableness		-1.652	(1.571)	.288	.456		-4.637	(3.375)	.182	.182
	Neuroticism		-4.650	(1.453)	***.002***	***.004***		-7.466	(3.173)	***.012***	***.042***
	Cognition		2.551	(2.469)	.511	.880		.637	(2.483)	.802	.802
	Bachelor's degree or above		.654	(1.078)	.311	.311		30.530	(5.399)	***.000***	***.000***
	Joint				***.029***	.116				***.000***	***.000***

Notes: Calculations are based on the Terman data. Statistically significant p-values at the 10% level are bolded. p-values are calculated using bootstrap techniques, and further adjusted using the stepdown procedure in Romano and Wolf (2005). Earnings measures are net of tuition paid for schooling and taxes, in 2010 US dollars.