MASTER THESIS: ECONOMETRICS & MANAGEMENT SCIENCE
SPECIALISATION: ECONOMETRICS

INTERGENERATIONAL TRANSMISSION OF SKILLS AND OCCUPATIONAL CHOICE USING STRUCTURAL EQUATION MODELLING

August 2, 2021

ERASMUS UNIVERSITEIT ROTTERDAM

Erasmus School of Economics

Author: Aarthi S.Iyer

Student ID number: 523320

Supervisor: dr. AJ Koning

Second assessor: dr. O Kleen

Abstract

Development of skills by the youth can help address the issues of employment, entrepreneurship, improve economic, equitable and sustainable growth. Skills can be categorised into cognitive, non-cognitive and technical skills. This paper seeks to answer whether the occupation choice of an individual is influenced by the transmission of skills from their parents or their occupation. Since these skills are unobserved variables, the methodology used for answering the research question is Structural Equation Modelling(SEM) which makes use of latent and observed variables. The results found indicate that positive intergenerational transmission of cognitive skills which impact the occupational choice of an individual. However, for non-cognitive skills, no evidence of intergenerational transmission of skills from mother is found, but positive evidence of influence of parental investment, self-productivity and cross-productivity is found which impacts the occupational choice of an individual.

Contents

1 Introduction

In the labor market, workers are often classified as "White Collar" or "Blue Collar" workers which helps in understanding the social standing of the person, approach to life, life choices among others. This classification also tells about the kind of skill the individual possess which in turn determine their success in the labor market. The heterogeneity in the labor market can be classified into two types, namely, labor services defined in terms of different knowledge, tasks, output or equipment, and in terms of individuals who supply labor in the market. Economic theory recognises that individual exhibit differences in both their productive capabilities and their preference for varieties of utility and disutility associated with supply of labor. Because of which individuals are not expected to suit each role and these differences contribute to determinants of individual's occupational outcomes, thereby choosing varied labor market roles (Ham et al., 2009b). One of the important factors contributing to the country's economic growth and positive outlook is when it achieves increased employability along with labor productivity which can be attained if a worker achieves productivity by working towards their full working potential. With global changing trends such as technological advancements, climate change and urbanization, the skills needed are evolving continuously and rapidly. In this ever changing environment, the skills in high demand are cognitive skills such as critical thinking and problem solving, and socio-economic skills such as leadership, teamwork and grit along with relevant technical skills. Due to the presence of heterogeneity in the labor market, breaking down the job roles into the required skill set can allow employers to understand viable job transition pathways and at the same time also allow the employers to make decisions regarding reskilling or upskilling required for such transitions (World Economic Forum, 2021). However, the labor market especially in developing countries workers are not able to reach their full working potential due to low skills levels which is partly due to low education levels or due to skill mismatch. Such a lack of skilled workers has limited the innovation that stems for employers and also affects a country's economic growth. Reskilling of workers or matching their skill with respective occupation can help in alleviating the problem of skill mismatch highlighting the problem of occupational choice of an individual.

There are multiple sources of heterogeneity affecting the labor market, one of them identified is human capital between individuals. Becker (2009) argues that there are diverse array of factors such as education level, on the job skill training, and experience which can be seen as investments in human capital increasing the productivity of an individual, though the effects of education may have non-linear effects over the years (Heckman et al., 2003).

Recently, the literature in economics has expanded the research in understanding the role of personality traits, a significant part of non-cognitive skills to explain multiple economic behavior as well as one of the sources of heterogeneity affecting labor market outcomes. Over the years, empirical economics has paid a lot of attention to cognitive skills such as reading, mathematics, science overlooking the other abilities which are as important as such abilities in determining an individual's development and success in varied life outcomes Brunello and Schlotter (2011) which include abilities such as social skills, motivation, leadership known as non-cognitive skills and personality traits. Heckman and Rubinstein (2001) define these non-cognitive skills as "dark matter" in economics which has potential explanatory power in individual economic and social outcomes. A growing amount of literature has shown evidence that non-cognitive skills play a crucial role in explaining various individual outcomes and social achievements Heckman, Stixrud, and Urzua (2006) having high predictive outcomes in the areas of education, occupation, health, wage determination, and crime rates. As suggested by Heckman et al. (2006) both cognitive and non-cognitive skills play multiple roles in explaining schooling decisions and educational attainment which indirectly affects occupational choice of an individual.

The literature evidences from Europe and US supports the idea that certain level of non-cognitive skills are pre-requisite in avoiding labor market failures (Brunello & Schlotter, 2011). By using the HILDA dataset, Cobb-Clark and Tan (2011) demonstrate that personality traits can explain sorting of individuals in different occupations. Similar personality traits can be found within same groups of occupation (John & Thomsen, 2014). Occupational choice of an individual chosen according to their personality traits affects their performance and overall individual productivity, and understanding the mechanism behind this can help in improving the labor market outcomes. Cubel, Nuevo-Chiquero, Sanchez-Pages, and Vidal-Fernandez (2016) through a laboratory experiment illustrate the correlation between Big Five personality traits and productivity. According to them the study of link between personality and productivity is important for two reasons: (i) First, employers themselves are interested in understanding this relationship, (ii) Second, understanding to what extent personality affects labor market outcomes and productivity can help in laying foundation for policy making. Employers are also now looking beyond academic achievement while hiring potential employees. Kuhn and Weinberger (2005) report the findings of National Association of Colleges and Employers survey which states that employers most valued skills are communication, motivation/initiative, teamwork, and leadership skills than academic achievement or grade point average.

A widely accepted taxonomy in the empirical economics literature to measure personality traits is the Big Five Model which includes the factors : agreeableness, conscientiousness, openness, extraversion and neuroticism. *Agreeableness* is defined as the tendency to act in an unselfish, cooperative manner. *Conscientiousness* is when an individual acts in an organised, responsible and hardworking manner. *Openness* is the tendency of an individual to be open to new culture, experience, and aesthetic. *Extraversion* associates with individuals who have a preference for human contacts, empathy, and assertiveness. Individual with extraversion prefer more outer world people and things rather than inner world of subjective things. They are sociable and have positive effect. *Neuroticism* refers to an individual has emotional instability (Todd & Zhang, 2020). By using the Big Five Model as a measure of personality traits, John and Thomsen (2014) provide evidence that occupational sorting is influenced by non-cognitive skills and that there is an inter-dependency of personality, occupation and wages, underlying the importance of occupation specific evaluation of returns. Ham et al. (2009b) use the HILDA survey to study the effects on the probability of being in a white collar occupation and find that along with human capital, parental status and personality traits have significant effects on occupational outcomes, and that effect of conscientiousness is larger. Ham, Junankar, and Wells (2009a) find that human capital has non-linear effects on occupational choice, and that parental status has minimal effect whereas the Big Five Model has a significant, persistent effect over occupational outcomes. Using the HILDA survey data, they find are that managers are less agreeable and more antagonistic; labourers are less conscientiousness; and sales people are more extraverted. Similar kind of evidence is also found by Nieken and Störmer (2010) using the German Socio-Economic Panel(GSOEP) data set providing the evidence of individuals with different personality traits associated with different occupations.

Apart from human capital and personality traits, another source of heterogeneity contributing to varied occupations in labor market is the role of parents. The influence of parents on occupational choice can be considered via two channels : firstly, the effect of status of individual's parents within the society which is referred as "dynasty hysteresis", and secondly through intergenerational transmission skills. In the first way of influence, dynasty hysteresis can be defined as an individual's parents achievements, abilities, and skills can influence an individual's occupational choice and due to such a transfer mechanism, parents and offspring's are often found in the same occupation. Existing literatures provide evidences of dynasty hysteresis which can be caused due to different reasons, such as human capital transfer, religion and its associated characteristics, social groups, and preferences (Ham et al., 2009a). Regardless of different causes, dynasty is an important phenomenon which has a huge potential to influence an individual's decisions regarding occupational choice. Constant and Zimmermann (2003) examine the effects of parental social status and find that it has significant effect on an individual's occupational choice. They find that family background affects occupational choice through genetic endowment, social connections, wealth, and indirectly through education. Though, evidences of causes and presence of dynasty hysteresis has been documented in literatures, the mechanism is still under debate.

In the second way of influence an individual's occupational choice is can be seen through intergenerational transmission of skills from parents which can be both cognitive and non-cognitive skills. There are two main channels of transmission of cognitive and non-cognitive skills between generations posited. It can be transmitted through inheritance of genes(nature) and through productivity effect of parental skills (nurture) (Anger, 2012). Cognitive skills are based on past learning and are more strongly transmitted when related to innate ability highlighting the importance of parental investments in children's cognitive outcomes (Anger & Heineck, 2010). Cunha and Heckman (2008) find evidence that parental input affects the formation of both cognitive and non-cognitive and that parental inputs affect the formation of cognitive skills more strongly at earlier ages. Chevalier et al. (2002) conclude that occupational choices of UK graduates are same as that of their father after 6-11 years of graduation and that their decision is mainly based on their father's education and occupation. An individual's skill level changes with time from childhood into adulthood as a result of education, work experience and training. de Coulon et al. (2011) conduct a research to study how strong is parent's adult skill levels help in predicting their children's cognitive and non-cognitive skills and their results indicate that parent's with high numeracy and literacy skills have children with higher cognitive and non-cognitive skills.

The literature review highlights that in the labor market occupational choice of an individual to a great extent influence their productivity levels. It also highlights that occupational choice of a personal can be influenced through multiple factors such as cognitive skills, non-cognitive skills, occupation of parents among others. With this, the objective of the thesis is to understand the intergenerational transmission of both cognitive and non-cognitive skills from parent to child and its role in influencing the occupational choice of the individual and among these two skills which is more influential in impacting the occupational choice of an individual?

The outline of rest of the thesis follows in this way: In section 2, LISREL methodology is described in detail following which in section 2.3, an example is illustrated to elucidate the methodology. In section 3, an overview of the model is described consisting of three parts, namely, measurement model, structural model, and occupational choice model along with identification of factor loadings. Section 4 discusses the estimation techniques for the measurement model and structural model. It also discusses the reliability of factors and various measures of goodness-of-fit indices. Section 5 presents the data along with the different sets of variables considered which is followed by results of the estimated model in section 6. Finally, in section 7 conclusion along with discussion is presented with further scope of research.

2 Methodology

The methodology employed to answer the research question is the Structural Equations Modelling(SEM) also known as Linear Structural Relations Model
(LISREL). It uses various theoretical models that defines different hypothesis on how different sets of variables define constructs and how these constructs are related to each other. The goal of the SEM is to determine the extent to which theoretical model is supported by the sample data. To understand SEM further, two important variables need to be defined. Firstly, the observed variables or also known as indicator variables are a set of variables that are used to define or infer the latent variables. Secondly, latent variables (constructs or factors) which are not directly observed or measured but are indirectly observed or measured which are inferred from a set of observed variables such as surveys, tests and so on.

Both observed and latent variables can be endogenous or exogenous. In econometrics, an exogenous variable is a variable that is not caused by other variables in the solution and an endogenous variable is caused by one or more variables in the model. Thus, exogenous variables can be viewed as independent variables and endogenous variables can be viewed as dependent variables (Brown & Moore, 2012). Within the context of structural modelling, exogenous variables represent the constructs that exert the influence on other constructs under study and are not influenced by other factors in the model and endogenous variables are the variables which are affected by exogenous and other endogenous variables present in the model (Schreiber et al., 2006). To understand the concept of exogenous and endogenous variables in the context of structural equation model with the help of an example from (Schumacker & Lomax, 2016, p. 180). Let *Intelligence* indicated as latent independent variable which is supposed to predict *Stochastic Achievement* known as latent dependent variable can be depicted as,

$$Intelligence \longrightarrow Achievement_1$$

A latent dependent variable when has one arrow pointing from another latent variable is often referred to as endogenous variable ($Achievement_1$) and when one latent variable does not have any arrow pointing to it then it is often referred to as exogenous variable (Intelligence). When a third latent variable is added which is depicted as follows ,

$$Intelligence \longrightarrow Achievement_1 \longrightarrow Achievement_2$$

In this model, Intelligence is still the exogenous variable, while $Achievement_2$ is the latent dependent variable and hence endogenous. However, $Achievement_1$ has one from Intelligence while it point one arrow to $Achievement_2$ which makes it first a dependent variable and then an independent variable. This model illustrates the indirect effects of latent variables. In this case $Achievement_1$ becomes the mediating latent variable.

One of the main assumptions of the SEM models is that the observed variables used to measure the latent variables should be reflective in nature, .i.e., they share the same underlying concept as the latent variables. However, in many cases formative indicators are applied (also known as causal measures). Formative indicators are the measures that form or cause the creation or change in a latent variables. Inclusion of formative indicators becomes problematic as the main assumption is that correlations among the observed variables for a particular latent variable is only caused by that particular latent variable (Chin, 1998).

The LISREL model consists of two parts, a measurement model and a structural equation model. In the measurement model, latent variables or constructs are specified on how they depend on observed variables. It describes the measurement properties, reliabilities, and validities of the observed variables in describing the latent variables. The structural equation component of the model is a regression model, specifying the causal relationship among latent variables and assigning explained and unexplained variances (Tsai, 2006). If the fitted model is poor, then it will be more likely due to the misspecification in the measurement model and hence it is important that an acceptable measurement model is established before interpreting the structural relationship among latent variables. SEM models provide flexibility in terms of the interplay between theory and data. More specifically, (i) it models relationship among multiple predictor and criterion variables, (ii) construct unobservables latent variables, (iii) model errors in the measurements for observed variables and,(iv) statistically apriori assumptions for empirical data (Chin, 1998).

2.1 Measurement Model

The first part of the LISREL model is the measurement model which evaluates how the observed variables combine to identify the underlying hypothesised constructs The main objective of the measurement model is to establish the reliability and validity of the observed variables in relation with the latent variables. The relationship between observed variables and latent variables is indicated by factor loadings highlighting the extent to which the observed variables are able to measure the latent variables. Along with the factor loadings, the measurement model also produces measurement error associated with the observed variables. It specifically highlights the extent to which observed variables are measuring something other than the proposed latent variable (Shanmugam & Marsh, 2015). While generating the measurement model defining the latent variables the following questions needs to addressed: To what extent are the observed variables actually measuring the latent variables? Which of the observed variables are the best predictors of the latent variables? Are the observed variables measuring something other than latent variables?

2.1.1 Confirmatory factor analysis

The measurement model uses confirmatory factor analysis (CFA) model to confirm that a hypothesised latent variable can be inferred from observed variables. In CFA, one can find out the extent to which the items used to measure a latent variable are related to one another but also the latent variable can be determined by examining the loadings of the observed variables. CFA is an effect indicator model. The use of CFA goes beyond rejecting or confirming a factor model, where one can also revise, refine and retest the model using a well defined set of criteria (Phakiti, 2018). It specifies the number of factor loadings to reflect that only certain factors can influence certain factor indicators by fixing many or all cross loadings equal to zero. And by incorporating the prior knowledge in the form of restrictions, the definition of latent variables is more subjective and also leads to a parsimonious models. However, fixing many or all cross loadings equal to zero may lead to a parsimonious model than suitable to the data which then leads to multiple respecification Asparouhov and Muthén (2009). In the confirmatory factor model, the relationship between observed and latent variables is explained. Each factor indicates gives the information to what extent an observed variable is able to measure the latent variable. To avoid misspecification in the model, one has to define latent variable accurately so that the measures defined are strongly correlated with each. If weakly correlated, then the constructs will be poorly defined and lead to model misspecification (Weston & Gore Jr, 2006).

In the CFA model, there are three types of parameters specified, namely, free, fixed, or constrained. A free parameter is unknown which has to be estimated by minimizing the differences between the observed and predicted variance-covariance matrices. A fixed parameter is pre-specified to a specific value mostly equalling to 1 or 0. A constrained parameter is unknown, however the parameter is not free to be any value, some restrictions are placed on the value it may assume. One of the most common constraints is the equality constraints in which some of the parameters in CFA solutions is restricted to be equal in value (Brown & Moore, 2012).

CFA is confined to the analysis of variance-covariance structures and the analysis of the covariance structures is based on the assumption that indicators are measured as deviation from the means (Brown & Moore, 2012). While conducting CFA, the researcher uses a hypothesised model to estimate a population covariance matrix which is then compared with the observed covariance matrix. The main goal is to minimise the difference between the estimated and observed matrices(Schreiber et al., 2006).

2.2 Structural Model

The second part of the LISREL model is the structural model which is similar to the regression model explaining the causal relationship between independent and dependent variables as specified in the measurement model. The structural model explains how the latent variables are related among each other. After specifying the hypothesised structural model, it can be tested to determine the extent to which these priories are supported by the sample variance-covariance data. Each structural equation contains a prediction error which explains the portion of variance that is not explained by the latent variables (Schumacker & Lomax, 2016, p. 187).

2.3 LISREL Illustrated Example

In this section two examples have been illustrated in figure 1 and 2 to explain the LISREL model combining the measurement and structural part of the model. The example in figure 1 is a properly specified structural equation some parameters are fixed while some are free which needs to be estimated and figure 2 explains a simple form of a MIMIC model.

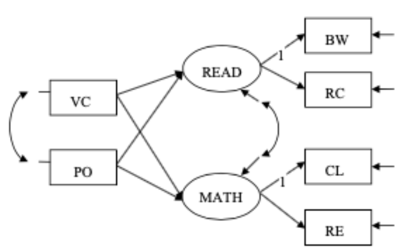


FIGURE 2. Conceptual model of math and reading.

Figure 1: Example of a simple LISREL model from (Lei & Wu, 2007)

Figure 1 shows a model that predicts reading(READ) and mathematics(MATH) latent ability variables from two observed scores of intelligence namely, verbal comprehension(VC) and perpetual organisation(PO). The latent variable READ is indicated by basic word reading(BW) and reading comprehension(RC) scores, while the latent variable MATH is indicated by calculation(CL) and reasoning(RE) scores. The paths denoted by directional arrow from VC and PO to READ and MATH, from READ to BW and RC, and from MATH to CL and RE along with curved arrows between VC and PO, and between residuals of READ and MATH are the free parameters to be estimated in the model as well as the residuals of the endogenous and exogenous variables. The remaining paths which are not shown will not be estimated and are fixed to zero. The scale of the latent variable can be standardised by fixing its variance to 1 or the latent variable can take the scale of one of its indicator variables by fixing the first factor loading of one indicator to 1. When the parameter value of a path is fixed to a constant, the parameter will not be estimated from the data.

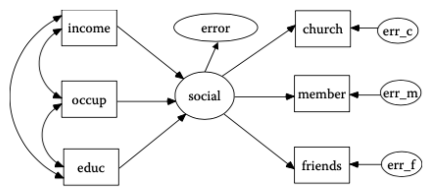


FIGURE 15.1a
MIMIC Model.

Figure 2: Example of a MIMIC model from (Schumacker & Lomax, 2016, p. 294)

MIMIC model is a special case of SEM models known as multiple indicators and multiple causes model which involves using latent variables predicted by the observed variables. Figure 2 is a representation of a simple MIMIC model in which latent variable social participation is defined by observed variables church attendance, memberships, and friends. Further the latent variable social participation is predicted by observed variables income, occupation and education. In the above MIMIC model, social which is a latent variable has arrows pointing out to three observed variables(church, member,and friends) along with their respective measurement error for each. This part is the measurement part of the model. Now, the latent variable social also has arrows pointed towards it from three observed predictor variables which are correlations among them. This is the structural part of the model which uses observed variables to predict latent variables (Schumacker & Lomax, 2016).

3 Model

The model used to answer the research objective combines multiple indicator multiple cause (MIMIC) and LISREL model helping in the process of identification of restrictions of cross equations. MIMIC is a special case of the SEM models which permits the specification of one or more latent variables with one or more observed variables as predictors of latent variables (Schumacker & Lomax, 2016, p. 293). It incorporates additional variables which are assumed to influence the latent factors. The MIMIC model introduces the causes of latent factors. As proposed by Jöreskog and Goldberger (1975) , in the MIMIC model, one observes multiple indicators and multiple causes of a single latent variable. The observed variables result from the latent factors and the latent factors themselves are caused by other exogenous variables. Thus, there is a measurement equation and a causal relationship (Krishnakumar & Nagar, 2008). The model comprises of three parts, the first part is the measurement model using confirmatory factor model, second part is the structural model which uses factor score regression and the third part is a factor score regression to understand the influence of intergenerational transmission of skills from parents and occupation of parents in the individual's choice of occupation.

3.1 Part I: Measurement Model

The first part of the model is a measurement model for which a confirmatory factor model is fitted to understand the extent to which the underlying indicators measure the latent variable. The CFA model was fitted for individual and mother's cognitive, non-cognitive skills along with parental investments. The measurement model to account for the individual's cognitive and non-cognitive skills is given by,

$$X^k_{i,t-1} = \Lambda^k_{0i,t-1} + \Lambda^k_{1i,t-1}\theta^k_{t-1} + \varepsilon^k_{1i,t-1} \tag{1}$$

$$X^k_{i,t} = \Lambda^k_{0i,t} + \Lambda^k_{2i,t}\theta^k_{t} + \varepsilon^k_{2i,t} \tag{2}$$

with $k \in (C, N, I, PC, PI)$. X represents the observed measures of latent variables where $i = 1, ...m^k$ denoting different indicators of specific latent variable. Here C is the cognitive skills of the individual, N is the non-cognitive skills of the individual, PC is the mother's non-cognitive skills, PN is the mother's non-cognitive skills and I is the parental investment. θ is the factor for the latent variable and $\Lambda_{i,t}$ is the respective factor loading.

The following assumptions have to be made:
(i) The factors i.e. $(\theta_t^C, \theta_t^N, \theta_t^{PC}, \theta_t^{PN}, \theta_t^I)$ and the error term $\varepsilon_{j,t}^k$ are uncorrelated and the mean of error term is zero.
(ii) $\varepsilon_{j,t}^k$ has mean zero and is independent across agents and over time for $k \in$(C,N,PC,PN,I).
$(iii)\varepsilon_{j,t}^k$ is independent from $\varepsilon_{j,t}^l$ for $k = l$.

3.2 Part II: Structural Model

The structural model is a regression model which uses the factor scores obtained from the confirmatory factor model to estimate the cognitive and non-cognitive skills.

$$\theta_{t+1}^C = \delta_1\theta_t^C + \delta_2\theta_t^N + \delta_3\theta_t^I + \delta_4\theta_t^P C + \delta_5\theta_t^P N + \eta_t^C \tag{3}$$

$$\theta_{t+1}^N = \delta_1\theta_t^C + \delta_2\theta_t^N + \delta_3\theta_t^I + \delta_4\theta_t^P C + \delta_5\theta_t^P N + \eta_t^N \tag{4}$$

where θ_t^k represents the cognitive and non-cognitive skills of an individuals and parents with $k \in (C, N, I, PC, PN)$ respectively at time period t and θ_{t+1}^k represents the cognitive and non-cognitive skills of individual at time period $t+1$. Equation 3 and 4 determines how stocks of cognitive and non-cognitive skills at time period t affect the next time period $t+1$ allowing to examine the self-productivity and cross-productivity.

3.3 Part III: Impact of skills on occupational choice

After obtaining the estimates of θ_t^k , anchor the estimates of the factor scales using the occupational choice of an individual.

$$Y_a = \theta_{t+1}^C + \theta_{t+1}^N + Y_f + Y_m + \epsilon \tag{5}$$

where Y_z are the occupations with $z \in (a, f, m)$ respectively of individual, father, and mother. Y_z takes the value 1 is the occupation is a white collar job and 0 if the occupation is a blue collar job

3.4 Identification of factor loadings

The identification problem is to understand whether or not θ is uniquely identified by the covariance matrix (Jöreskog, 1978). Identification must be demonstrated before estimation is done. A necessary condition for identification is that the number of parameters estimated is less than or equal to the number of non-redundant elements of the sample covariance matrix of the observed variables. Identification is usually solved by solving for the parameters of the model in terms of the variances and covariances of the observed variables. This can be done by first identifying the parameters of the measurement model, including the variances and covariances of the factors. Once the covariance among the factors are identified then the structural parameters can be identified by solving in terms of covariances among the factors (Long, 1983). If the information matrix obtained from the likelihood estimation is positive definite then the model is high likely to be identified (Joereskog & Sörbom, 1984).

Identification for two measurements per latent factors i.e. $m_t^k = 2$ where $k \in (C, N, PC, PN)$. The identification of $Cov(X_{1,t}^k, Y_{2,\tau}^l)$ can be computed for all k, l pairs and for all time periods t, τ ∈1,...,T.

This is the case where $k \neq l$

$$Cov(X_{1,t}^k, X_{1,t+1}^l) = Cov(\theta_t^k, \theta_t^l) \tag{6}$$

$$Cov(X_{1,t}^k, X_{2,t}^l) = \lambda_t^k Cov(\theta_t^k, \theta_t^l) \tag{7}$$

$$\frac{Cov(X_{1,t}^k, X_{2,t}^l)}{Cov(X_{1,t}^k, X_{1,t}^l)} = \lambda_t^k \tag{8}$$

From equation 8 factor loadings are obtained from which the covariance across latent skills can be obtained written as the ration of observed indicators and the identified loading,

$$Cov(\theta_t^k, \theta_t^l) = \frac{Cov(X_{1,t}^k, X_{2,t}^l)}{\lambda_t^k} \tag{9}$$

When $k = l$,

$$Var(X_{1,t}^k, X_{1,t}^k) = Var(\theta_t^k) + \epsilon_t^k \tag{10}$$

$$Cov(X_{1,t}^k, X_{2,t}^l) = \lambda_t^k Var(\theta_t^k) \tag{11}$$

$$Var(\theta_t^k) = \frac{\lambda_t^k Var(\theta_t^k)}{\lambda_t^k} \tag{12}$$

where ϵ_t^k is also identified.

In case of parental investments (I), the latent variables are correlated across ages in which the identification is as follows,

$$Cov(X_{1,t-1}^I, X_{1,t}^I) = Cov(\theta_{t-1}^I, \theta_t^I) \tag{13}$$

$$Cov(X^I_{2,t-1}, X^I_{1,t}) = \lambda^I_{t-1} Cov(\theta^I_{t-1}, \theta^I_t) \tag{14}$$

$$\frac{Cov(X^I_{2,t-1}, X^I_{1,t})}{Cov(X^I_{1,t-1}, X^I_{1,t})} = \lambda^I_{t-1} \tag{15}$$

$$Cov(\theta^I_{t-1}, \theta^I_t) = \frac{Cov(X^I_{2,t-1}, X^I_{1,t})}{\lambda^I_{t-1}} \tag{16}$$

Identification of variance and error terms

$$Var(X^I_{1,t-1}, X^I_{1,t-1}) = Var(\theta^I_{t-1}) + \epsilon^I_{t-1} \tag{17}$$

$$Cov(X^I_{1,t-1}, X^I_{2,t-1}) = \lambda^I_{t-1} Var(\theta^I_{t-1}) \tag{18}$$

$$Var(\theta^I_{t-1}) = \frac{\lambda^I_{t-1} Var(\theta^I_{t-1})}{\lambda^I_{t-1}} \tag{19}$$

4 Estimation

4.1 Estimation of measurement models

In the estimation of LISREL, it is assumed that the multivariate distribution is normally distributed. To check the normality assumption, the distribution of the observed variables can be examined. Estimation of model based on incorrect assumptions can lead to drawing incorrect conclusions. Multivariate statistical models are are most widely used in linear structural relations among observed and latent variables where these variables are usually non-normally distributed because of which classical multivariate analysis based on error free variables which have no simultaneous interactions is not the right way to deal with such datasets (Yuan & Bentler, 1997). When non-normality situation arises distribution free methods can be used such as the asymptotic distribution method (ADF). The ADF method is a weighted least squares method in which the weight matrix has to specified properly in order to guarantee the asymptotic properties of the standard normal theory estimators and test statistics obtained (Bentler & Dudgeon, 1996). The ADF method minimises the following generalised least squares function

$$Q = (s - \sigma(\theta))' W (s - \sigma(\theta)), \tag{20}$$

to get an optimal parameter estimate $\hat{\theta}$ and where W is the weight matrix given by $W = \hat{V}^{-1} = V(\hat{\theta})^{-1}$ which is a consistent estimator. The matrix V is the asymptotic distribution of residual given by,

$$\sqrt{n}(s - \sigma) \xrightarrow{D} Normal(0, V). \tag{21}$$

The elements in the V matrix is given by

$$v_{ij,kl} = \sigma_{ijkl} - \sigma_{ij}\sigma_{kl}, \tag{22}$$

where σ_{ij} is the sample covariance matrix. σ_{ijkl} is given by

$$\sigma_{ijkl} = E(z_i - \mu_i)(z_j - \mu_j)(z_k - \mu_k)(z_l - \mu_l) \tag{23}$$

The S matrix which is the sample covariance matrix obtained from the variables which are observed independently. In some situations the independence assumption can be difficult to achieve which might require special methods. The paper by Bentler and Dudgeon

(1996) the argument is that the values of the parameters can be estimated from the sample covariance matrix S and can be tested for the fit of the model in the population covariance matrix $\Sigma(\theta)$ by minimizing some scalar function $F = F[S, \Sigma(\theta)]$. The scalar function F indicates the discrepancy between S and $\Sigma(\theta)$. The discrepancy function F has the following properties: (i) the value of F will be greater than or equal to zero ,(ii) F will only be equal to zero if $S = \Sigma(\theta)$, and (iii) F must be twice differentiable with respect to both S and $\Sigma(\theta)$. Some of the examples of the discrepancy function are ML and GLS. In case of a distribution free deployed, the results obtained will always be optimal as the discrepancy function would always be correctly specified which was introduced by Browne (1982) and minimum distance method by (Chamberlain, 1982).

A study by Benson and Fleishman (1994) compared the robustness of both ML and ADF methods using Monte Carlo simulation to investigate the combined effects of sample size, magnitude of correlation among observed indicators, number of indicators, magnitude of skewness and kurtosis. Their results indicated a little bias in the factor loadings in both ML and ADF estimation under all conditions studied. The bias was seen more in ADF than ML as the number of indicators increased. Increase in skewness and kurtosis showed underestimation of standard errors with ML standard errors being more biased than ADF under non-normality conditions, and ML chi-square was also inflated. A comparison study between pseudo-likelihood (p-ML) and ADF by Molenaar and Nesselroade (1998) shows that both the methods tend to give consistent parameter estimates, but the standard errors and chi-square statistics appears to be more consistent in the ADF method.

The model fit is of the CFA is found to be inadequate, can be improved by using the modification indices in which certain parameters are added or deleted. The value given by the modification indices is the minimum amount that the chi-square value is expected to decrease if the corresponding parameters is fixed. A parameter is freed to improve the fit of the model and this process continues until an adequate fit is obtained.

4.2 Estimation of structural regression model

The classic approach of modelling structural equation models is by using one-step approach where both measurement and structural part are estimated simultaneously. However, simultaneous estimation of both parts can suffer from interpretational confounding which gets reflected in the parameter coefficients of the estimates. Interpretational confounding "occurs as the assignment of empirical meaning of unobserved variable which is other than meaning of assigned to it by an individual which is other than meaning assigned to it by an individual a priori to estimating unknown parameters" (Burt, 1976, p. 4). It can be explained as the changes in the coefficients when alternate structural models are estimated. But the interpretational confounding can be minimized by prior separate estimation of the measurement model to put no constraints on the structural parameters. A two-step approach focuses on the tradeoff between goodness of fit and strength of the causal inference. And separate assessment of measurement and structural model preclude having good fit of one model compensate for poor fit for other (Anderson & Gerbing, 1988). According to Bollen (1996) misspecification errors in one part of the equation can spillover to other parts of the equation when simultaneously estimated. The first part consists of measurement equation model in the two-step procedure factor analysis is performed using CFA through which factor scores are calculated, where the factor scores are estimated for the true latent variables. The factor scores are usually predicted using either regression predictor or Bartlett predictor. In the second part, the factor scores are used in linear regression considering them to be true latent variable scores. However, the use of factor scores directly leads to a bias in the estimates obtained. The bias is due to the covariance matrix of the factor scores and the true latent variables being different.

4.2.1 Bias Corrected factor score regression

According to Skrondal and Laake (2001), the performance of conventional factor score regression is quite bad. To avoid bias, their methodology of revised factor score comprised of using regression factor scores for the explanatory latent variables and for the response latent variables Bartlett scores were used to provide consistent estimators for all parameters. This method however fails when there are correlations between independent variables. On the other hand, Croon (2002) corrects bias. According to him, there is a difference between variances and covariances of the factor scores and the latent variables and hence uses an estimation of variances and covariances of true latent variable instead of the factor scores to estimate the regression parameters. This method can be used any estimator and predictor. Devlieger and Rosseel (2017) summarise the method of Croon,

1. Perform factor analysis for all the latent variables separately and obtain the respective factor scores.

2. Calculate the variance-covariance matrix of factor scores.

3. Estimate the true variance and covariances for all the elements in this variance-covariance matrix.

4. Perform regression using the estimated variance-covariance matrix as the input for the covariance in the model.

The bias corrected formula proposed by Croon for the variance and covariance is given as follows,

$$\widehat{var(\xi)} = \frac{var(F_\xi) - A_\xi \Theta_\delta A'_\xi}{A_\xi \Lambda_x \Lambda'_x A'_\xi}, \tag{24}$$

$$\widehat{cov(\xi, \eta)} = \frac{cov(F_\xi, F_\eta)}{A_\xi \Lambda_x \Lambda'_y A'_\eta}. \tag{25}$$

where, Λ_x and Λ_y are the factor loadings, A_ξ and A_η are the factor score matrices, and Θ_δ is the variance-covariance matrix of the unique factor scores of δ. Equation 25 show that corrected covariance can be obtained by dividing the covariance among factor scores and products of the factor scores and loading matrices. One of the key insights of equation 25 is that matrix multiplication in the denominator gives the product of construct reliabilities, i.e., each matrix multiplication of the individual factor score and loading matrices produces empirical estimates of construct reliabilities on the basis of measurement models. To correct the variances, we can set each term of the latent variable to one on the diagonal of the covariance matrix.

4.3 Reliability of factors

The reliability part in the measurement model is the part containing no random error. In terms of SEM models, the reliability of an indicator is defines as the variance of the indicator that is not accounted by the measurement error, represented by squared multiple correlation coefficient ranging from 0 to 1 (Raines-Eudy, 2000). Measurement is an important aspect in the structural equation modelling and presence of measurement error can produce biased estimates among the constructs. The most commonly used reliability measure is the alpha, where the the measurement model is assumed to be tau equivalence which i.e., true score equivalent where the factor loadings are equal. This has found to be quite misleading due to it being restricted and hence unrealistic. An alternative is the coefficient omega by McDonald (1999) in which the reliability estimates are calculated from the parameter estimates of the factor models which are specified to represent associations between the items and constructs. Here the measurement model is assumed to be congeneric model in which the items are affected by a single factor but with varying degrees termed as one factor confirmatory factor model. As the model expands, consisting of more than one factor, it can be termed as multiple regression equation with each x_j regressed on multiple factors (Viladrich, Angulo-Brunet, & Doval, 2017). The coefficient omega is given as follows,

$$\omega_u = \frac{(\sum_{j=1}^{J} \hat{\lambda_j})^2}{\hat{\sigma_X}^2}, \tag{26}$$

$$\omega_h = \frac{(\sum_{j=1}^{J} \hat{\lambda_{jg}})^2}{\hat{\sigma_X}^2} \tag{27}$$

The factor loading is the strength of the association between the items and factors, meaning the extent to which a set of items reliably measures the factors and hence the reliability can be estimated from the parameter estimates fitted to the item scores. Equation 26 is reliability of the one factor model where the numerator represents amount of total variance explained by the common factor as a function of estimated factor loadings, while the denominator represents the estimated variance of the observed total score. On the other hand, equation 27 represents the reliability of a two factor model where λ_{jg} is the factor loading of x_j on a general factor g. The total score variance represented by $\hat{\sigma_X}^2$ can be calculated from the model implied variance of X.

Raykov (1998) proposed a method to obtain standard errors and confidence intervals in a structural equation modelling framework for estimating the reliability of congeneric measures using bootstrap method. In this approach he takes repeated samples from a given sample to estimate reliability. The advantages of this method is that firstly, it is applicable to any set of measures assessing a common latent variable which basically means that it should be a one factor confirmatory model. Secondly, it is applicable to non-normal observed variable distributions. Thirdly, this method can be applied to weighted and unweighted scales.

4.4 Goodness-of-fit indices

Goodness-of-fit and its assessment is a key in the measurement model so as to one can understand how well the model fit is. There are various ways to assess the model fit of the model which determines to which extent the observed (S) and model implied (Σ) variance-covariance matrix differs. The most commonly used measures of model fit chi-square (χ^2), goodness-of-fit index(GFI), adjusted goodness-of-fit index(AGFI) and root-mean-square residual index(RMR). A significant χ^2 value indicates that the observed and implied covariance matrices differ. Usually, one is interested in obtaining a non-significant χ^2 value so that the difference between observed and implied covariance matrix. However, the χ^2 value is very sensitive to sample size and when the observed variables do not follow multivariate normality because of which it can lead to erroneous conclusions. Given the sensitivity of χ^2 to sample size, other measures of goodness-of-fit model are proposed which are some function of chi-square and degrees of freedom. GFI is based on the ratio of sum of squared differences between the observed and model implied covariance matrix to observed variance. The GFI measures the amount of variance and covariance in S that is predicted by the Σ. To assess how well the given model approximates the true model, Root Mean Square Error of Approximation(RMSEA) was introduced. If the approximation is good, then RMSEA should be small where an acceptable value should be below 0.05. Another measure is the Root Mean Square Residual (RMR) which is square root of the mean-squared differences between the matrix elements in S and Σ. An acceptable level of RMR is usually 0.05 and most often standardised RMR are reported.

Several studies such as (Anderson and Gerbing (1984);Bearden et al. (1982)) suggests that parameter estimates are not affected by sample size or model characteristics, but the goodness of fit indices have been affected, especially the chi square statistic, and GFI, RMR. Assumptions of chi-square tests violated by excessive kurtosis, moderate sample sizes, unknown statistical distributions, confidence intervals, little knowledge about the distribution and behaviour of fit indices for misspecified models are some of the reasons. The underlying distribution of test statistic is unknown in many instances and is often mathematically intractable. Bootstrapping is a general procedure for determining the sampling distribution of a parameter estimate whose theoretical sampling distribution is unknown. Previous studies have found that goodness of fit indices such as chi square statistics, root mean square error are affected by model characteristics and as well sample size.

Bootstrap technique can be used to evaluate the goodness of fit indices of a confirmatory factor model. A bias may be introduced if the bootstrapping sampling distribution is non-normal and in such cases confidence interval estimated by percentile method i.e., bias corrected percentile method can be introduced (Bone et al., 1989). The general bootstrapping method can be described where there is a random sample $X_1, X_2, ..., X_N$ of size N with each X_i being drawn independently from the same population having a cumulative distribution G with parameter θ. We want to know the sampling distribution of the estimator $\hat{\theta}$. For the given random variables, there is a sample $x_1, x_2, ..., x_N$. The bootstrap method samples from the population distribution defined by G_N to estimate the sampling distribution of $\hat{\theta}$. We bootstrap sample, $X_1^*, X_2^*, ..., X_N^*$ by taking N draws with replacement from $x_1, x_2, ..., x_N$. The bootstrap estimate $\hat{\theta^*}$ is computed using the bootstrap sample. Repeating the process B times gives $\hat{\theta}^*(1), \hat{\theta}^*(2), ..., \hat{\theta}^*(B)$. With these bootstrap replicates, mean and variance of the $\hat{\theta}$ can be obtained (Bollen & Stine, 1992). The general bootstrapping procedure works in many cases, it can however fail as well.

The naive bootstrapping (completely non-parametric resampling from the original distribution) is inaccurate as the distribution of the bootstrapped model test statistics follows a noncentral chi-square distribution instead of central chi-square distribution. (Bollen & Stine, 1992) formulated a transformation on the original data to adjust for this inaccuracy which forces the resampling space to satisfy the null hypothesis (i.e., making the model-implied covariance matrix the true underlying covariance matrix in the population). The transformation is given by

$$Z = YS^{-1/2}\hat{\Sigma}^{-1/2}, \tag{28}$$

where Y is the original data matrix from the parent sample. Analytically T_{ML} values from bootstrap samples drawn from transformed data matrix Z have an expectation equal to the model df. They also proposed a method to adjust for the p-value associated with T_{ML}. The bootstrap adjusted p-value is calculated as the proportion of bootstrap model test statistics that exceed the value of T_{ML} obtained from the original parent sample.

5 Data

There are two different datasets used for this analysis, one for the parents and the second for the individual. Due to the limited availability of data, the dataset obtained is for the individual with their respective mothers. For the mother, National Longitudinal Survey of Youth 1979 (NLSY79) from where their cognitive skills, non-cognitive skills and occupations variables have been collected. While for the individual NLSY79 Child and Young Adult cohort(NLCY79) is used to obtain their cognitive skills, non-cognitive skills and occupations keeping only those observations which match with mother's ID. The NLSY79 is a longitudinal survey of men and women born during 1957-1964 in the United States when the survey began. Data collected in this survey provides information about the participants choices regarding schooling, decisions regarding education and training, labor market, family life and expectations about their retirement. NLCY79 is the longitudinal project which follows the biological mothers from the NLSY79. The survey starts in 1986 where the mothers were interviewed every two years. The assessment include home environment, PIAT math and reading , Peabody Picture Vocabulary Test, the Digit Span scale of the Wechsler along with detailed health information. Then starting in 1994, children aged 15 and older were interviewed based on the modified questionnaire of NLSY79 designing it according to the current generation. Both these datasets were chosen due to its richness in the availability of variables and its range covering wide topics. For the purpose of this analysis individuals born to NLSY79 mothers after 1985 was only used which was 5722 observations, but for actual estimation only 3754 observations was only available due to the presence of missing values.

Mother's cognitive ability is measured by the SAT and ACT scores in both mathematics and verbal. Both these test scores were chosen to capture the mother's cognitive ability through their performance in standardised tests. Individual's cognitive skills are measured by Peabody Individual Achievement Test(PIAT) for mathematics, reading comprehension and reading recognition. PIAT is a wide range measuring academic achievement for children aged five and older, and the test if of high reliability and validity for the assessment of academic achievement. PIAT Math measures the child's attainment in mathematics taught in mainstream education. PIAT Reading Recognition measures word recognition and pronunciation ability while PIAT Reading Comprehension measures the child's ability to derive meanings from sentences. The items used to measure parental investments is the relationship of the individual with biological parents, time parents spend with the individual on various occasions.

To measure the non-cognitive ability of mother, Rotter Locus of Control and Big Five personality traits are used. Rotter Locus of Control measures the extent to which believe they have control over their lives either through self-motivation and self-determination defined as internal control as opposed to external control that conducts their life such as fate, luck. The higher the score, the individual will be more external. The Big Five personality traits measures the Big Five personality trait. For the individual,to measure non-cognitive skills along with an additional measure Pearlin Mastery Scale and Big Five personality traits. Pearlin Mastery Scale refers to the measure of self-concept and references to the extent to which individuals perceive themselves in control of forces that significantly impact their lives. At t, the individual's age is 19 years in which an individual is assumed to make their educational choices while at $t+1$ individual's age is 23 where the individual will be entering the job market. For simplicity, parental investment and mother's skills are taken to be constant for both t and $t+1$

The Big Five personality trait consists of the following traits Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism which the respondent rates on a scale of 1(Strongly Disagree) to 7 (Strongly Agree). The following are the indicators of each personality trait are give by: Extraverted, enthusiastic (E); Dependable, self-disciplined (C); Anxious, easily upset (N); Open to new experiences, complex (O); Sympathetic, warm (A); Calm, emotionally stable (N, reversed). The indicators for the Rotter Locus of Control are given as follows: Degree of control one has over own life, importance of planning, importance of luck, degree of influence one has over own life. While the indicators of the Pearlin Mastery scale is given as follows: No way I can solve the problems I have, I sometimes feel that I am being pushed around, I have little control over what happens to me, I can do just about anything that happens in my life, I often feel helpless in dealing problems in my life, What happens to me in future mostly depends on me, Little I can do to change things in my life.

Table 1 contains the classification of occupation into white collar and blue collar job for the individual, mother and father. The occupations that come under white collar jobs consists of office-related, non-manual jobs, while blue collar jobs are more of manual and industrial occupations.

Category of Occupation	Occupations
White Collar	Managers, Executives in public and private companies or legislative,Professionals,Scientists, Academics, Lawyers, Doctors and Healthcare professionals
Blue Collar	Servers, Cook, Personal Care Workers, Labourers, Transportation, Setters, Construction workers, Entertainers, Sports, Media and Communication Workers

Table 1: Occupational Groups for both parents and individuals

6 Results

6.1 Summary Statistics

Indicators	Mean	SD	Skewness	Kurtosis
Extraversion	5.114	1.754	-0.837	-0.056
Conscientiousness	5.986	1.507	-1.933	3.300
Agreeableness	5.867	1.543	-1.747	2.617
Neuroticism	3.276	1.942	0.524	-0.873
Openness	5.196	1.665	-0.941	0.301
Neuroticism (Reversed)	5.459	1.633	-1.156	0.675
SAT Mathematics	5.549	1.112	0.218	-0.713
SAT Verbal	4.945	0.904	0.104	0.146
ACT Mathematics	4.986	1.329	-0.838	-0.451
ACT Verbal	4.743	1.034	-0.800	0.214

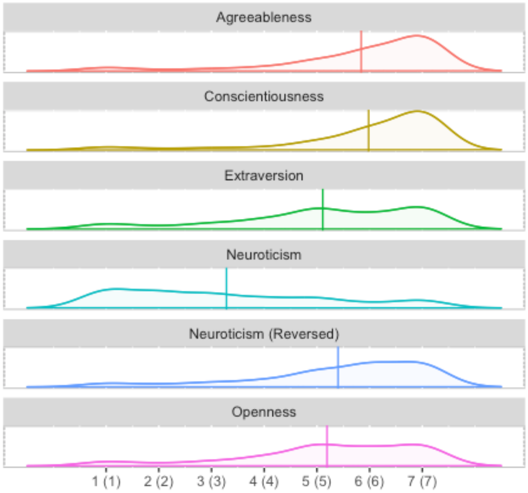
Table 2: Summary statistics for observed indicators for mother's skills

Table 2 and 3 show the summary statistics of the indicators of skills for both mother and the individual. The skewness and kurtosis for a normal distribution is 0 and 3 respectively. The summary statistics indicate that the observed indicators are non-normal. The skewness and kurtosis values indicate that the indicators are moderately non-normal.

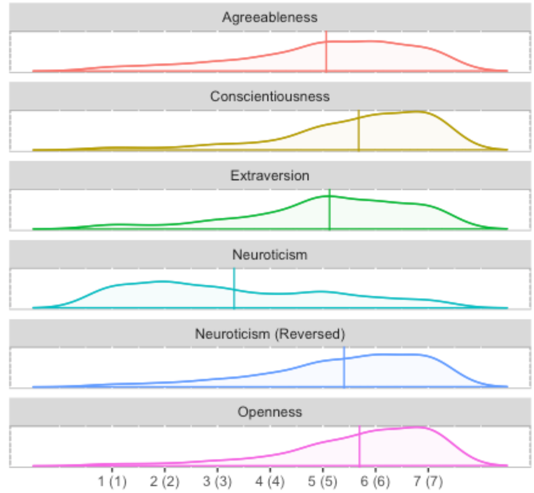
Indicators	**Mean**	**SD**	**Skewness**	**Kurtosis**
Extraversion	5.208	1.478	-0.815	0.401
Conscientiousness	5.810	1.288	-1.378	2.054
Agreeableness	5.152	1.545	-0.817	0.164
Neuroticism	3.516	1.847	0.327	-0.961
Openness	5.705	1.367	-1.261	1.549
Neuroticism (Reversed)	5.322	1.492	-0.939	0.467
Pearlin Mastery Indicator 1	1.987	0.777	0.526	-0.015
Pearlin Mastery Indicator 2	2.065	0.767	0.294	-0.376
Pearlin Mastery Indicator 3	1.875	0.697	0.535	0.344
Pearlin Mastery Indicator 4	3.385	0.601	-0.675	0.848
Pearlin Mastery Indicator 5	1.955	0.683	0.454	0.403
Pearlin Mastery Indicator 6	3.365	0.611	-0.736	1.186
Pearlin Mastery Indicator 7	2.027	0.720	0.417	0.128
PIAT Mathematics	6.369	2.867	-0.351	-1.086
PIAT Reading Recognition	6.684	2.819	-0.513	-0.902
PIAT Reading Comprehension	5.654	2.781	-0.102	-1.171

Table 3: Summary statistics for observed indicators for individual's skills

Figures 3 and 4 shows the density plots of the observed indicators of cognitive and non-cognitive skills for both mother and individual. The vertical lines indicate the means of each group while the curves are density plots which the show the distribution values similar to the histogram for Likert scale data. To check for the normality assumption, chi-square Q-Q plots are plotted which can be seen in figure 5. For both mother and the individual's cognitive and non-cognitive skills the observed indicators in the Q-Q plot are curved rather than being a straight line confirming the results found in table 3 that observed indicators are non-normal.

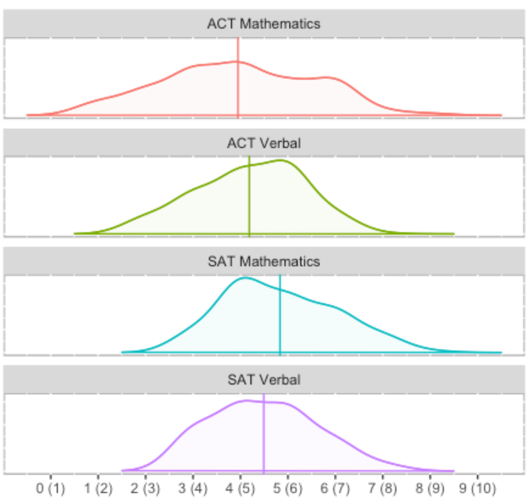


(a) Indicators of mother

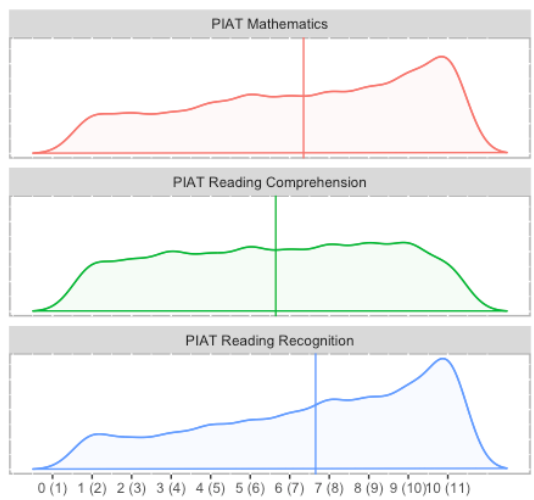


(b) Indicators of individual

Figure 3: Density plot for observed indicators of non-cognitive skills



(a) Indicators of mother



(b) Indicators of individual

Figure 4: Density plot for observed indicators of cognitive skills

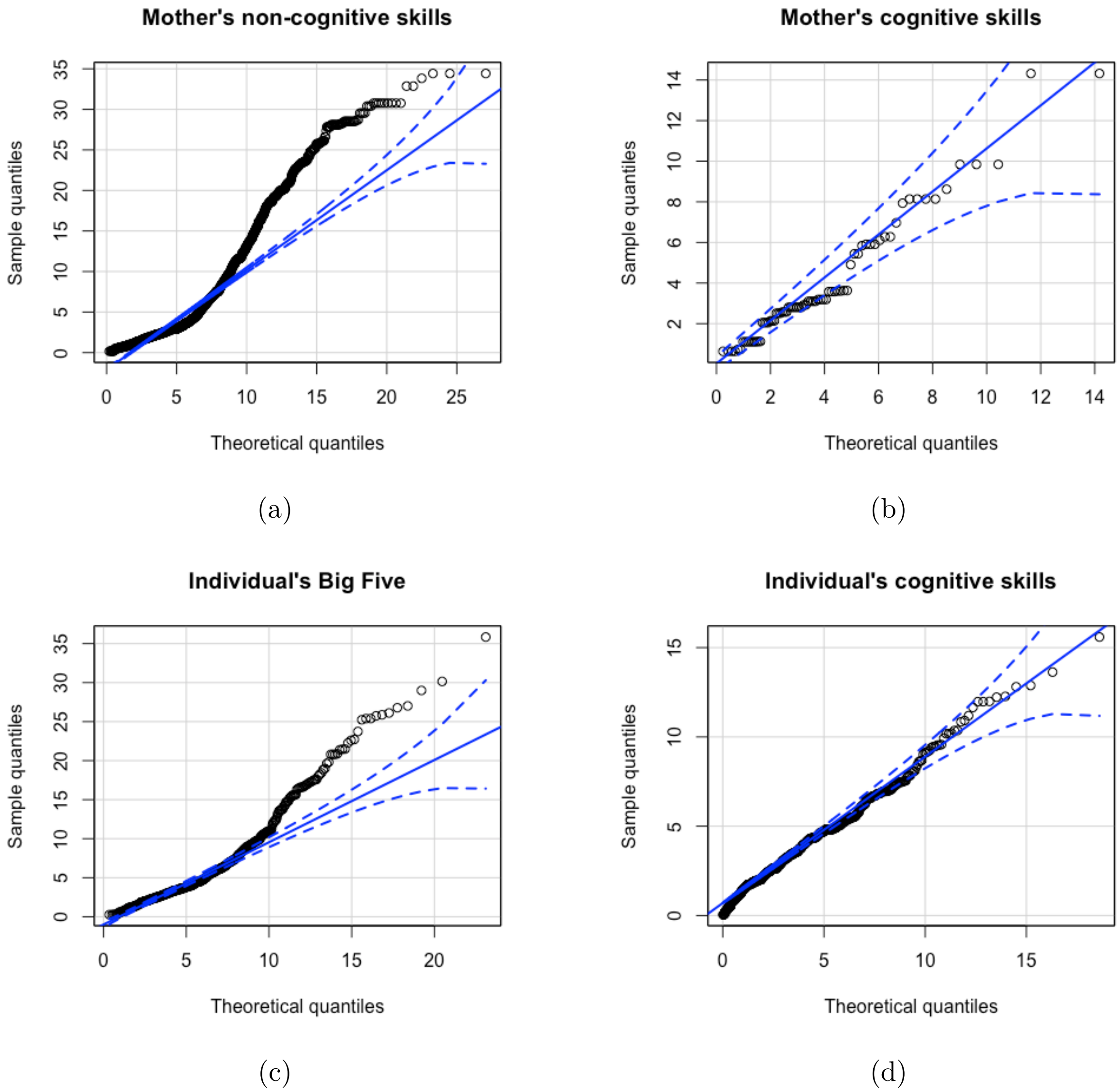


Figure 5: Chi-Square Q-Q plots for observed indicators

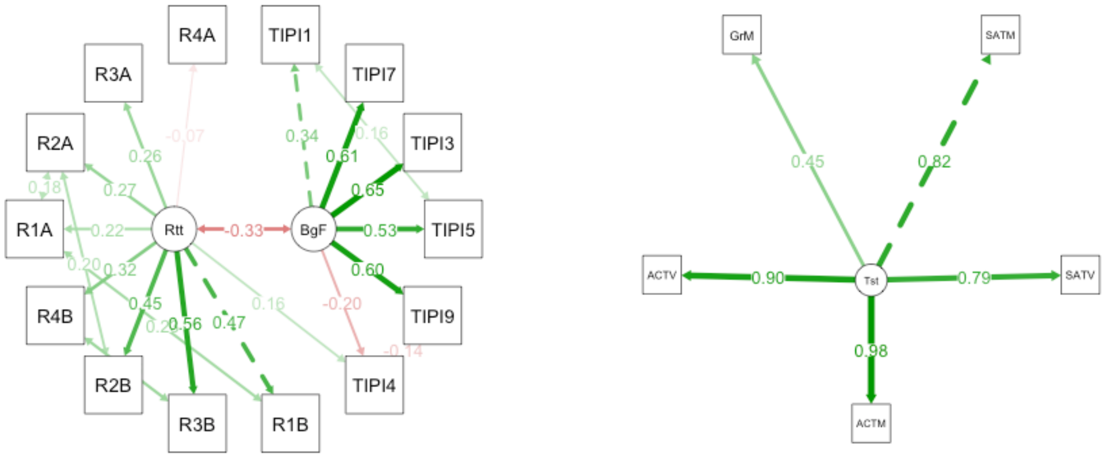
6.2 Part I: Measurement Model

Latent Factor	Indicator	Factor Loadings
Non-Cognitive Skills (θ_t^{PN})		
Big Five ($\theta_t^{PN_b}$)	Extraversion	0.343(-)
	Agreeableness	0.610(0.137)
	Conscientiousness	0.64(0.136)
	Neuroticism	-0.203(0.099)
	Openness	0.530(0.113)
	Neuroticism (Reversed)	0.603(0.131)
Rotter Scale ($\theta_t^{PN_r}$)	Rotter Locus of Control Indicator 1	0.223(0.045)
	Rotter Locus of Control Indicator 2	0.467(-)
	Rotter Locus of Control Indicator 3	0.266(0.058)
	Rotter Locus of Control Indicator 4	0.452(0.072)
	Rotter Locus of Control Indicator 5	0.262(0.035)
	Rotter Locus of Control Indicator 6	0.561(0.095)
	Rotter Locus of Control Indicator 7	-0.065(0.053)
	Rotter Locus of Control Indicator 8	0.316(0.071)
	Neuroticism	0.163(0.240)
Cognitive Skills (θ_t^{PC})		
Test Scores	SAT Mathematics	0.816(-)
	SAT Verbal	0.787(0.076)
	ACT Mathematics	0.984(0.166)
	ACT Verbal	0.900(0.124)
	Highest Grade of Mother	0.446(0.134)

Table 4: Factor Loadings: Mother's Non-Cognitive and Cognitive Skills

A confirmatory factor model was fitted for mother's skills in the measurement model part. Both mother's cognitive and non-cognitive skills are latent variables which needs to be estimated from the observed variables. Table 4 shows the factor loadings along with standard errors in bracket obtained for both cognitive and non-cognitive skills. In the case of non-cognitive skills a two factor model was employed, the first one being Big Five personality traits ($\theta_t^{PN_b}$). The items used to measure the Big Five personality traits are Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness, Neuroticism (Reversed). For Rotter Scale Locus of Control ($\theta_t^{PN_r}$)items such as control over own life, importance of planning, importance of luck, and influence over one's life were considered. The estimation method used was asymptotic distribution free method (ADF) with robust standard errors. Robust standard errors were applied as a scaling correction to the non-normal standard errors. The correction involves applying a scaling constant to the covariance matrix of the parameter estimates and then robust standard errors are obtained by taking the square root of the elements along the diagonal of the covariance matrix (Nevitt & Hancock, 2001).The first factor loading was fixed to 1 as stated in the assumption and hence there is no standard errors available for those indicators. The factor loadings for the $\theta_t^{PN_b}$ range from 0.3 to 0.6 which is on an average scale and are positive except for Neuroticism which is a negative trait and hence its factor loading being negative is justified. In case of cognitive skill (θ_t^{PC}) a one factor model was used with indicators only available for the mother. The indicators considered are the test scores in SAT and ACT both in mathematics and verbal along with the highest grade completed by the mother. The factor loadings range between 0.4 to 0.9 which is quite high.

The model was not kept entirely restricted, it allowed interaction between two factor models of the measurement model in the non-cognitive skills to see if by using modification indices, the model fit increases. Based on the modification indices, interactions of covariances between the observed indicators were taken into account. For instance, the covariance between Neuroticism and Neuroticism (Reversed) was considered due to a possible correlation as they are reveres coded items between them for which the factor loading obtained was -0.141. Reverse coded items can often share excess covariance between them. Covariance between Extraversion and Openness was considered as people who are more extroverted can be more open to new experiences, in which the factor loading was 0.157. In case of Rotter locus of control, covariance was taken between degree of control one has over direction in their life and importance of planning . Both of them can be related with each other as with proper planning, one can attain some direction in their life and the factor loading obtained was 0.180. The factor loadings obtained are quite weak for these covariance terms, but by considering these interactions between the observed indicators, the value of the χ^2 does get reduced, improving the model fit to some extent as a high χ^2 value is an indicator of large discrepancy between implied covariance matrix and sample covariance matrix.



(a) Parental Non-cognitive skills (b) Parental Cognitive skills

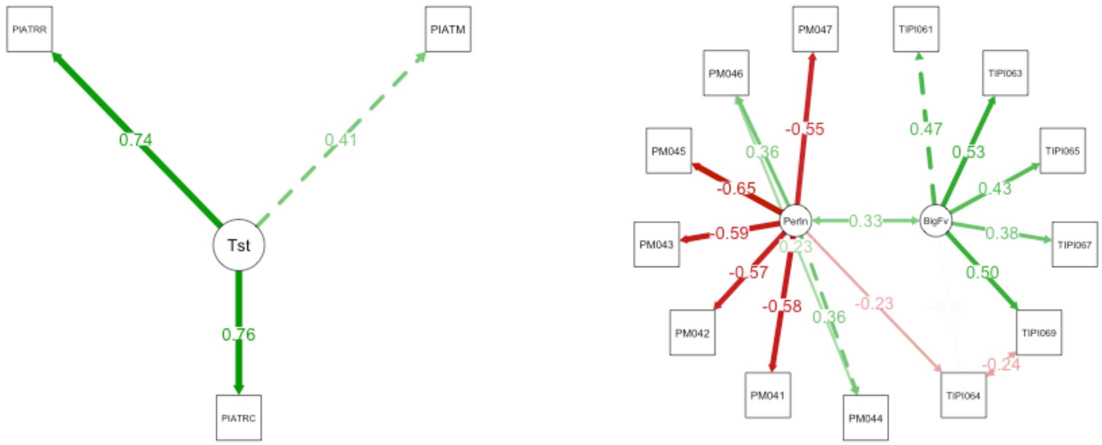
Figure 6: Path diagram for mother's skills

Figure 6 shows the path diagram for parental cognitive and non-cognitive skills. In the path diagram the rectangle boxes represents the observed variables, the indicators for the latent variable while the circles represent the latent variables. The arrows represent the association between observed and latent variables. The dark green arrow means that the factor loading is positive and quite high which means that the observed variable is a good indicator of the latent variable. Figure 6b shows the path diagram for the mothers cognitive skills in which the dark green paths indicate a positive high factor loadings for all the standardised test except for highest grade completed by the mother. Figure 6a shows the path diagram for mothers non-cognitive skills. Most of them have a positive factor loadings and are also above 0.6 except for one item which is the Neuroticism(TIPI9), however the factor loading is low.

Latent Factor	Indicator	Factor Loadings θ_t	Factor Loadings θ_{t+1}
Non-cognitive Skills (θ_t^N)			
Big Five Personality ($\theta_t^{N_b}$)	Extraversion	0.466(-)	0.543(-)
	Conscientiousness	0.532(0.087)	0.513(0.058)
	Neuroticism	-0.014(0.089)	-0.028(0.060)
	Openness	0.433(0.075)	0.410(0.054)
	Agreeableness	0.384(0.080)	0.403(0.056)
	Neuroticism (Reversed)	0.505(0.090)	0.477(0.060)
Pearlin Mastery ($\theta_t^{N_p}$)	Pearlin Mastery Indicator 1	-0.588(0.226)	-0.613(0.108)
	Pearlin Mastery Indicator 2	-0.568 (0.145)	-0.598(0.102)
	Pearlin Mastery Indicator 3	-0.591(0.132)	-0.657(0.102)
	Pearlin Mastery Indicator 4	0.360(-)	0.414(-)
	Pearlin Mastery Indicator 5	-0.648 (0.137)	-0.676(0.103)
	Pearlin Mastery Indicator 6	0.362(0.076)	0.4100.054)
	Pearlin Mastery Indicator 7	-0.549(0.139)	-0.539 (0.093)
	Neuroticism	-0.230(0.244)	-0.207(0.176)
Cognitive Skills (θ_t^C)			
Tests Scores	PIAT Mathematics	0.407(-)	0.418(-)
	PIAT Reading Recognition	0.738(0.091)	0.727(0.194)
	PIAT Reading Comprehension	0.762(0.092)	0.793(0.182)

Table 5: Factor Loadings: Individual's Non-Cognitive and Cognitive Skills

A two factor confirmatory factor model was fitted for individual's non-cognitive skills, the first one being Big Five personality trait ($\theta_t^{N_b}$) which has the same observed indicators as the mother's Big Five personality trait. The second factor is the Pearlin Mastery Scale ($\theta_t^{N_p}$). For the cognitive skills, one factor confirmatory model was fitted whose results are shown in table 5. Again the estimation method used is ADF taking into consideration the non-normality of data. The factor loadings for $\theta_t^{N_b}$ range from 0.3 to 0.5 which is an average range. The factor loading for Neuroticism is quite weak with a loading of -0.014. In case of $\theta_t^{N_p}$, the factor loadings range from -0.6 to 0.3 which is also in an average range. Negative factor loadings can be associated with the fact that items express a negatively associated trait. The factor loadings for the cognitive skills (θ_t^C), range between 0.4 to 0.7 with PIAT Reading Comprehension having the highest factor loading.



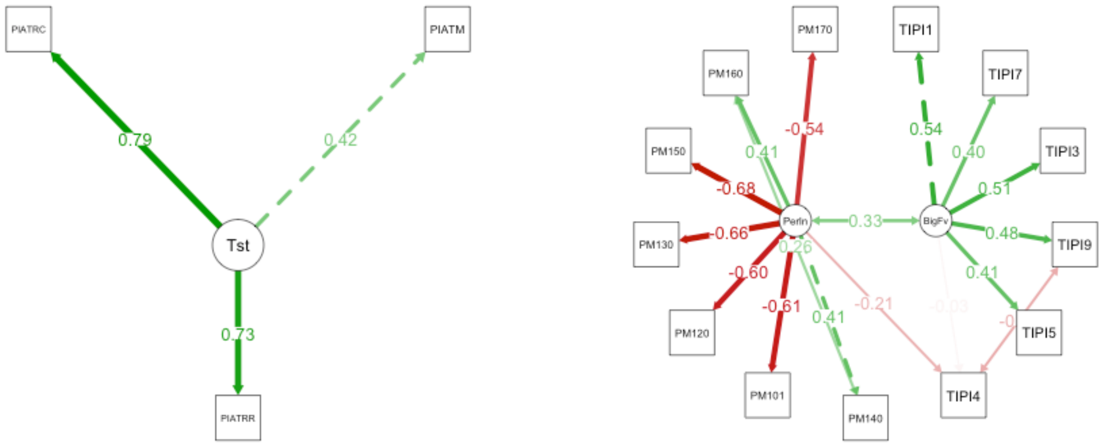
(a) Cognitive skills

(b) Non-Cognitive skills

Figure 7: Individual skills for θ_t

With the help of modification indices, covariance interactions between the observed variables were considered to account for any correlation present among them. Inclusion of covariance between Neuroticism and Neuroticism(Reversed) in $\theta_t^{N_b}$ and $\theta_t^{N_p}$ shows a factor loading of -0.242 while allowing a covariance between Pearlin Mastery observed indicators have a factor loading of 0.226. The covariances between the observed indicators have low factor loadings, despite which the χ^2 value reduces to some extent to improve the model fit.

Figure 7 and 8 show the path diagram of individuals skills both cognitive and non-cognitive. In the path diagram the rectangle boxes represents the observed variables while the circles represent the latent variables. The arrows represent the association between observed and latent variables. The dark green arrow means that the factor loading is positive and quite high and is a good indicator of the latent variable while the dark red arrow means that the factor loading is negative and quite high. From figure 7a and 8a a high factor loading is seen for reading items shown in green. From figure 7 and 8 it can be inferred that, since most of the factor loadings are high which means that the observed variables are good indicators of the respective latent variable.



(a) cognitive skills (b) Non-Cognitive skills

Figure 8: Individual skills θ_{t+1}

One factor confirmatory model was fitted for parental investments whose factor loadings along with standard errors in brackets are shown in table 6. The factor loadings range from 0.4 to 0.6 which is satisfactory, and only with one factor loading being in the range of 0.1. The results of the CFA model show that strong indicators of parental investment is the amount of time spent parents spend with the individual in various forms.

Latent Factor	**Indicator**	**Factor Loadings**
	Parental Investment	
Parental Investment	Relation of Individual with Parents	0.106(-)
	Parents and Individual spend time together (Movies)	0.468(0.721)
	Parents and Individual spend time together (Dinner)	0.646(0.901)
	Parents and Individual spend time together (Shopping)	0.487(0.632)
	Parents and Individual spend time together (Outing)	0.497(0.796)

Table 6: Factor Loadings: Parental Investments

6.2.1 Reliability of factors

Factors	Omega2	Omega3
$\theta_t^{PN_b}$	0.572	0.551
$\theta_t^{PN_r}$	0.268	0.256
θ_t^{PC}	0.898	0.943
θ_t^{C}	0.683	0.431
$\theta_t^{N_b}$	0.531	0.506
$\theta_t^{N_p}$	0.438	0.396
θ_{t+1}^{C}	0.701	0.467
$\theta_{t+1}^{N_b}$	0.537	0.512
$\theta_{t+1}^{N_p}$	0.448	0.415
θ_t^{I}	0.408	0.407

Table 7: Omega reliability for the factors

Table 7 shows the reliability of factors in the measurement model in terms of both omega2 and omega3. Factor reliability given under omega2 consists of the denominator which is the model implied variance of the total score while denominator under omega3 is given by the observed sample variance. The reliability result of θ_t^{PC} represents the proportion of total score variance that is due to a single factor, how well the given indicators measure the latent factor θ_t^{PC}. For the interpretation of non-cognitive skills, which are $\theta_t^{PN_b}$ and $\theta_t^{PN_r}$, the explanation given is proportion of non-cognitive skills total score variance is due to $\theta_t^{PN_b}$ over and above the influence of other factors. With the reliability estimates, it can be seen that $\theta_t^{PN_b}$ is more reliable than $\theta_t^{PN_r}$ in case of the mother. Similar results can be seen for the individual. The bootstrapped confidence intervals estimated for θ_t^{PC} is given as $(0.811, 0.918)$ and the reliability estimate of model implied variance is given in that interval. For θ_t^{I} as well the estimate lies in the confidence interval sampled using the bootstrapped method. However for θ_t^{C} andθ_{t+1}^{C}, the reliability estimates lie outside of the confidence interval.

6.2.2 Goodness-of-fit indices

	χ^2	**CFI**	**GFI**	**RMSEA**
θ_t^{PN}	440.216(0.000)	0.852	0.999	0.040
θ_t^{PC}	9.514(0.090)	0.946	0.968	0.119
θ_t^{C}	247.861(0.000)	0.681	0.941	0.403
θ_{t+1}^{C}	48.469(0.000)	0.745	0.948	0.359
θ_t^{N}	229.240(0.000)	0.890	0.973	0.031
θ_{t+1}^{N}	319.663(0.000)	0.883	0.971	0.034
θ_t^{I}	7.994(0.157)	0.991	0.998	0.021

Table 8: Goodness-of-fit indices indices

Table 8 represents the fit indices for all the CFA models used. By looking at the χ^2 values along with the p-value in brackets for all the models, the null hypothesis is rejected in all except for the models θ_t^{PC} which is the mother's cognitive ability and θ_t^I which is the parental investment where the models are accepted. Despite using the modification indices, the χ^2 values indicated are quite high due to which the models are being rejected. The discrepancy function F between the sample covariance matrix S and the observed covariance matrix Σ is still big which is why the model is being rejected and the goodness-of-fit measures are also below the acceptable threshold limit. It has been noted that χ^2 can be inflated due to violations of multivariate normality and large samples sizes ((Anderson & Gerbing, 1984);(Moshagen, 2012)) and for this other goodness of fit indices can be considered for the model fit. The CFI values of greater than 0.9 indicate a good fit which is only achievable for θ_t^{PC} and θ_t^I. Similar acceptable levels hold for GFI and all of the factors fulfil that criteria.

	χ^2	**CFI**		**GFI**		**RMSEA**	
		Mean CFI	**CFI C.I**	**Mean GFI**	**GFI C.I**	**Mean RMSEA**	**RMSEA C.I**
θ_t^{PN}	109.924(0.000)	0.997(0.003)	(0.987,1.000)	0.999(0.005)	(0.997, 0.999)	0.003(0.004)	(0.000,0.011)
θ_t^{PC}	29.843(0.258)	0.981(0.033)	(0.883,1.000)	0.984(0.108)	(0.956,0.997)	0.051(0.063)	(0.000,0.193)
θ_t^{C}	10.445(0.000)	0.998(0.002)	(0.991,1.000)	0.999(0.0003)	(0.998,0.999)	0.008(0.015)	(0.000, 0.050)
θ_{t+1}^{C}	13.434(0.000)	0.996(0.008)	(0.973, 1.000)	0.999(0.001)	(0.995,0.999)	0.017(0.031)	(0.000, 0.098)
θ_t^{N}	103.885(0.000)	0.996(0.005)	(0.981, 1.000)	0.992(0.001)	(0.989,0.994)	0.0043(0.0045)	(0.000,0.013)
θ_{t+1}^{N}	100.539(0.000)	0.997(0.0034)	(0.988, 1.000)	0.994(0.001)	(0.992, 0.996)	0.0032(0.0036)	(0.000, 0.011)
θ_t^{I}	18.382(0.148)	0.996(0.007)	(0.974, 1.000)	0.999(0.00005)	(0.994,0.999)	0.008(0.014)	(0.000, 0.034)

Table 9: Bootstrapped goodness-of-fit indices

Table 9 shows the Bollen-Stine adjusted bootstrapped results for the goodness-of-fit indices of the measurement model for which the mean, standard deviation(in brackets) and confidence intervals are given. The bootstrapped results have been presented for three goodness-of-fit indices, namely, CFI, GFI and RMSEA. The values obtained from the bootstrap procedure can be used to determine the discrepancy between the goodness-of-fit indices of the actual data and those due to sampling error and hence determine the fit of the model (Bone et al., 1989). The results show a increase in goodness-of-fit as compared to the original values, indicating a presence of sampling error. Additionally, chi-square values were also bootstrapped and the relevant p-values were also calculated. In case of the chi-square values, after applying the Bollen-Stine bootstrapping method, the values did reduce substantially, however we still rejected and accepted the same models.

6.3 Part II: Structural Model

The structural model is a regression between the latent variables obtained from the confirmatory factor model in the measurement part of the LISREL model. For the estimation of the structural model, factor scores obtained from the confirmatory factor model is used. As mentioned in the methodology, the bias is corrected in the covariance matrix of the factor scores using the Croon method, and corrected covariance matrix is used for the estimation of the structural model.

	$\theta_t^{N_b}$	$\theta_t^{N_p}$	θ_t^{I}	$\theta_t^{PN_b}$	$\theta_t^{PN_r}$	θ_t^{PC}	θ_t^{C}
$\theta_{t+1}^{N_b}$	-0.015(0.162)	0.070(0.311)	0.346(0.377)	-0.024(0.235)	-0.535(0.484)	-0.070(0.076)	-0.027(0.030)
$\theta_{t+1}^{N_p}$	0.025(0.065)	-0.158(0.132)	0.119(0.124)	-0.029(0.094)	0.057(0.193)	-0.006(0.030)	-0.007(0.011)
θ_{t+1}^{C}	0.046(0.121)	-0.064(0.295)	-0.929(0.654)	-0.268(0.444)	-1.502(0.927)	0.015(0.143)	0.119(0.151)

Table 10: Parameter Estimates of the structural model

The structural relationship between latent variables is shown in table 10. From the estimated structural regressions, the effects of self-productivity and cross-productivity can be understood. Self-productivity refers to effect of any past period's cognitive/non-cognitive skills on the present period's cognitive/non-cognitive skills, while cross productivity refers to any past period's non-cognitive/cognitive skills on the present period's cognitive/non-cognitive skills. In the case of non-cognitive skills, for $\theta_{t+1}^{N_b}$ the effect of $\theta_t^{N_b}$ is negative with a value of -0.015 indicating a negative self-productivity of skills, while the effect of $\theta_t^{N_p}$ is positive which means that there is a positive cross-productivity between non-cognitive skills. The same results can be seen in the case of $\theta_{t+1}^{N_p}$. The effect of θ_t^{C} on $\theta_{t+1}^{N_p}$ and $\theta_{t+1}^{N_b}$ is negative indicating a negative cross-productivity of cognitive skills on non-cognitive skills. This means that cognitive skills are not helping in improving the non-cognitive skills of the individual. In the case of θ_{t+1}^{C}, there is a positive effect of self-productivity of skills, and a positive effect of $\theta_t^{N_b}$ traits on the present period's cognitive skills, however negative effect of $\theta_t^{N_p}$. The results in the case of intergenerational transmission of skills show that parental investments has an impact in improving the non-cognitive abilities of the individual. Mother's non-cognitive abilities $\theta_t^{PN_r}$ and $\theta_t^{PN_b}$ has a negative effect on the individual's $\theta_t^{N_b}$, while a positive effect on the individual's $\theta_t^{N_p}$. The intergenerational transmission of cognitive skills has a positive effect on the individual's cognitive ability.

6.4 Part III: Estimation of occupational choice

The factors scores the non-cognitive and cognitive skills were anchored along with occupations of mother and father to understand the mechanism behind the choices of the individual's occupation whose results are presented in table 11. The factor scores of the latent variable used were already corrected for bias using Croon method.

Parameter	Estimates	Standard Error
$\theta_{t+1}^{N_b}$	0.069	0.039
$\theta_{t+1}^{N_p}$	-0.132	0.105
θ_{t+1}^{C}	0.015	0.016
Y_m	-0.013	0.024
Y_f	-0.019	0.049

Table 11: Parameter estimates of the impact on occupational choice

The estimates obtained show that occupation choice of an individual is largely based on their Big Five personality traits and their cognitive ability. Pearlin Mastery skills are having a negative effect which is in a good sign since the indicators of this latent variable contain statements which are more negative in nature. Parents occupation have a negative effect on the occupational choice of the individual. The results indicate that intergenerational transmission of skills in the cognitive abilities are more strongly impacting the occupational choice of an individual while the for the non-cognitive skills, the cross-productivity among non-cognitive skills and parental investment are better off.

7 Discussion & Conclusion

This paper seeks to answer the research question whether there is a presence of intergenerational transmission of skills in terms of both cognitive and non-cognitive, and its impact on the occupational choice of an individual. The focus was to understand whether parents skills get transmitted to their offspring and can it help in understanding what factors contribute more in an individual making occupational choices which can help in improving the labor market outcomes. The methodology used for this purpose was Structural Equation Modelling(SEM) and the purpose to use this methodology is that certain variables such as non-cognitive ability are defined as latent variables which are not directly observable and hence they need to be defined from the observed variables. In the measurement model, confirmatory factor analysis was used to understand how strong are the observed indicators are for the latent variables. For the non-cognitive skills, the indicators were average while for the cognitive skills, the indicators were quite high. The average indicators can be explained by the fact that non-cognitive skills are more of personality traits for which a measurement value has not yet been defined. The results in the structural model indicated a positive effect of transmission of cognitive skills from mother to the individual, while for the non-cognitive skills it was negative. Instead for the non-cognitive skills, self-productivity and cross-productivity of skills were more dominant along with parental investments. In the final part where the latent variables were used to predict the factors contributing to the occupational choice of an individual, among the non-cognitive skills, Big Five personality traits stand out followed by the cognitive skills. This implies that the choice made by an individual to large extent influenced by their underlying personality and their cognitive abilities such as test scores, GPA, highest grade completed.

One of the limitation was in the estimation of the confirmatory factor model does however reject the null hypothesis with high χ^2 value which means that the current hypothesized model is wrong. As discussed in the section 6, the rejection of the models can be due to large sample sizes in which χ^2 is sensitive. Despite applying the Bollen-Stine bootstrap procedure, the rejected models were still rejected. For the purpose of the analysis, the model was kept restricted allowing interactions only between the observed indicators as covariances. This could be one of the reasons for the rejection of the models. However, when more interactions were allowed between the latent variables, the factor loadings became weak. There was a tradeoff between the model significance and factor loadings which was consecutively used for further analysis of the model. Moreover, in the case of non-cognitive skills, which are personality traits model significance can be tough to achieve. The models for parental investment and mother's cognitive skills were however accepted which could be due to low sample size present in them.

Another limitation was the availability of the dataset of the father's cognitive and non-cognitive skills which allowed only the estimation of mother's skills. The analysis could only answer the transmission of mother's skills to the individual which showed that mother's cognitive ability positively affects the individual's cognitive ability influencing the occupational choice of the individual. Further, for the estimation of the structural model, instead of using the factor scores, the model could have been estimated implicitly but it did lead to error as the inverse of the covariance matrix was singular. The reliability of the measurement model differs in terms of the model implied variance and observed sample variance. The reliability based on the model implied variance gives a higher reliability factor than the observed sample variance from which it can be inferred that the model can be relied, however, the presence of model misspecification cannot be completely ruled out.

The paper does present econometric challenges in terms of the rejection of models, average factor loadings in the measurement model, possible presence of misspecification in the measurement model. However, given that all of the variables under consideration are latent variables most of which are personality traits since these variables are not of quantifiable nature, econometric challenges are more likely to occur. For further research, the existing model could be expanded more to include more observed variables such as race, ethnicity, income of parents, effect of siblings, and other varied background. This can potentially help in reducing any potential model misspecification as well. In the empirical economics literature non-cognitive skills are often termed as black box which has become research interest of many even in terms of improving econometric estimation techniques for them. Opening up these black boxes, can help in improvement of educational programs, job market preparedness which can overall improve the quality of life of an individual.

References

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, *49*(2), 155–173.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, *103*(3), 411.

Anger, S. (2012). The intergenerational transmission of cognitive and non-cognitive skills during adolescence and young adulthood.

Anger, S., & Heineck, G. (2010). Do smart parents raise smart children? the intergenerational transmission of cognitive abilities. *Journal of population Economics*, *23*(3), 1105–1132.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural equation modeling: a multidisciplinary journal*, *16*(3), 397–438.

Bearden, W. O., Sharma, S., & Teel, J. E. (1982). Sample size effects on chi square and other statistics used in evaluating causal models. *Journal of marketing research*, *19*(4), 425–430.

Becker, G. S. (2009). *Human capital: A theoretical and empirical analysis, with special reference to education.* University of Chicago press.

Benson, J., & Fleishman, J. A. (1994). The robustness of maximum likelihood and distribution-free estimators to non-normality in confirmatory factor analysis. *Quality and Quantity*, *28*(2), 117–136.

Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. *Annual review of psychology*, *47*(1), 563–592.

Bollen, K. A. (1996). An alternative two stage least squares (2sls) estimator for latent variable equations. *Psychometrika*, *61*(1), 109–121.

Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research*, *21*(2), 205–229.

Bone, P. F., Sharma, S., & Shimp, T. A. (1989). A bootstrap procedure for evaluating goodness-of-fit indices of structural equation and confirmatory factor models. *Journal of Marketing Research*, *26*(1), 105–111.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. *Handbook of structural equation modeling*, 361–379.

Browne, M. W. (1982). Covariance structures. *Topics in applied multivariate analysis*, 72–141.

Brunello, G., & Schlotter, M. (2011). Non-cognitive skills and personality traits: Labour market relevance and their development in education & training systems.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological methods & research*, *5*(1), 3–52.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of econometrics*, *18*(1), 5–46.

Chevalier, A., et al. (2002). Just like daddy: the occupational choice of uk graduates. In *Royal economic society annual conference 2002.*

Chin, W. W. (1998). *Commentary: Issues and opinion on structural equation modeling.* JSTOR.

Cobb-Clark, D. A., & Tan, M. (2011). Noncognitive skills, occupational attainment, and relative wages. *Labour Economics*, *18*(1), 1–13.

Constant, A. F., & Zimmermann, K. F. (2003). Occupational choice across generations. *Available at SSRN 487416*.

Croon, M. (2002). Using predicted latent scores in general latent structure models. *Latent Variable and Latent Structure Models*, 195–224.

Cubel, M., Nuevo-Chiquero, A., Sanchez-Pages, S., & Vidal-Fernandez, M. (2016). Do personality traits affect productivity? evidence from the laboratory. *The Economic Journal*, *126*(592), 654–681.

Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of human resources*, *43*(4), 738–782.

de Coulon, A., Meschi, E., & Vignoles, A. (2011). Parents' skills and children's cognitive and non-cognitive outcomes. *Education economics*, *19*(5), 451–474.

Devlieger, I., & Rosseel, Y. (2017). Factor score path analysis. *Methodology*.

Ham, R., Junankar, P. N., & Wells, R. (2009a). Antagonistic managers, careless workers and extraverted salespeople: an examination of personality in occupational choice.

Ham, R., Junankar, P. N., & Wells, R. (2009b). *Occupational choice: Personality matters* (Tech. Rep.). IZA discussion papers.

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2003). *Fifty years of mincer earnings regressions* (Tech. Rep.). National Bureau of Economic Research.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the ged testing program. *American Economic Review*, *91*(2), 145–149.

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

Joereskog, K., & Sörbom, D. (1984). *Analysis of linear structural relations by maximum likelihood, instrumental variables, and least squares methods, users guide.* Uppsala:

University of Uppsala.

John, K., & Thomsen, S. L. (2014). Heterogeneous returns to personality: the role of occupational choice. *Empirical Economics*, *47*(2), 553–592.

Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, *43*(4), 443–477.

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *journal of the American Statistical Association*, *70*(351a), 631–639.

Krishnakumar, J., & Nagar, A. L. (2008). On exact statistical properties of multidimensional indices based on principal components, factor analysis, mimic and structural equation models. *Social Indicators Research*, *86*(3), 481–496.

Kuhn, P., & Weinberger, C. (2005). Leadership skills and wages. *Journal of Labor Economics*, *23*(3), 395–436.

Lei, P.-W., & Wu, Q. (2007). Introduction to structural equation modeling: Issues and practical considerations. *Educational Measurement: issues and practice*, *26*(3), 33–43.

Long, S. J. (1983). *Covariance structure models* (No. 04; e-book.). Sage Beverly Hills.

McDonald, R. P. (1999). Test theory: A unified approach.

Molenaar, P. C., & Nesselroade, J. R. (1998). A comparison of pseudo-maximum likelihood and asymptotically distribution-free dynamic factor analysis parameter estimation in fitting covariance-structure models to block-toeplitz matrices representing single-subject multivariate time-series. *Multivariate Behavioral Research*, *33*(3), 313–342.

Moshagen, M. (2012). The model size effect in sem: Inflated goodness-of-fit statistics are due to the size of the covariance matrix. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(1), 86–98.

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural equation modeling*, *8*(3), 353–377.

Nieken, P., & Störmer, S. (2010). Personality as predictor of occupational choice: Empirical evidence from germany. *Diskussionspapiere des Schwerpunktes Unternehmensführung am Fachbereich BWL*, *2*(8).

Phakiti, A. (2018). Confirmatory factor analysis and structural equation modeling. In *The palgrave handbook of applied linguistics research methodology* (pp. 459–500). Springer.

Raines-Eudy, R. (2000). Using structural equation modeling to test for differential reliability and validity: An empirical demonstration. *Structural equation modeling*, *7*(1), 124–141.

Raykov, T. (1998). A method for obtaining standard errors and confidence intervals of composite reliability for congeneric items. *Applied Psychological Measurement*, *22*(4), 369–374.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of educational research*, *99*(6), 323–338.

Schumacker, E., & Lomax, G. (2016). *A beginner's guide to structural equation modeling. 4th edtn.* New York, NY: Routledge.

Shanmugam, V., & Marsh, J. (2015). Application of structural equation modeling to the social sciences: a brief guide for researchers. *Mesure et évaluation en éducation*, *37*(3), 99–123.

Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, *66*(4), 563–575.

Todd, P. E., & Zhang, W. (2020). A dynamic model of personality, schooling, and occupational choice. *Quantitative Economics*, *11*(1), 231–275.

Tsai, J. (2006). Decoding the ged signal: Labor market returns to non-cognitive ability (graduate seminar draft).

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Annals of Psychology*, *33*(3), 755–782.

Weston, R., & Gore Jr, P. A. (2006). A brief guide to structural equation modeling. *The counseling psychologist*, *34*(5), 719–751.

World Economic Forum. (2021). Building a common language for skills at work a global taxonomy.

Yuan, K.-H., & Bentler, P. M. (1997). Mean and covariance structure analysis: Theoretical and practical improvements. *Journal of the American statistical association*, *92*(438), 767–774.

8 Appendix

The list of Pearlin Mastery Indicators:

- Pearlin Mastery 1: No way I can solve the problems I have
- Pearlin Mastery 2: I sometimes feel that I am being pushed around
- Pearlin Mastery 3: I sometimes feel that I am being pushed around
- Pearlin Mastery 4: I can do just about anything that happens in my life
- Pearlin Mastery 5: I often feel helpless in dealing problems in my life
- Pearlin Mastery 6: What happens to me in future mostly depends on me
- Pearlin Mastery 7: Little I can do to change things in my life

The Rotter Locus of Control indicators is as follows:

- Rotter Locus of Control Indicator 1: Degree of control individual has over direction of own life
- Rotter Locus of Control Indicator 2: Is the above statement much or slightly closer to individual's opinion?
- Rotter Locus of Control Indicator 3: Importance of Planning
- Rotter Locus of Control Indicator 4: Is the above statement much or slightly closer to individual's opinion?
- Rotter Locus of Control Indicator 5: Importance of Luck
- Rotter Locus of Control Indicator 6: Is the above statement much or slightly closer to individual's opinion?
- Rotter Locus of Control Indicator 7: Degree of Influence R has over own life
- Rotter Locus of Control Indicator 8: Is the above statement much or slightly closer to individual's opinion?

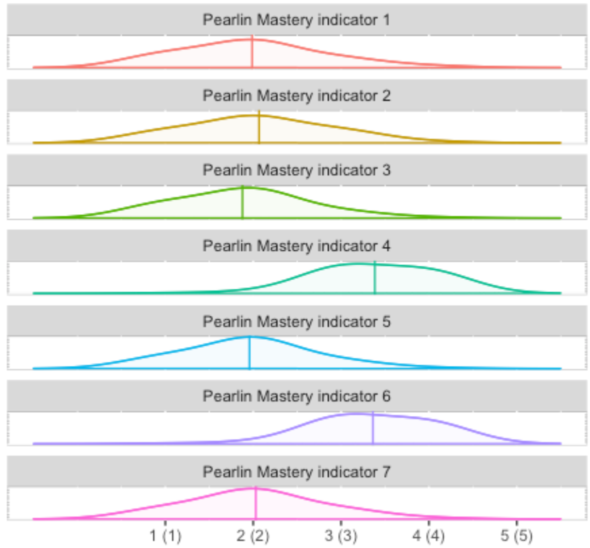


Figure 9: Density Plot of Observed Indicators of Pearlin Mastery

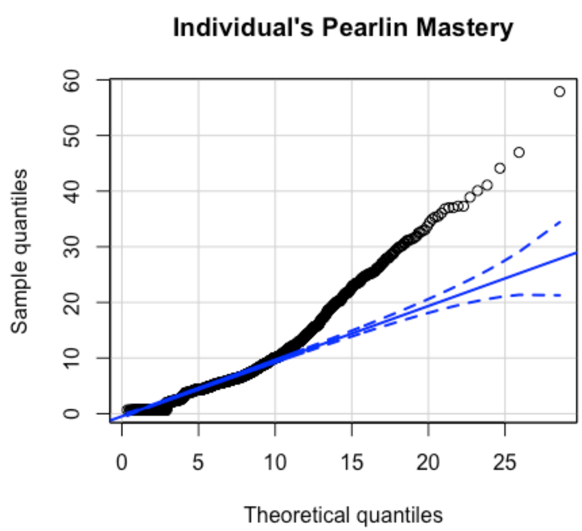


Figure 10: Chi-Square Q-Q plots for observed indicators of Pearlin Mastery