

The relationship between conscientiousness and income

Rajius Idzalika¹

April 2014

Abstract

We provide the initial evidence on the relationship between conscientiousness and income for Indonesia. Using Indonesian Family Life Survey in 2000, we utilize respondent's item response rate which is the facets of conscientiousness of Five Factor Model. Our findings suggest that type of primary activity is important determinant of how many questions answered in the survey and self-employed individuals answer more questions than workers. The estimation of item response rate for earnings yields positive association together with other traditional predictors of earnings for workers, but not for self-employed people.

Keywords: non-cognitive abilities, Five Factor Model, conscientiousness, item response rate, wages.

JEL classification: J30, C46, Z00

1. Introduction

Non-cognitive ability has been long recognized as the potential factor that contributes to individual and society performance in economics besides cognitive ability and traditional human capital predictors such as education. Bowles & Gintis (1976) underlined that attitudes, motivation and personality traits are more important for labor market performance than academic achievement. Similar to that, according to Heckman & Rubinstein (2001) persistence, reliability and self-discipline are among

¹ Georg-August Universität Göttingen. Platz der Göttinger Sieben 3, Oeconomicum Room Oec 2.122
37073 Göttingen. Tel. +49 (0) 551 39 7343. Email: rajius.idzalika@wiwi.uni-goettingen.de.

personality traits that define success more than IQ. However, the study of the impact of non-cognitive ability on wealth is less rich than the ones for cognitive-ability due to the challenge to establish a proxy for non-cognitive ability.

A recent literature about survey based measurement of conscientiousness, one of five widely known personality traits, shed a light that we are able to use administered survey to enrich the literature of non-cognitive ability in the economic perspective. Our contribution in this paper is to employ item response rate of the detailed Indonesian Family Life Survey (IFLS) as the more objective and inexpensive proxy of non-cognitive ability and its association with wealth. The paper unfolds as follows. Section 2 discusses the theoretical background of non-cognitive ability with most weight is for conscientiousness and evidence from psychological economics literature. Item response rate takes place at section 3. Results are presented in section 4. We give concluding remarks in section 5.

2. Non-cognitive ability and earnings

The term non-cognitive ability is generally used to distinguish individual traits and behavior which are not the part of academic skills. Because of this broad concept, non-cognitive skills consist of a large body of what we observe from an individual such as motivation, perseverance and self-confidence. In psychology, those traits are grouped into five comprehensive but non-overlap factors called as Five Factor Model (5FM). They are Neuroticism, Extraversion, Conscientiousness, Agreeableness and Openness to experience (Digman, 1990).

Extraversion is a trait that related to sociability such as venturesomeness, affiliation, positive affectivity, energy, ascendance, and ambition. Agreeableness depicts the dimension of humanity including altruism, nurturance, caring and emotional support. Conscientiousness is generally described as strong-willed and the

people who carry those traits are usually thorough, neat, well-organized, diligent and achievement-oriented. Openness to experience includes aspect of intellect in the broader scope such as high intellectual ability, enjoying aesthetic impressions, has wide interests, and unconventional thought (McCrae & John 1992). Five Factor Model does not mean to represent the whole personality traits, yet it serves as the major categorization and is widely acceptable in psychology and other fields. This model also gets more attention in human capital research such as in Almlund et al. (2011).

The measurement of non-cognitive ability has been so far focusing on developing specific scale for the skill of interest in self-reported questionnaires. To name a few, self-concept is measured by survey instrument such as Self-Description Questionnaire (SDQ) (Marsh, 1990; 1992). Motivation as the representation of goal orientation is measured by The Patterns of Adaptive Learning Scale (PALS) for students and teachers (Midgley et al, 2000). Self-control, associated by conscientiousness, is measured by Self-Control Scale (Tangney, Baumeister, & Boone, 2004). It is questionable if ones can construct a single component of non-cognitive skills (Brunello & Schlotter, 2011).

Basically, economists are quite reluctant to use subjective data when they are dealing with personality traits due to measurement error and being unfamiliar with the scale in psychological questionnaires (Nyhus & Pons 2005). Nevertheless, measurement error is able to be (partially) corrected using Cronbach's alpha reliabilities (Cronbach, 1951).

Although there is a scarcity of the more objective personal traits measurement, economists themselves are not virtually on the zero ground. Blanden, Gregg & Macmillan (2006) approach non-cognitive ability to account for intergenerational income persistence by mediating factors: cognitive test scores, educational

performance and early labor market attachment. Meanwhile, Brunello & Schlotter (2011) argue that in the absence of performance based incentive, good scores in test tend to represent effort and motivation more than the level of cognitive skills.

More researches that relate personality traits to earnings using the variety of measurement are Groves (2005), Nyhus and Pons (2005), Mueller and Plug (2006), Semykina & Linz (2007) and Heineck & Anger (2010). However, we always need to keep in mind that as Carneiro & Heckman (2004) noticed, most of measurements of non-cognitive abilities on earnings are self-reported ex-post assessments. Also, its direction to labor market outcomes is not well known to be the causes or the consequences.

Among the five factors, Conscientiousness and Openness to experience are considered to have the most relevant link with educational achievement (Digman, 1990). Judge et al. (1999) draw conclusion from the organizational psychology literature that conscientiousness, extraversion, and neuroticism affect career success most. Jencks (1979) argued that industriousness, perseverance and leadership have independent impact from socio-economic background of the family, cognitive ability and years of schooling on earnings. Industriousness and perseverance are the facets of conscientiousness.

Conscientiousness by itself is a valid predictor of job performance (Heineck & Anger 2010; Bowles, Gintis, & Osborne, 2001b). Moreover, Barrick and Mount (1991) as well as Salgado (1997) demonstrated that conscientiousness is positively associated with job performance which occurred across sectors. Costa et al. (1991) argued it is because this trait is related to self-control, persistence, hard work, careful, organized and neat. When the population are divided by gender, evidence from Dutch DNB Household Survey (DHS) suggest that conscientiousness benefits men at the

beginning of employment relationship and openness to experience is more important at the later stage; agreeableness is associated with lower wages for women (Nyhus & Pons 2005). On the other hand, limited evidence of highly educated Wisconsin white male and female from Mueller & Plug (2006) suggested that among men, antagonism as the other side of agreeableness, emotional stability as the other side of neuroticism and openness to experience matter most for earnings while the most important traits related to earnings for women are conscientiousness and openness to experience. Finally, Bouchard and Loehlin (2001) claimed that agreeableness and neuroticism consistently appear to play role in the largest gender differences. With somewhat conflicting evidence for gender differences, the strong support from theoretical perspective and evidence in general put researches to argue that the impact of personality should have a more serious place in economics (Borghans et al, 2008).

Besides its relations to job performance, conscientiousness also play important role in academic success from primary school to college (Bowen, Chingos & McPherson 2009; Nofle and Robbins, 2007; Poropat, 2009). Thus the association of conscientiousness and earnings has two possible paths, directly and indirectly through education.

The story to find a good measurement for conscientiousness itself typically involved self-reported questionnaires such as Ten Item Personality Inventory (TIPI) and Grit Scale (Duckworth and Quinn, 2009) or behavioral checklist like Behavioral Indicators of Conscientious (Jackson et al. 2010) for a more precise proxy. However, researchers are supposed to be aware that self-reported item is a weak proxy while behavioral checklist is tediously long and not appropriate for most surveys.

For our study we propose a recent, more objective task based proxy which is item response rate in a survey. This approach is introduced by Hedegreen & Stratmann (2012) who claim that item response rate is a function of cognitive ability (i.e. IQ) and

non-cognitive ability (i.e. conscientiousness). They observed that when respondents forget or do not want to answer items in questionnaire, it gives a clue about who they are. This is because survey items are typically not cognitively challenging, long and boring as well as give low incentive to finish. To complete them, ones need persistence and attention span which are the facets of conscientiousness. If they actually know the answers but simply ignore them or left it unanswered, that is a sign of losing interest or effort. Therefore, item response rate arguably represents effort or conscientiousness, besides cognitive ability required to complete the survey.

Hedegreen & Stratmann (2012) find that firstly the correlation between item response rate and income is positive. The only identified drawback of this approach is if person with higher income refused to answer many questions. The plausible interpretation of the estimation would be the lower bound of item response rate's effect on earnings. Item response rate captures a fraction of facet of conscientiousness, not all of them (Hill & Trivitt 2013) might help to explain why this does not work for everybody. Secondly, item response rate is associated with conscientiousness after controlling for cognitive ability. Since respondents only need the minimum level of cognitive ability to complete a survey, we conclude that item response rate captures conscientiousness while at the same time already controlling for the same level of cognitive ability.

3. Item response rate

To measure conscientiousness based on administered survey, one simply calculates the item response rate. Item response rate is the fraction of questions that the respondents fill up. The opposite is item nonresponse rate, means that the fraction of questions left unanswered for the variety of reason such as do not know, forget or unwilling. Surveys typically allow respondents to do that, sometimes with specific

code for each reason. This is to distinguish them with missing values due to the skipping pattern.

In IFLS, there are two types of question. The first one, respondents are given options and they only need to choose one, or sometimes more, relevant circumstances. The second type is open question. Respondents need to write something such as the salary for the past month or the amount of working hours last week. For this kind of question, IFLS database recorded them twice. One is if the respondents provide the answer or not. The next is the answer written by the respondents. By design, this arrangement gives us benefit to establish more weight for the second type of question since this one needs more effort than just choose the answer provided in the questionnaire.

As the fraction of the whole questions, item response rate is bounded. This opens an alternative to model the response variable by assuming beta distribution besides the traditional assumption of normal distribution. Regarding the shape of distribution, Budria & Ferrer-i-Carbonell (2012) found that for conscientiousness is skewed to the left.

Due to the long list and complicated skipping pattern, we provisionally select some sections in IFLS 2000 after carefully considering the tradeoff between the number of question included in the calculation of item response rate and the attrition rate. The sections we chose are Subjective Wellbeing, Migration and Employment to control the number of respondents. Therefore, we assume that given the other sections occupy the same respondents as in the sections we selected, the distribution of item response rate does not extremely change.

To measure the association between conscientiousness and income, we consider earnings or salary yield from people who work in private and government office and those who are register themselves are self-employed. Typically self-

employed people provide service across sectors or having farms. There is a possible interdependency between earnings and personality. Psychological literatures suggest that our personalities are inherited but only partially (Jang et al. 1996) and become stable by the age of thirty (James, 1890, pp. 125-126; see also McCrae and Costa, 1990, 1994, 1996, 2003; Costa, McCrae and Siegler, 1999; Cobb-Clark and Schurer, 2012). This traditional view has been challenged by arguing that personality traits has life cycle and is not going to stable before fifty (Roberts & DelVecchio, 2000) which also Srivastava et al. (2003) and Borghans et al. (2008) considered. To control the possible interdependency between earnings and item response rate, we use instrumental variable method. The next possible way to avoid reverse causality is to establish causal model by involving IFLS 2007 and carry out panel analysis. At this point, however, we only do cross-section analysis.

To scrutiny the wage of individual's conscientiousness we used augmented Mincerian earnings equation. The model is:

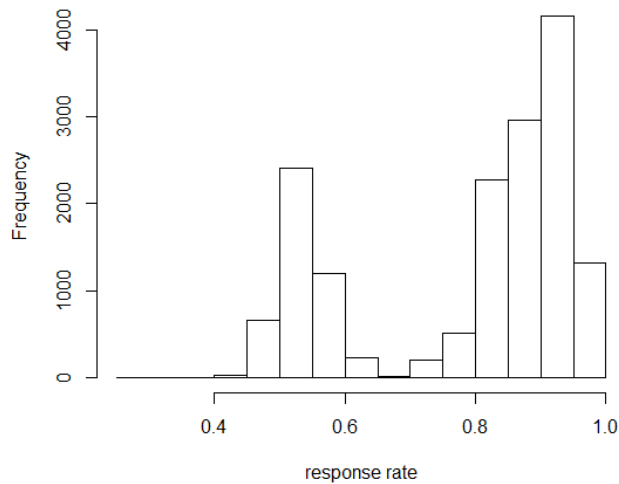
$$y_i = \beta_1 x_i + \beta_2 c_i + u_i \tag{1}$$

y_i is log hourly wage for workers or log hourly profit for self-employed. x_i is vector of covariates including years of education, c_i is individual's item response rate which represents conscientiousness given the similar level of cognitive ability and u_i denotes the idiosyncratic error term. We use Heckman's correction procedure to correct selection bias for respondents who earn money (Heckman, 1979). To accommodate the gender differences in traditional wage literatures and personalities (e.g. Filer, 1986; Osborne, 2000), we add gender dummy in covariates. We also include regional dummies, working field dummies, ethnicity dummies and main language dummies to control working environment, culture and demographic.

4. Results

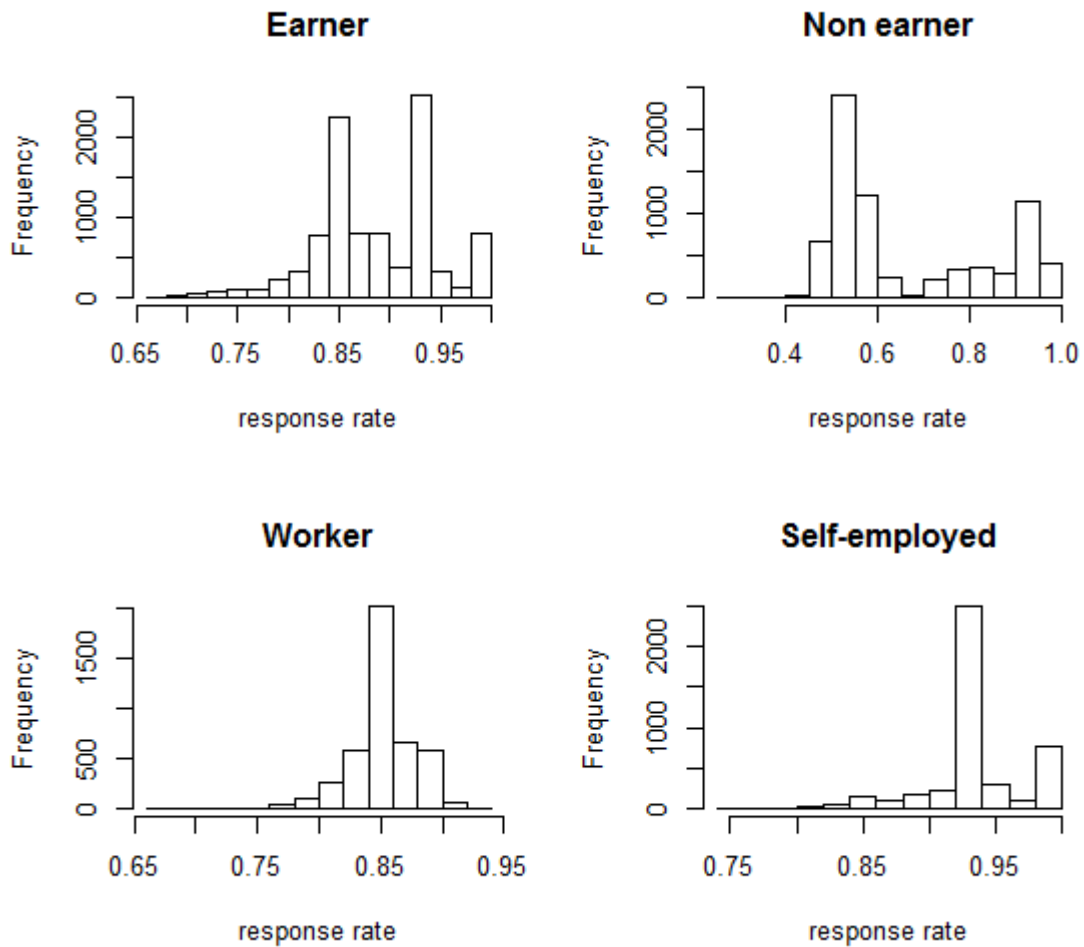
The distribution of item response rate is relatively left skewed and the overall mean is 0.8984. From visual inspection, we found two spikes which leads to the possibility of two subpopulations.

Figure 1a. The distribution of item response rate all samples



A closer look at Figure 1b demonstrates that lower spikes comes from non-earner of income respondents while higher spike represents earner respondents. Furthermore, earner respondents consist of worker which has lower mode and self-employed individual which has higher mode. Summary statistics for samples are available in Appendix 1.

Figure 1b. The distribution of item response rate for sub populations



Since we want to establish association between item response rate and income, factors such as demographic, cultural and working environment are not included in the equation to explain item response rate. Therefore we only establish the relationship between response rate and the type of primary activity. We model the response rate following two distribution assumptions: normal distribution and beta distribution. For regression with beta distribution, there are two models. Each of them has different assumption about response rate variance: homocedasticity and heterocedasticity.

Diagnostic residuals of OLS indicate some outliers and non normality as the backup evidence to model the alternative assumption of beta distribution. Moving to

beta regression, diagnostic residuals assuming heteroscedasticity exhibit a little model improvement from the assumption of homoscedasticity (see Appendix 2).

Table 1. Factors explaining item response rate

Item response rate	OLS (1)		OLS (2)		Beta regression			
	Coefficient	Clust std.error	Coefficient	Clust std.error	Homoscedastic		Heteroscedastic	
					Coefficient	Std. Error	Coefficient	Std. Error
Mu equation								
Primary activity (earner)								
- Student/just graduated	-0.2563***	0.005			-1.2606***	0.0159	-1.4639***	0.0176
- Housekeeping	-0.2657***	0.0031			-1.3356***	0.0117	-1.5078***	0.0125
- Others (retired, sick etc)	-0.2655***	0.0067			-1.3141***	0.0212	-1.4993***	0.0256
Working age (D)	.0455***	0.0033	0.003**	0.0013	0.2352***	0.0138	0.1848***	0.0134
Worker (D)			-0.8692***	0.0008				
Sigma equation								
Primary activity (earner)								
- Student/just graduated							0.97***	0.0086
- Housekeeping							0.9116***	0.0209
- Others (retired, sick etc)							0.9306***	0.0339
Observations	13541		7984		13541		13541	
Adjusted R-squared	0.6444		0.6581					
Root MSE	0.102		0.0314					
Global deviance					-24296		-28463	
AIC					-24280		-28441	
BIC					-24220		-28359	

Note: Significance level are 0.01(***), 0.05(**) and 0.1(*). Models used longitudinal weight without survey design. Estimation of beta regression used the logit link function. To calculate standard error for beta regression, variance covariance matrix qr instead vcov used.

We found that every model agrees on the significance of primary activity and if the respondent is within the interval working age that we defined as between 20 to 65 years old. Together with working age dummy, the type of primary activity explains the variation of response rate up to 0.6444. Considering only the earners in the equation which we divided into worker and self-employed individual, the adjusted R-squared is even slightly higher, 0.6581. Between the types of earner, self-employed people on the average scored eight percent more in responding the questions compared to workers. Thus it is possible to consider that psychological side contributes to determine whether one chooses to work or being self-employed since we found that self employed people are more persistent and more attentive than those who are workers.

When we assume heteroscedasticity by modeling type of primary activity as explanatory variables of sigma equation in beta regression, we found that earner type has smaller response rate's variance compared to those of other types.

Table 2. Regression model of income²

Overall sample	OLS		Heckman		IV	
	Coefficient	Clust. Std. Error	Coefficient	Clust. Std. Error	Coefficient	Clust. Std. Error
Log income per hour						
Response rate	0.8577***	0.2753	0.5671**	0.2261	0.9998***	0.3623
Year of schooling	0.0928***	0.0043	0.0933***	0.0034	0.0919***	0.0043
Age	0.0588***	0.0075	0.0646***	0.0058	0.0604***	0.0078
Age squared	-0.0006***	0.0001	-0.0006***	0.0001	-0.0006***	0.0001
Female (D)	-0.2863***	0.0380	-0.2625***	0.0339	-0.295***	0.0388
Married (D)	0.0936**	0.0420	0.0829**	0.0355	0.0894**	0.0429
Household head (D)	0.1511***	0.0406	0.1194***	0.0339	0.1459***	0.0410
House status (self-owned)						
- Occupying	-0.1539***	0.0429	-0.1166***	0.0350	-0.1230***	0.0422
- Rented/contracted	-0.0618	0.0494	-0.0666	0.0421	-0.0450	0.0505
- Other	0.1606	0.4368	0.5745	0.5416	0.1351	0.4211
Select						
Household head			0.5424***	0.0292		
Married			0.3171***	0.0246		
Age			0.0043***	0.0009		
Year of schooling			0.0234***	0.0030		
Sex			-0.1996***	0.0130		
Observation	7378		15883 (uncensored: 7799)		7707	
Adjusted R-squared	0.1818				0.1919	
Root MSE	1.496				1.205	
Log pseudolikelihood			-21741			

Note: Significance level are 0.01(***), 0.05(**) and 0.1(*).

Regarding income equation, the result from OLS, Heckman selection model and two least square instrumental variable (IV) regression suggest that controlling for traditional predictors³, item response rate is highly significant in explaining the variance of hourly income. This finding is consistent with psychology economics literature that connects personality trait and wealth.

² Item response rate is instrumented using the fields of work.

³ Including ethnicity, working field, region and main language. Complete coefficients see Appendix 3

The estimation of overall sample could also be thought as depiction of income variation due to being worker or being self-employed which are naturally different, since this factor contributes significantly to explain the variation in response rate. We carry out separated regression for each earner type to examine whether response rate explain wages for workers the same as it explains profit for self-employed people. Our finding suggests that response rate is significant predictor of earnings for workers, but not for self-employed (see Appendix 4). Therefore we argue that persistence and attention span are important determinant for wages. Profit makers, on the other hand, seem to require either more types or different types of non-cognitive ability to explain profit variation even though they have higher response rate compared to workers.

5. Concluding remarks

Item response rate, a recent task based measurement of conscientiousness, provides better opportunity to examine to what extent psychological side determines income. Since this approach is relatively more objective, we expect to avoid measurement error given there is no incentive for the respondents to exaggerate the result.

Our findings are as follow:

1. Primary activity is important in determining item response rate in a survey. Those who work and earn money are more persistent and pay more attention to the questionnaire. Moreover, self-employed are better in responding survey than workers.
2. Using two different methods to deal with selection bias and endogeneity issue, we found consistent evidence that item response rate is a strong predictor of earnings. This finding applies only for workers, though. Additionally, the estimation needs careful interpretation due to endogeneity and selection bias along with the

interpretation that the coefficient of item response rate is the lower bound effect on earnings. Nevertheless, the findings shed a light that personality traits matter for economics success and should be given more place in any economic research of wealth.

Finding good and inexpensive measurement of personality traits is currently the main task. While item response rate could capture conscientiousness, it actually represents only some of its facet. Hill & Trivitt (2013) speculated that another proxy of conscientiousness, coding speed test, captures the different facets which are decisiveness and mindfulness. If both approaches can be combined, it might be a stronger representation of conscientiousness.

Considering the importance of non-cognitive ability for individual and aggregate economic performance, it is important for government policies to include this aspect into education system more seriously. Particularly those are for the worker training programs or for people with lower cognitive ability.

References

- Almlund M., Duckworth A.L., Heckman J.J. & Kautz T. (2011). Personality Psychology and Economics. *IZA Discussion Papers*, 5500.
- Blanden J., Gregg, P. & Macmillan, L. (2007). Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education. *Economic Journal*, Royal Economic Society, 117(519), C43-C60, 03.
- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1–26.
- Borghans, L., Duckworth A.L., Heckman, J.J., Weel, B. ter. (2008). The Economics and Psychology of Personality Traits. *Journal of Human Resources*, 43(4), 972-1059.
- Bouchard, T.J., Jr. & Loehlin, J.C. (2001). Genes, Evolution, and Personality. *Behavioral Genetics*. 31(3), 243–73.
- Bowen, W.G., Chingos, M.M & McPherson, M.S. (2009). *Crossing the Finish Line*. Princeton, Princeton Univ. Press.
- Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America* (Vol. 75). New York, Basic Books.
- Bowles, S., Gintis, H., & Osborne, M. (2001a). Incentive-enhancing preferences: Personality behavior and earning. *American Economic Review*, 91, 155–158.
- Bowles, S., Gintis, H., & Osborne, M. (2001b). The determinants of earnings: A behavioral approach. *Journal of Economic Literature*, 39, 1137–1176.
- Brunello, G. & Schlotter, M. (2011). Non Cognitive Skills and Personality Traits: Labour Market Relevance and their Development in Education & Training Systems, *IZA Discussion Papers*, 5743.
- Budria, S. & Ferrer-i-Carbonell, A. (2012). Income Comparisons and Non-cognitive Skills, *SOEPpapers on Multidisciplinary Panel Data Research*, 441.
- Carneiro, P. & Heckman, J.J. (2004). Human Capital Policy In: James J. Heckman, Alan B. Krueger, and Benjamin M. Friedman, eds., *Inequality in America: What Role for Human Capital Policy?* Cambridge, MIT Press.
- Cobb-Clark, D.A. & Schurer, S. (2012). The stability of big-five personality traits, *Economics Letters*, Elsevier, 115(1), 11-15.
- Costa, P.T.Jr. & McCrae, R.R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology*, 54, 853-863.

- Costa, P.T.Jr. & McCrae, R.R.(1990). *Personality in adulthood*. New York, Guilford Press.
- Costa, P.T.Jr., McCrae, R.R. & Dye, D.A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences*, 12(8), 887-898.
- Costa, P.T.Jr. & McCrae, R.R. (1994). Set like plaster: Evidence for the stability of adult personality. In: Heatherton TF, Weinberger JL (Eds), *Can personality change?* Washington, American Psychological Association.
- Costa, P.T.Jr., McCrae, R.R. & Siegler, I.C. (1999). Continuity and change over the adult life cycle: Personality and personality disorders. In: Cloninger CR, editor. *Personality and psychopathology*, Washington, American Psychiatric Press.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*. 16, 297-334.
- Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41,417–440
- Duckworth, A.L & Quinn, P.D. (2009). Development and Validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, 91(2), 166-174.
- Filer, R. K. (1986). The role of personality and tastes in determining occupational structure. *Industrial and Labor Relations Review*, 39, 412–424.
- Groves, M.O. (2005). How important is your personality? Labor market returns to personality for women in the US and UK. *Journal of Economic Psychology*, 26, 827-841
- Hedengren, D. & Strattman, T. (2012). The Dog that Didn't Bark: What Item Nonresponse Shows about Cognitive and Non-Cognitive Ability. *Unpublished Manuscript*. Available at <http://ssrn.com/abstract=2194373> or <http://dx.doi.org/10.2139/ssrn.2194373>.
- Heckman, J.J. (1979). Sample Bias as a Specification Error. *Econometrica*, 47, 153-161.
- Heckman, J.J. & Rubinstein, Y. (2001). The importance of non-cognitive skills. *The American Economic Review*, 91, 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.
- Heineck, G. & S. Anger (2010). The returns to cognitive abilities and personality traits in Germany. *Labour Economics*, 17(3), 535-546.
- Jackson, J.J., Wood, D., Bogg, T., Walton, K.E., Harms, P.D. & Roberts, B.W. (2010). What do conscientious people do? Development and validation of the Behavioral Indicators of Conscientiousness (BIC). *Journal of research in personality*, 44 (4), 501-511.
- Jang KL, Livesly WJ, Vernon PA. (1996). Heritability of the Big Five personality dimensions and their facets: A twin study. *Journal of Personality*, 64, 577-591.

- James, W. (1890). *The principles of psychology, Vol. 1*. Cambridge, Harvard University Press.
- Jenks, C. (1979). *Who gets ahead?* New York. Basic Books.
- Judge, T.A., Higgins, C.A., Thoresen, C.J. & Barrick, M.R. (1999). The Big Five Personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52, 621-652.
- Marsh, H.W. (1990). *SDQ manual : self-description questionnaire*. Campbelltown, University of Western Sydney.
- Marsh, H.W. (1992). *Self-Description Questionnaire-2 (Short)*. Australia, University of Western Sydney.
- McCrae, R.R. & Costa, P.T.Jr. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In: Wiggins JS (Ed), *The five-factor model of personality: Theoretical perspectives*, New York, Guilford.
- McCrae, R.R. & Costa, P.T.Jr. (2003). *Personality in adulthood (2nd ed.): A Five-Factor Theory perspective*. New York, Guilford.
- McCrae, R.R. & John, O.P. (1992) *Journal of personality* 60 (2), 175-215.
- Midgley, C., Maehr, M.L., Hruda, L., Anderman, E.M., Anderman, L., Freeman, K.E., Gheen, M., Kaplan, A., Kumar, R., Middleton, M.J., Nelson, J., Roeser, R., & Urdan, T. (2000). *Manual for the Patterns of Adaptive Learning Scales (PALS)*. Ann Arbor, University of Michigan.
- Mueller, G. & E. Plug (2006). Estimating the Effect of Personality on Male-Female Earnings. *Industrial and Labor Relations Review*, 60(1), 3-22.
- Noftle, E.E. & Robins, R.W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, 93(1), 116-130.
- Nyhus, E.K. & E. Pons (2005). The effects of personality on earnings. *Journal of Economic Psychology*, 26, 363-384
- Osborne, M. (2000). The power of personality: Labor market rewards and the transmission of earnings. *Unpublished doctoral dissertation*.
- Poropat, A.E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322-338.
- Roberts, B.W. & DelVecchio, W.F. (2000). The Rank-Order Consistency of Personality Traits From Childhood to Old Age: A Quantitative Review of Longitudinal Studies. *Psychological Bulletin*, 126 (1), 3-25.
- Salgado, J.F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, 82, 30-43.

Semykina, A. & Linz, S. (2007). Gender differences in personality and earnings: Evidence from Russia. *Journal of Economic Psychology*, 28(3), 387-410.

Srivastava, S., John, O.P., Gosling, S.D. & Potter S. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, 85(5), 1041-1053.

Tangney, J.P., Baumeister, R.F. & Boone, A.L. (2004). High Self-Control Predicts Good Adjustment, Less Pathology, Better Grades, and Interpersonal Success. *Journal of Personality*, 72, 271–324.

Tritt, C. & Hivitt, J.R. (2013). Don't know or don't care: predicting educational attainment using survey item response rate and coding speed test as measures of conscientiousness. *EDRE working paper*, 2013-05.

Appendix 1

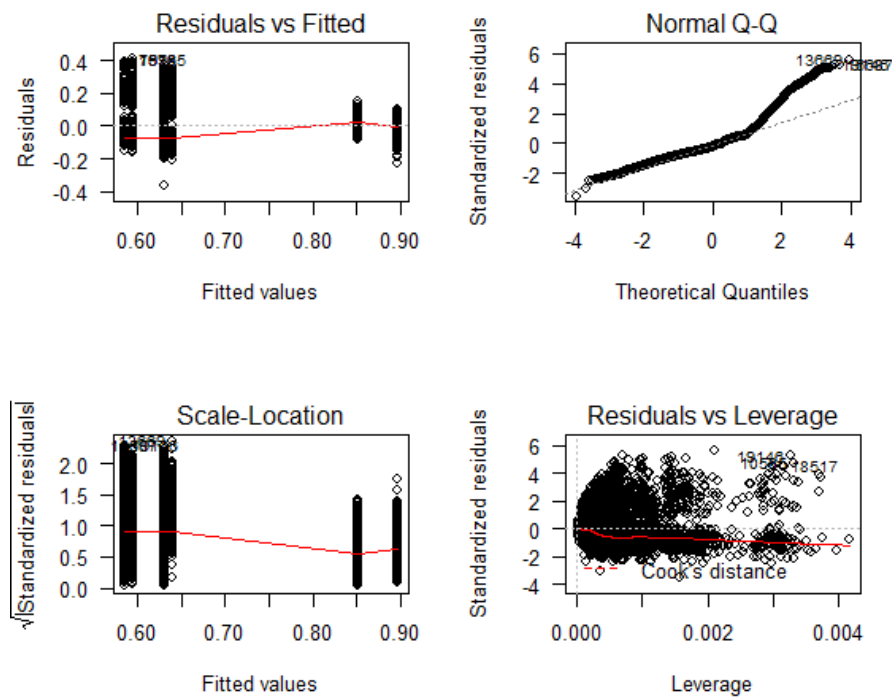
Sample	Observation	Mean	Std. Dev.	Min.	Max.
All	15966	0.7908	0.1687	0.2632	1
Earner	8737	0.8933	0.0543	0.6667	1
- Worker	4325	0.8497	0.0261	0.6667	0.9223
- Self-employed	4412	0.9361	0.0384	0.7439	1
Non earner	7229	0.6669	0.1768	0.2632	1

Note: summary statistics has neither weight nor survey design

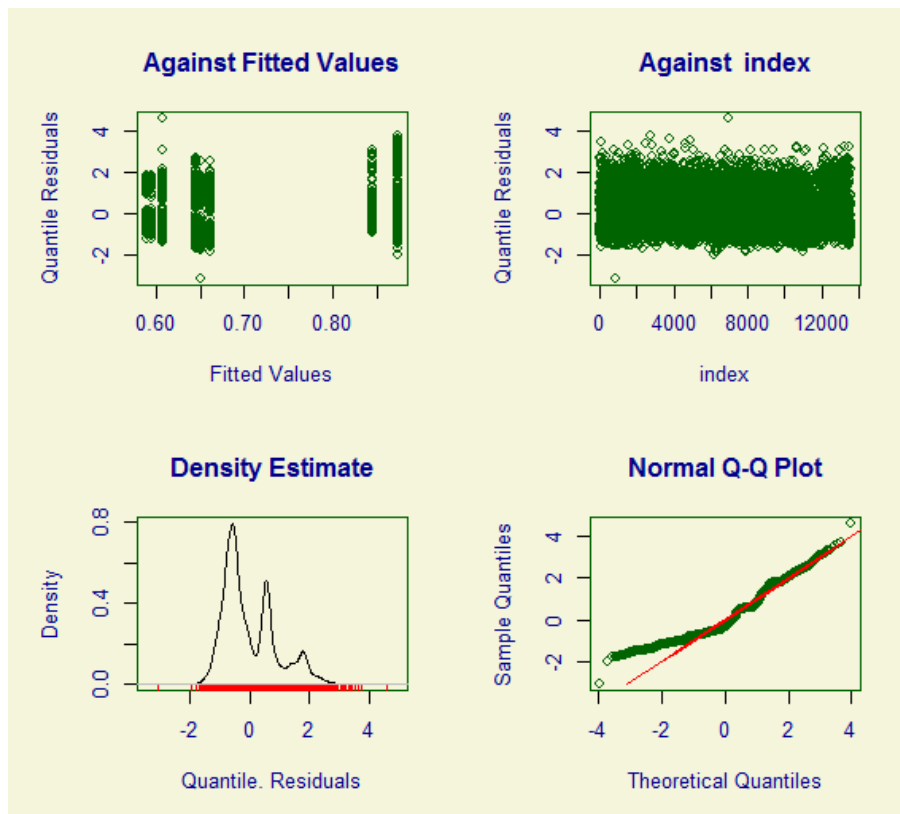
Appendix 2

A. Diagnostic residuals OLS

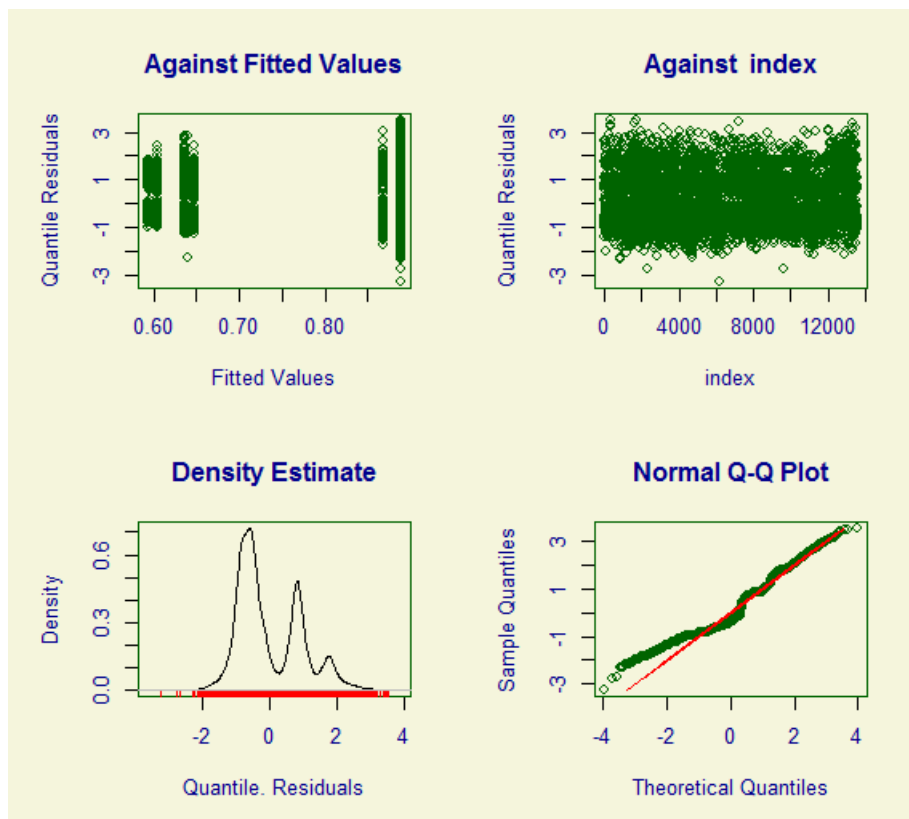
$\text{lm}(\text{resp} \sim \text{factor}(\text{primact}) + \text{factor}(\text{workage}))$



B. Diagnostic residuals beta regression assuming homoscedasticity



C. Diagnostic residuals beta regression assuming heteroscedasticity



Appendix 3

Overall sample	OLS		Heckman		IV	
	Coefficient	Clust. Std. Error	Coefficient	Clust. Std. Error	Coefficient	Clust. Std. Error
Log income per hour						
Response rate	0.8577***	0.2753	0.5671**	0.2261	0.9998***	0.3623
Year of schooling	0.0928***	0.0043	0.0933***	0.0034	0.0919***	0.0043
Age	0.0588***	0.0075	0.0646***	0.0058	0.0604***	0.0078
Age squared	-0.0006***	0.0001	-0.0006***	0.0001	-0.0006***	0.0001
Female (D)	-0.2863***	0.0380	-0.2625***	0.0339	-0.295***	0.0388
Married (D)	0.0936**	0.0420	0.0829**	0.0355	0.0894**	0.0429
Household head (D)	0.1511***	0.0406	0.1194***	0.0339	0.1459***	0.0410
House status (self-owned)						
- Occupying	-0.1539***	0.0429	-0.1166***	0.0350	-0.1230***	0.0422
- Rented/contracted	-0.0618	0.0494	-0.0666	0.0421	-0.0450	0.0505
- Other	0.1606	0.4368	0.5745	0.5416	0.1351	0.4211
Ethnic (Java)						
Sunda	0.0307	0.0579	0.0248	0.0535	0.0472	0.0575
Bali	0.3059***	0.1137	0.1923	0.1306	0.229**	0.1106
Batak	0.2026*	0.1058	0.1288	0.0828	0.2133**	0.1080
Bugis	-0.0153	0.1479	-0.0674	0.1347	-0.0146	0.1447
Tionghoa	0.2037	0.2084	0.1372	0.2161	0.2165	0.2073
Madura	0.0618	0.1023	0.0242	0.0901	0.0765	0.1024
Sasak	-0.2541	0.2256	-0.2676	0.2041	-0.1484	0.2367
Minang	0.3074**	0.1219	0.2403**	0.1160	0.2924**	0.1246
Banjar	0.0892	0.2047	-0.1660	0.1890	0.0888	0.1986
Bima-Dompu	0.1104	0.2102	-0.0726	0.1955	0.1231	0.2218
Nias	0.2358	0.5162	-0.2094	0.6200	0.2141	0.5222
Palembang	-0.2478	0.1918	-0.1937	0.1874	-0.0968	0.1990
Sumbawa	0.6038***	0.2152	0.3732*	0.2111	0.6023***	0.2205
Toraja	1.4529***	0.0804	1.526***	0.0639	1.4874***	0.0845
Betawi	0.1568*	0.0927	0.0546	0.0723	0.1734*	0.0955
Melayu-Deli	0.6694**	0.3001	0.3765	0.2431	0.3701*	0.2046
Komering	0.4697*	0.2856	0.4106	0.3134	0.485*	0.2841
Ambon	0.9986**	0.5095	0.9368**	0.4620	0.2879***	0.1118
Manado	-0.3054***	0.0786	-0.3883***	0.0666	-0.2766***	0.0812
Other South Sumatra	-0.2392	0.1602	-0.1917	0.1510	-0.2177	0.1641
Other	-0.1096	0.1059	-0.0208	0.1024	-0.1302	0.1047
Region (North Sumatera)						
West Sumatera	-0.3867**	0.1847	-0.2830	0.1645	-0.3484*	0.1877
South Sumatera	0.2855**	0.1365	0.3211***	0.1232	0.2762**	0.1385

Lampung	-0.0870	0.0959	-0.0884	0.0869	-0.0998	0.0971
DKI Jakarta	0.1897**	0.0869	0.2007***	0.0746	0.1709**	0.0891
West Java	0.1375*	0.0800	0.1266***	0.0696	0.1492*	0.0810
Central Java	-0.0411	0.0817	-0.0325	0.0708	-0.0231	0.0828
DI Yogyakarta	-0.1641*	0.0889	-0.1666**	0.0768	-0.1710	0.0891
East Java	-0.0384	0.0816	-0.0381	0.0707	-0.0395	0.0826
Bali	-0.2221*	0.1235	-0.1575	0.1329	-0.1508	0.1219
NTB	-0.1324	0.2052	-0.0170	0.1917	-0.1430	0.2143
South Kalimantan	0.2807*	0.1594	0.3465**	0.1469	0.2762*	0.1613
Main language (Indonesian)						
Javanese	-0.16***	0.0485	-0.2099***	0.0420	-0.1503***	0.0479
Sundanese	-0.0864	0.0730	-0.13*	0.0670	-0.0969	0.0736
Balinese	0.0084	0.1116	0.0229	0.1055	-0.0172	0.1123
Batak	-0.1063	0.2445	-0.1247	0.2422	-0.1073	0.2445
Maduranese	-0.4226***	0.1541	-0.4023***	0.1447	-0.4147***	0.1552
Sasak	0.1660	0.1409	0.0588	0.1167	0.0906	0.1427
Minang	0.2230	0.1669	0.1784	0.1462	0.2412	0.1714
Banjar	-0.0059	0.2054	0.0752	0.1804	-0.0134	0.2058
Bima	-0.1569	0.3061	-0.0769	0.2801	-0.1197	0.2977
Nias	-1.796**	0.8201	-1.4499	0.8839	-1.7659**	0.8227
Palembang	-0.0742	0.1559	-0.1122	0.1423	-0.1225	0.1607
Lahat	-0.1823	0.4138	-0.3864	0.3343	-0.1772	0.4131
Other South Sumatera						
Betawi	0.5853***	0.0724	0.5652***	0.0622		
Select						
Household head			0.5424***	0.0292		
Married			0.3171***	0.0246		
Age			0.0043***	0.0009		
Year of schooling			0.0234***	0.0030		
Sex			-0.1996***	0.0130		
Observation	7378		15883 (uncensored: 7799)		7707	
Adjusted R-squared	0.1818				0.1919	
Root MSE	1.496				1.205	
Log pseudolikelihood			-21741			

Note: Significance level are 0.01(***), 0.05(**) and 0.1(*).

Appendix 4

OLS	Log wage per hour		Log profit per hour	
	Coefficient	Clust. Standard error	Coefficient	Clust. Standard error
Response rate	1.4595**	0.6479	0.4946	0.7088
Year of schooling	0.1139***	0.0047	0.0569***	0.0079
Age	0.0682***	0.0072	0.04424***	0.0147
Age squared	-0.0007***	0.0001	-0.0005***	0.0002
Female (D)	-0.1889***	0.0389	-0.4854***	0.0815
Married (D)	0.1609***	0.0438	-0.0393	0.0801
Household head (D)	0.0852**	0.0432	0.1306	0.0820
House status (self-owned)				
- Occupying	-0.1884***	0.0467	0.0396	0.0851
- Rented/contracted	-0.0636	0.0573	0.0207	0.0996
- Other	1.3633***	0.1890	-0.5932***	0.2012
Ethnic (Java)				
Sunda	0.0490	0.0691	0.0890	0.0964
Bali	0.3633***	0.1203	0.1429	0.1995
Batak	0.4953***	0.1478	-0.0312	0.1631
Bugis	-0.2551	0.1569	0.3142	0.2173
Tionghoa	-0.1621	0.1927	0.5416*	0.3183
Madura	-0.0024	0.1124	0.1918	0.1732
Sasak	0.0594	0.2528	-0.4082	0.3253
Minang	0.2205**	0.0904	0.4271	0.2812
Banjar	0.2141	0.3053	-0.0354	0.3522
Bima-Dompu	0.7311***	0.1960	-0.4751*	0.2828
Nias	0.7116	0.3564	-0.1297	0.7341
Palembang	-0.2573	0.2576	0.1606	0.2998
Sumbawa	0.8732***	0.2230	0.2665	0.3044
Toraja			1.9661***	0.1552
Betawi	0.0891	0.0834	0.2736	0.1841
Melayu-Deli	0.5276	0.3221	0.2235	0.4873
Komering	0.6846***	0.0871	0.3740	0.5762
Ambon			0.5673***	0.1613
Manado	-0.0275	0.0818	-1.0804***	0.1762
Other South Sumatra	-0.3509*	0.1880	-0.0886	0.2570
Other	-0.0204	0.1126	-0.4025*	0.2432
Region (North Sumatera)				
West Sumatera	-0.2403	0.1848	-0.5700	0.3673
South Sumatera	0.3053*	0.1580	0.1814	0.2168
Lampung	-0.0235	0.1097	-0.2719*	0.1653

DKI Jakarta	0.2515***	0.0861	-0.0578	0.1787
West Java	0.1794**	0.0758	-0.0149	0.1571
Central Java	0.0985	0.0767	-0.1969	0.1570
DI Yogyakarta	-0.0507	0.0813	-0.4071**	0.1728
East Java	0.0284	0.0744	-0.1957	0.1590
Bali	-0.1554	0.1265	-0.3105	0.2267
NTB	-0.4737**	0.2006	0.0581	0.3071
South Kalimantan	0.2657*	0.1540	0.2134	0.3033
Main language (Indonesian)				
Javanese	-0.1518***	0.0508	-0.1418*	0.0844
Sundanese	-0.0521	0.0761	-0.1204	0.1351
Balinese	-0.2181	0.1492	0.1498	0.1618
Batak	0.9923***	0.2363	-0.1366	0.2484
Maduranese	-0.2409	0.1678	-0.6122**	0.2555
Sasak	0.1821	0.1956	0.1157	0.1907
Minang	0.3209	0.1988	0.1383	0.2874
Banjar	-0.0552	0.2590	0.0336	0.4182
Bima	-0.2135	0.3452	0.1375	0.5974
Nias			-1.6332*	0.9880
Palembang	0.0408	0.1752	-0.3426	0.2585
Lahat	0.7825**	0.3619	-0.6893	0.4982
Observation	3770		3237	
Adjusted R-squared	0.3403		0.1110	
Root MSE	0.8083		1.2285	

Note: Significance level are 0.01(***), 0.05(**) and 0.1(*).