This paper is a draft submission to the

WIDER Development Conference

Human capital and growth

6-7 June 2016 Helsinki, Finland

Does education affect time preference?*

Tushar Bharati[†] Seungwoo Chin[‡] Dawoon Jung[§]

January 13, 2015

Abstract

Individual time and risk preferences are important determinants for economic behavior and outcomes. The neoclassical economic theory assumes individual time preference is constant across time for the individuals and does not pay much attention to the determinants of individual time preference. Using exogenous variation in access to schools caused by spatial variation in Indonesia primary school construction project (INPRES), we estimate the causal impact of education on time preference. We find significant effect of education on time preference for females. IV-Probit model suggests that one more year of schooling makes people 5 to 6 percent points more patient while a conventional IV-2SLS estimates suggest a increase in patience by 6 to 7 percentage points. Factor analysis (taking advantage of behavior measure information) and bound analysis produce consistent results. Components of cognitive ability and health seem to be probable mechanisms. Our results suggest that educational opportunities in developing countries can affect time preference that may, in turn, affect investments in human capital production and health, starting off a virtuous cycle.

JEL Classification Codes: D01, D90, I25

Keywords: Individual preferences, education, Indonesia

*We are grateful to Professor Jeff Nugent and Professor John Strauss for their valuable feedback. We are also grateful to Professor Manisha Sha, Lisa Cameron and Craig McIntosh for their valuable comments. This paper has benefited substantially from feedback provided by Ryan Kendall, Riddhi Bhowmick, Teresa Molina and participants of the informal Development Economics seminar at University of Southern California. This paper has also benefited from participants at Western Economic Association International at Hawaii 2015. We are thankful to Professor Esther Duflo for providing us INPRES data. All errors are our own.

[†]Department of Economics, University of Southern California; tbharati@usc.edu

[‡]Department of Economics, University of Southern California; chinseun@usc.edu

[§]Department of Economics, University of Southern California; dawoonju@usc.edu

1 Introduction

Individual preferences(time or risk preference) are very important in the sense that they are important inputs for economic behavior and socioeconomic outcomes (Stigler and Becker (1977)). For example, investments in stock market, insurance market, or health care are associated with individual preferences, which are also crucial determinants of many other outcomes such as employment, income and wealth accumulation. In the context of economic growth, individual preferences are obviously of great importance. According to Ramsey's growth model, since individual time preference is closely related to saving and consumption behavior, it can explain the economic growth and inequality. Further, individual time preference can explain the microeconomic behavior such as mechanisms through which education causally affects health (Cutler and Lleras-Muney (2006)). Thus, understanding the effect of or the determinants of individual time preference is critical in analyzing the macroeconomic condition and microeconomic behavior consequences.

However, neoclassical economic theory does not pay much attention to the determinants of individual time preference. Individual time preference is assumed to be constant across time period. Economists have raised questions about this assumption and have investigated the endogenous formation of individual preferences. Also, several studies have examined the stability of individual preferences either by empirically studying the observational data or by conducting relevant experiments. In a recent study by Meier and Sprenger (2015), they found out temporal instability of time preference for some individuals even though the reason for the instability is not certain whether that is due to demographic backgrounds or not.

In this study, we empirically examine the causal relationship between education and individual time preference by testing the time preference hypothesis. Proponents of time preference hypothesis argue that more patient individuals (those who have a low discounting rate) are likely to have more education (Grossman (2006)). As Becker and Mulligan (1997) argue, however, :

> "*Schooling also determines....[investments in time preference] partly through the study of history and other subjects, for schooling focuses students' attention on the future.*"

These debates over the effect or the determinants of individual time preference are unanswered empirical questions. Becker and Mulligan (1997) have provided a theoretical framework for endogenous formation of time preference. Based on this initial study, several empirical studies found out the difference in individual time preferences across individuals (Leigh (1986); Barsky et al. (1997); Dohmen et al. (2011)) due to different socioeconomic backgrounds. However, relatively less studies have been studied in the context of developing countries (Hamoudi (2006); Ng (2013)). These studies have shown the correlation between individual preferences and many socioeconomic factors such as education, marriage, drinking/smoking behavior, age, sex, household wealth and so on. Some studies have also checked the association between individual preferences and some historical events such as a natural disaster (Cassar et al. (2011); Cameron and Shah (2015); Hanaoka et al. (2015)) and financial crisis (Guiso et al. (2013)). However, these associations among these variables might not reflect the casual relationship, which have rarely been investigated empirically. Hryshko et al. (2011) examined the causal relationship between parent education and children's risk preference while Perez-Arce (2011) tried to explain the casual link of own education and own preference at the first time using experiment data (Mexico public college lottery entrance system). Very few studies try to look at the causal impact of education on individual preferences. In this paper, we want to explain the endogenous formation of individual preferences (focus mainly on time preference) affected by education level in Indonesia using IFLS4 (Indonesia Family Life Survey).

The question that if education affects time preference has always been an interesting hypothesis which has rarely been examined empirically. Actually, there are two difficult problems in studying the relationship between education and individual preferences (Perez-Arce (2011)). First, estimating the causal effect of education on time preference is nontrivial because decision of education is endogenous so that unobservable factors could affect both education and individual preferences, which restricts causal interpretation. Second, reverse causality may drive this relationship. That is, more patient person or more risk averse person is more likely to invest their time on education. Third, individual preferences are not easily measurable. Many studies that have investigated individual preferences use hypothetical questions which may fail to capture true underlying preferences. And these hypothetical time preference questions are not clear as to whether it reflects only time discounting rate or it reflects true time preference or uncertainty, or the perception of opportunity costs (Frederick et al. (2002)). Even though some studies already document that hypothetical survey questions are good predictors of actual behavior and preferences (Barsky et al. (1997); Donkers et al. (2001); Anderson and Mellor (2008); Dohmen et al. (2011)), we should be careful about imperfect measurement. In order to minimize the measurement error in the survey responses of individual preferences, we provide several different specification results such using factor analysis model and bound analysis with theoretical tools. This is provided in the appendix. In sum, the primary goal of this paper is mainly focused on the reduced form result on the causal effect of education on individual preferences.

This paper uses a natural experiment, a large-scale primary school construction project launched in 1973/4 and continued until 1978 in Indonesia. This project was the fastest and one of the largest primary school construction program in the world (Peters (1990)). This large-scale project offered great opportunity to attend the primary school for those who would or could have not attended. Using this exogenous variation in education level, Duflo (2001) estimated the returns to school (about 0.2 years increase). Being inspired by this seminal work, we could leverage this natural experiment to reveal the causal link from education to individual preferences. We mainly focus on female samples and found out that education makes female respondents more patient by five to six percentage point. This amounts to about 7% decrease in impatience. We found very consistent empirical results across various specifications such as IV-2SLS, IV-Probit and through several robustness checks. We suggest the plausible mechanism such as individual cognitive skills and health status through which education affects patience.

Our study makes two contributions. First, we find out the causal link from education to time preference in the developing countries. This helps to explain the structural formation of individual time preference. By providing the empirical evidence on this topic, we could have a better understanding how education policy affects individual behavior and outcomes in the long run. Second, in addition to reduced form estimation result, we provide suggestive mechanisms. We address the common topic regarding to human capital formation, cognition and non cognition (Heckman (2007)) by mentioning that education affects cognitive ability and again affects the formation of non-cognitive skills, which is called "skill begets skill" (Heckman (2000)). We also address the old issue about the relationship between education and health by considering individual time preference as potential mechanisms.

The remainder of the paper is organized as follows. Section 2 describes the INPRES program in Indonesia. Section 3 describes the data, IFLS4 and INPRES data. Section 4 explains the empirical strategy. Section 5 presents the empirical results and section 6 concludes.

2 Program Background

In 1973-74, Indonesian government launched Sekolah Dasar INPRES school construction program to provide equal opportunity of education. Between 1974 and 1978, over 61,000 new primary schools were constructed. According to Duflo (2001, 2004), one newly constructed schools were provided for about 500 children of ages 5 to 14 in 1971. In order to balance the quantity and quality of schools, Indonesian government recruited the teachers and the proportion of teachers who satisfied qualification criteria did not change much between 1971 and 1978 (Duflo (2001)). But unfortunately, the school quality measure were not measured, and teacher quality could not be estimated accurately. Indonesia primary school continues from 6-7 years old to 11-12 years old. Thus, children with ages before 11-12 (born in 1962 or after 1962) could have partially or entirely affected by the INPRES. Since especially the cohorts born between 1963 and 1967 could have been affected by INPRES or not, we are going to use cohorts born between 1968 and 1972 (treatment group, entirely affected by INPRES), and born between 1950 and 1962 (control group, entirely not affected by INPRES).

3 Data and Measures

3.1 IFLS4 and INPRES data

Indonesian Family Life Survey (IFLS) 4 is the latest survey (2007) of on-going longitudinal survey IFLS-series since 1993/4 by RAND. The survey respondents are representative of about 83% of Indonesian population and contains over 30,000 individuals living in 13 of the 27 provinces in the country (Strauss et al. (2009)). We use cohorts born between 1950 and 1972 (some of them were not affected by INPRES while others (born between 1968 and 1972) were affected). We used same cohorts that were used in Duflo (2001, 2004). The IFLS4 survey collected individual preferences using hypothetical questions. Individual preferences are the main dependent variables in this paper. We also exploited information from previous survey, IFLS1-3 for some of missing variables that were useful control variables (such as education, place of birth, migration, mother/father education, religion, ethnicity and so on).

INPRES data were obtained from Esther Duflo. She collected the data of number of constructed schools and calculated the intensity of school construction by year-Kabupaten (district) level. The intensity of school construction measure were recorded based on the province and Kabupaten code in 1995 Intercensal survey of Indonesia (SUPAS95). We match IFLS4 and INPRES based on the Propinsi(Province) and Kabupaten (District) code in SUPAS95. Since these Propinsi and Kabupaten name could be traced and matched from codes in IFLS1-4, we could link Propinsi by Kabupaten name of IFLS to SUPAS95. However, it would be a problem if name had been changed between 1995 and 2007 when they actually pointed the same area.

3.2 Outcome Variables

We look at individual preferences: time preference and risk preference, which are based on hypothetical questions with four different categories. You can check the question structure in the figure 1 and 2. The following is the samples of questions asked in the survey.

Time Preference A
You have won the lottery. You can choose between being paid
1. Rp 1 million today or 2. Rp 3 million in 1 year

If the answer to the question is '1', the interview continues asking a question with different amounts structure.

You have won the lottery. You can choose between being paid
1. Rp 1 million today or 2. Rp 6 million in 1 year

The following is the questions of time preference B.

Time Preference B
You have won the lottery. You can choose between being paid
1. Rp 1 million today or 2. Rp 4 million in 5 years

If the respondent answered '1', ask again the question:

You have won the lottery. you can choose between being paid
1. Rp 1 million today or 2. Rp 10 million in 5 years

The risk preference question also consists of two different sets. However, risk preference B questionnaire has a little problem with logical consistency of question array. We doubt about the responses to this risk preference B questions. We drop this question and only use risk preference A.

Risk Preference A
Suppose you are offered two ways to earn some money.
With option 1, you are guaranteed Rp 800 thousand per month
with option 2, you have an 50-50 chance of receiving either 1.6 million per month or Rp 400 thousand per month,
depending on how lucky you are.

Figure 1: Time preference cateogries

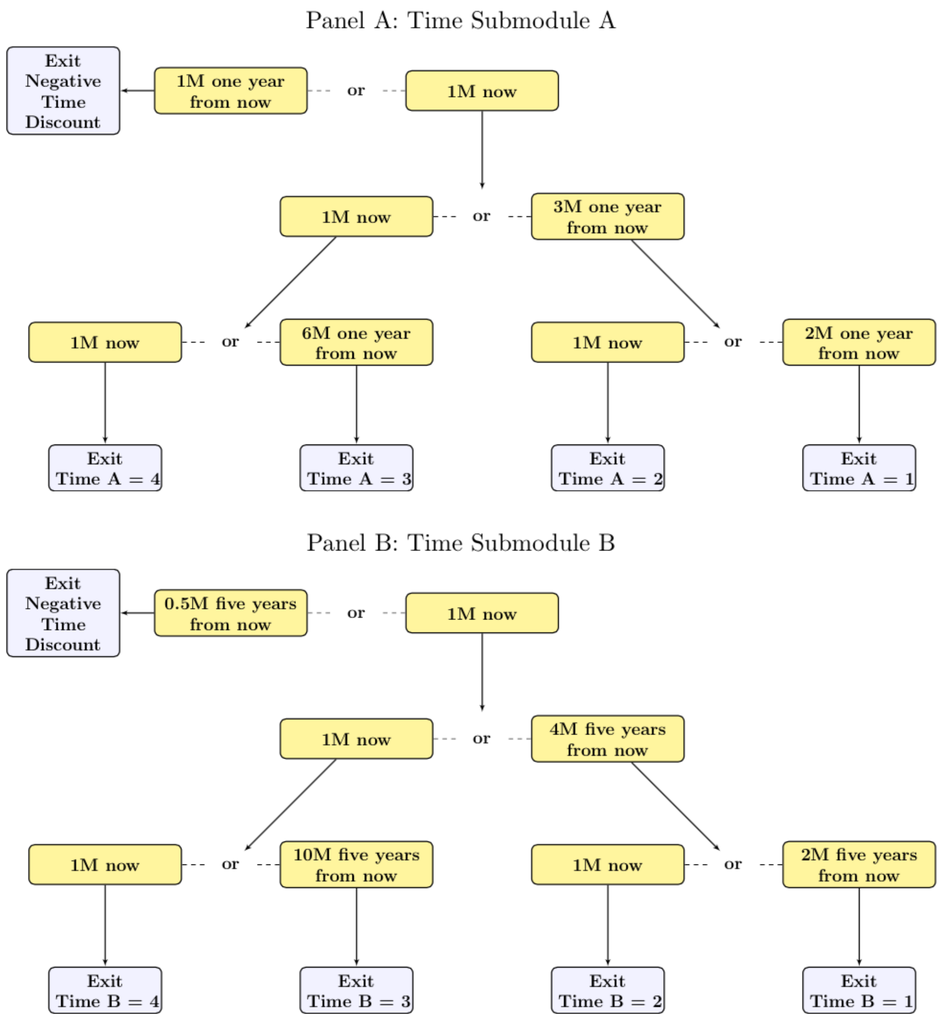
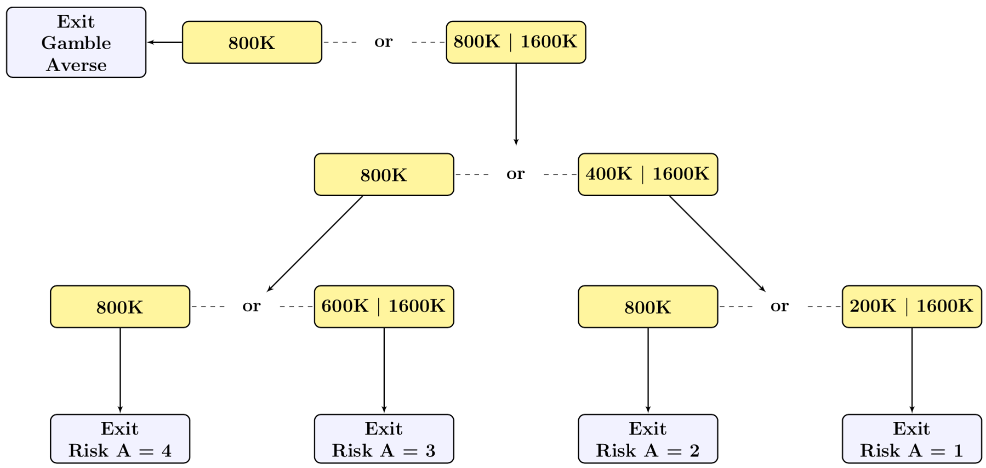


Figure 2: Risk Preference categories

Panel A: Risk Submodule A



Depending on the option that the respondent choose, the questions ask other combination of amounts with same odds. The figure 1 and 2 shows the four categories of each time preference and risk preference[1]. According to the structure of questions, we define the respondents belonging to category4 of time preference as the most impatient individuals. Similarly, the respondents belonging to category4 of risk preference are the most risk averse (the least risk tolerant) individuals.

For both time preference and risk preference questions, IFLS4 first asked respondents and checked if they are logically rational. That is, for time preference, the survey asked whether respondents choose Rp 1 million today or Rp 1 million in 1 year. If they chose the latter one, we don't trust their response (negative time discount) and we drop those respondents. Similarly, for risk preference, the survey asked whether respondents choose option 1 (Rp 800 thousand per month) or option 2 (50 chance of Rp 800 thousand per month or 50 chance of Rp 1.6 million per month). If they chose option 1, we don't trust their response and we again drop those individuals.

Table 1 shows the proportion of respondents in four different categories of both time preference and risk preference. Most of respondents are impatient (category 4) and are risk averse (category 4). More than 70 % of individuals are impatient, which are readily observed in developing countries. We also checked the consistency between time preference A and B and we found that if individual chose category 4 in time preference A, then they are likely to choose category 4 again in time preference B (Table 2). We also checked the relationship between time preference and risk preference responses and we found that most people are risk averse (category 4) conditional on their being impatient (category 4). The proportion of respondents who said they are risk averse are two times larger than the proportion of those who answered they are risk taking conditional on their being impatient.

[1]These charts are cited from Ng (2013).

Table 1: Time preference and risk preference response rate

Time preference A	Category 1	Category 2	Category 3	Category 4
proportion	0.072	0.056	0.140	0.733
observations	328	257	640	3,357
Time preference B	Category 1	Category 2	Category 3	Category 4
proportion	0.021	0.034	0.117	0.828
observations	96	159	539	3,824
Risk preference A	Category 1	Category 2	Category 3	Category 4
proportion	0.270	0.138	0.085	0.508
observations	711	364	224	1,343

Notes: For time preference A and B, category 4 is the most impatient while category1 is the most patient. For risk preference A, category 4 is the most risk averse while category 1 is the most risk taking. .

Table 2: Number of respondents between time preference measures A and B

	timeprefa1	timeprefa2	timeprefa3	timeprefa4
timeprefb1	90	10	23	8
timeprefb2	113	63	35	23
timeprefb3	130	99	399	153
timeprefb4	152	200	490	4822

3.3 Measurement of Individual Preferences

IFLS4 asked respondents to reveal their time and risk preference using hypothetical questions. Some studies suggested that answers to survey questions about individual preferences are consistent with estimates from experiments and actual behavior (Dohmen et al. (2011); Hamoudi (2006)). Thus, estimating the effect on individual preferences by using hypothetical survey questions could provide unbiased estimate of true effect. However, potential measurement error in individual preferences such as 'status quo bias' (Barsky et al. (1997)) or classical measurement error could underestimate the true effect. In order to overcome the potential measurement error, for robustness check, we suggest the evidence using factor analysis, bound analysis and so on. This is explained in the appendix with empirical results (Appendix section A, B, and C).

3.4 Independent Variables

Our independent variable of main interest is education or completed years of schooling. The main hypothesis that we try to test is whether education causes the formation of individual preferences. Since education is endogenously determined and is affected by many unobservable factors, the estimation of the effect of education on our dependent variables is going to be biased unless we control for endogeneity. Many studies have examined the returns to schooling in the context of labor market outcome (Card (1999); Duflo (2001)), health (Cutler and Lleras-Muney (2010)) and so on. However, only a few studies have examined associations between education and individual preferences (Becker and Mulligan (1997); Hryshko et al. (2011)). This study uses exogenous variation in education to estimate the causal effect of education on individual preference. To our best knowledge, Perez-Arce (2011) is the only paper trying to estimate the causal effect of education on individual preferences by using Mexico's public university entrance system. Mexico City started the lottery system for public university entrance in 2001. Using this exogenous variation in opportunity of studying in university, they could measure the causal effect of public university education on individual preference. Even though random assignment increases the probability of attending university, selection into university can not be just ruled out. In addition to this problem, those who are supposed to attend college might have already formed their individual preferences. While they found out the significant effect of education on individual preferences, the estimates might capture the effect of other factors that might be correlated with attendance of university and individual preferences.

We focus on primary school from which the initial education begin. In 1973-74, large primary school construction increased opportunity for cohorts born between 1968 and 1972 in this study sample. We leverage this exogenous change in education opportunity. In all our regressions, we include year of birth, year of birth square or year of birth fixed effects, gender dummy, urban dummy, religion dummy, ethnic dummy, season fixed effect, province-Kabupaten fixed effect. As Duflo (2001) pointed out, primary school construction program could have had dependent on children population, enrollment rate in periods before INPRES was introduced and on other national programs, we also control for population age of 5 to 14 in 1971, primary school enrollment rate in 1971 and the water and sanitation program (the second largest national-level program in 1973/4).

4 Empirical Analysis

4.1 Strategy

In this section, we describe our instrumental variables (IV) strategy. For all the IV regressions, we also present the validity of our instrument variables. As we already mentioned in the previous section, OLS estimates of the effect of education on individual preferences will likely be biased due to omitted variables, reverse causality and measurement error. We propose instrumenting for education using exogenous variation in school construction intensity. Here, we take advantage of two different specifications. One is control function methods or the instrumental variable probit model (IV-Probit) that is used when we have continuous endogenous variable and binary outcome variables (Lewbel et al. (2012). Another is a conventional IV-2SLS method. Both methods use the same strategy in the first stage and only the difference is wether second stage regression uses Probit or LPM (linear probability) model. In addition, IV-Probit is maximum likelihood estimation and provides consistent estimates only if the endogenous regressors in the model are continuous (Lewbel et al. (2012)).

As Duflo (2001) already showed the positive effect of primary school construction on education (schooling) by using difference-in-difference strategy, we replicate the exactly same identification strategy. The difference between Duflo (2001) and this study is the data set (SUPAS 95 VS IFLS4) and we focus on females mainly. After we estimate the effect on year of education, we could use predicted schooling level to estimate the second stage effect. In the second stage, we examine the effect of predicted schooling on individual preferences. The two-stage IV is specified as:

$$1^{st} stage : S_{ijc} = \alpha + \delta_j + \gamma_c + (P_j T_i)\rho + X_i + \epsilon_{ijc} \tag{1}$$

where i is individual, j is region, c is cohort, P_j denotes the intensity of the INPRES program in the region of birth measured by the number of school built per 1,000 children, T_i is a dummy indicating whether the individual belongs to the "young" cohort (born between 1968 and 1972; old cohort was born between 1950 and 1962), δ_j(province-Kabupaten) is region fixed effect and γ_c is cohort fixed effect(Some regressions use year of birth and year of birth quadratic). Individual controls X_i includes mother and father education, health status and an urban dummy. In addition, we use season fixed effect to absorb any static differences between seasons (rainy and dry season), religion fixed effect (or a dummy for muslim), ethnicity fixed effect as well. The second stage equation is,

$$2^{nd} stage : Y_{ijc} = \mu + \delta_j + \gamma_c + \beta_1 \hat{S_{ijc}} + X_i + \eta_{ijc} \tag{2}$$

where Y_{ijc} is an dummy variable which is equal to 1 if respondents are the most impatient (category 4 for time preference) or respondents are the most risk averse (category 4 for risk preference) and equal to 0 otherwise. We also use province-Kabupaten fixed effect, cohort fixed effect for the second stage. All the controls used in the first stage are included in the second stage. In addition, we include early life rainfall shocks (rainfall at birth), which was one of important factors in determining adult health and education (Maccini and Yang (2009)) and rainfall shocks between one year old and nineteen years old. Rainfall could affect school attendance especially for young children who go to primary school. We create rainfall variables by subtracting average rainfall amounts of province-Kabupaten and month specific from rainfall at the age of 0 to 19 years. The rainfall amounts are calculated by weighted average of rainfall within 100km stations' values. Since more rainfall is conducive to producing more agricultural products, we assume that more rain is a positive shock. In the main regression, we exclude some region specific variables that were used in Duflo (2001). However, as Duflo (2001) suggested that if other programs such as water and sanitation program which was the second largest national program in Indonesia around 70's influenced both INPRES placement and education level differently for treatment and control group, we could have a biased estimate. Although province-Kabupaten fixed effect absorbs any unobserved static differences across areas, time-varying unobservables could invalidate our estimates. For robustness check, we include interaction term with treatment group dummy and water sanitation program, enrollment rate in 1972, children population in 1971. Robust standard errors are clustered at the province level.

In the main regression, β_1 is the vector of coefficients of interest and it is the impact of education on individual preferences. Not only comparing cohorts affected by INPRES (born in between 1968 and 1972) to cohorts not affected by INPRES (born in between 1950 and 1962) but also comparing cohorts born in the same year but who were born in different areas, using different intensity level of INPRES construction allows us to use regional variation and time(year) variation as well.

4.2 Intuition

The hypothesis that we are going to test is very simple and intriguing questions, that is, whether education causally affects the formation of individual preferences. If we find the significant reduced form result, one of the mechanisms through which education affects health could be explained by individual preferences. This finding will provide very important policy implication that providing basic level of education or offering an opportunity to learn something from schools in developing countries help people form individual preferences, which can affect long-term economic outcomes such as saving, health care and parenting. Individual preferences are also known to be closely related to non-cognitive ability (Chiteji (2010)), which is recently getting the most attention due to its importance in determining health behaviors and human capital formation (Heckman (2007)). Education itself is not an outcome of human capital production function rather it is an investment. If we find some connections between education and individual preferences, the effect of education on non-cognitive or cognitive skill formation without considering individual preferences might be over- or underestimated. Thus, human capital production function should include individual preferences in its structural function to fully explain the formation of human capital.

On the contrary, if we don't find any casual effect of education on individual preferences, we could argue that the significant correlation between education and individual preferences could be explained by reverse causality. Economists have assumed, especially in developing countries, that more patient people are likely to be educated. This hypothesis has never been studied because finding an exogenous variation in individual preferences is very difficult. However, by suggesting this null result of the effect of education on individual preferences, we could provide meaningful interpretation about two variables. Further, individual preferences might not be the crucial factor that links education and health, which has rarely been tested (Cutler and Lleras-Muney (2006)). We discuss the results in the section below.

5 Results

5.1 Correlation between education and time preference

In table 4 and 5, we report our results estimating the association between individual time preference and education. In order to minimize omitted variable bias, we control several fixed effects that are included in the main regressions[2]. Education is significantly correlated with time preference. For example, in table 4, the female respondents who have one more year of schooling are 0.3-0.4 percentage point more likely to be patient. We have stronger association between time preference and education in table 5. The average years of schooling is 7.6 years for our samples and the coefficient can be thought of as about 2.3 percentage point of being more likely to be patient. Most of respondents (over 70%) are impatient, about 3 percentage point corresponds to less than 5%. Although the coefficient is very small, education and individual preferences are in a negative relationship. These results are consistent with previous studies(Barsky et al.

[2] We used season fixed effect, porvince-Kabupten fixed effect, ethnicity fixed effect, religion fixed effect and year of birth fixed effect.

(1997); Dohmen et al. (2010)). However, in this study, we don't find any significant relationship between education and risk preference (Appendix table A3).

5.2 Main result

In table 6 and 7, we suggest maximum likelihood estimates from IV-Probit model. The reason that we don't present male sample is that the calculation of MLE does not converge. Table 6 displays the result with respect to time preference measure B. We observe significant coefficient of education for female only. We provide the marginal effect of education separately for the clear interpretation because these specifications use probit model in the second stage. One more year of schooling makes people about 6 percentage point more likely be patient. It is about 7% (the proportion of people being impatient is 80%). Since the Wald test of exogeneity in column 3 does not reject the null hypothesis of no endogeneity, we should interpret it with caution. In this case, since there is not sufficient information in the sample to reject the null, a regular probit regression might be appropriate. However, we reject the Wald test in column 2 and 4, and we still have very close significance level (15 % significance) in column 3 and 5 with similar marginal effect. The differences come from the specification that one includes year of birth and square of year of birth to absorb any nonlinear trends in time preference and another includes year of birth fixed effect. In column 4 and 5, the results are similar to column 2 and 3. All the estimations include season fixed effect, province and Kabupaten fixed effect, religion fixed effect. We also include JAWA dummy variable (about 50% people's ethnicity is JAWA) to control for ethnicity[3]. However, we find significance only for the column 2 on time preference measure A while the magnitude of marginal effect is quite similar to table 6. The IV-Probit result for risk preference A is presented in the appendix table A4. We assume that time preference measure B captures true time preference more accurately compared to time preference measure A because time preference measure B is asking now and more further future than time preference measure A. That's why we observe the difference in significance level between time preference measure A and B. We argue that this is more likely to be the measurement issue not the issue of actual effect.

Based on IV-Probit results, now we turn to the conventional IV-2SLS result. Table 8 displays the results only for female. As several studies found out that early life circumstances are important for the long-term health and economic outcomes (Maccini and Yang (2009); Adhvaryu et al. (2014)), our estimate could confound the effect of education on individual preferences unless we do not control for early life environments. As explained in the Data section, we include rainfall shocks, that is, rainfall at birth, and rainfall at six years old to rainfall at nineteen years old, when children attend the primary school. Except the specification in column 2, we have significant effects in the second stage with the very good instrument variable (IV first stage F statistics). The magnitude of coefficient in the second stage is much larger than OLS estimates and similar or little bit greater than IV-Probit estimate of the marginal effects. One more year of schooling increases patience by 6 to 7 percentage point. What we observe is stronger negative relationship when we use IV approach. This might generate interesting discussion. First,

[3] When we include ethnicity fixed effect, MLE does not converge.

it is possible that in developing countries, more patient individuals might be forced to stop schooling and to participate in the labor market to support the family. This could drag down the magnitude of the correlation coefficient. Second, since IV estimate is LATE (Local Average Treatment Effect), those who were affected by INPRES program were potentially those who were at the margin. They would not have been educated so the effect of education turns out to be stronger for these group. In the appendix, table A5 and A6 also displays the 2SLS results for time preference measure A and risk preference measure A for female. No significant effects in the second stage have been found. Even though we found no significant effects in the second stage for time preference measure A, the magnitude of the effect of education on time preference is similar to the estimates from the IV-2SLS estimates of time preference B.

We also present the first stage results, that is, the effect of our instrument (INPRES intensity) on years of schooling. We found the significant effects mostly for female. We obtain one school built per 1,000 children (INPRES intensity) causes about 0.3 years to 0.5 years more schooling. Compared to Duflo (2001) where the effect is 0.2 years more schooling for male, it is a little bit greater. In addition, these effects are larger than those found in other studies in the context of developing countries (Miguel and Kremer (2004)) but it is less than the estimates in Behrman et al. (2011) and Field et al. (2009). The magnitude of estimates is not directly comparable give the fact that the data set and the target samples are different.

For robustness check, we estimate the regression without rainfall shock (Tables can be presented upon requests) and still have consistent results as main specifications. We also estimate the same specification with different clustering units (Table A7 -A9). Since the usual OLS standard errors can violate the standard assumption, we take into account the correlation within province level which is larger than prokab (province-Kabupaten) across all specifications so far. However, the number of clusters when we use province as a cluster unit is less than 50, robust clustered standard error may be incorrectly estimated (Cameron and Miller (2015)). We want to suggest the results using the standard error calculated from different cluster units. We run the exact same specification and find the similar results as the main specification although we lose some strength of our IV. In table A7-A9, the cluster level is narrowed down to province-Kabupaten. We lost some significance but still have a quite consistent outcome with bigger effects. For example, we obtain that one more year of schooling decreases impatience by about 7 percentage points with strong instrument variable (IV F statistics is 11.52). However, still, we don't find any significant effects for time preference A and risk preference A.

As described in the empirical strategy section, our estimate may be biased if the treatment effect could be affected by region specific trends or factors. We examine the results by following the specifications that Duflo (2001) used. The regression includes three interaction terms additionally as mentioned in the previous section. The empirical results are shown in the appendix table A10-A12. As we focus on time preference B so far, we want to analyze the result on time preference B. In the table A10, some of the specifications do not satisfy the first stage significance (IV F-statistics ¡ 10), and no significant results in the second stage are found. However, the sign is very consistently matched to the main results and the magnitude is comparable to the main and other specification results. For the robustness check, we also use different level of rainfall shocks that are calculated using stations data within 200km instead of 100km, producing very consistent results as the main results (Tables can be presented upon requests).

In sum, we have very consistent results across different specifications. We conclude in this section that education causally affects the formation of individual time preference. It has a meaningful policy implication in that an investment in education in developing countries can affect the formation of individual time preference so that individual saving behavior and health behavior that may be closely linked to individual time preference could be optimally adjusted. In the following section, we will look into behavior measure.

5.3 Behavior measure and factor analysis

From the previous main results, we suggest that education cases the formation of time preference but we don't have enough information to discern the risk preference formation. Several studies examined the relationship between individual preferences and behavior measure. For example. Smoking is considered as a risky behavior often driven by "affective" thinking (Oreopoulos and Salvanes (2011)). Saving is also considered as an investment that reflects time preferences(Fuchs (1980)). Self-employed generally have a riskier overall permanent income than wage earners so we could expect close relationships between self-employed and risk preference or time preference(Barsky et al. (1997)). We can easily imagine that persons who value future more are likely to get insurance (Barsky et al. (1997). Thus, first, we run the OLS regression of each behavior measure on individual preferences to see any associations in this study.

In table 9, we found that smoking is significantly correlated with time preference measure A. It seems that more patient people are likely to smoke. This is against our intuition and previous studies. However, it is important to note that only 3.16% of female are smoking. Or it is possible that more educated people are more likely to be exposed to social activities or job market opportunities so that the chance of smoking could be higher for these women. We can't find the exact mechanisms here. We also check insurance and self employ and found some significance. More patient people are likely to join insurance and more impatient people are likely to be self employed. In table 10, we also use saving, log of saving amount and community join as behavior measures. We found the significant correlation between time preference and community join. Those who are patient are likely to participate in community activities. Since we have very small sample size for log of saving amount, it is hard to derive implications from this estimation results.

In order to overcome some potential measurement error problems, we take an advantage of correlations between behavior measures and time preference A and B (All explanations and estimation results are presented in the appendix). Based on these correlation, we apply factor analysis by assuming that behavior measures and time preference measures share true preference parts. First, we assume that time preference A and time preference B has a common factor, which we thought is the true time preference. In order to estimate the true time preference, we assume that time preference measure A and time preference measure B is connected only through this true time preference. We factorize two measures and extract the true time preference measure using Bartlett scores (Bartlett (1938)). Table A13 shows the results of factor analysis using only two factors, time preference measure A and time preference measure B. In column 4, we found the significant effect of education on time preference. Even though the interpretation of magnitude is nontrivial (because this is a factor score), the negative sign suggests the similar interpretation as the main results. In addition, we combine all significant behavior measures with time preference measure A and time preference measure B (Table A14). We factorize insurance, community participation, self employment, smoking together with time preference measure A and time preference measure B. We assume that all these measures have a common factor, that is, true time preference and no other factors would be linked each other. Across different specifications, some have significant effects and some have no significant effects but we observe similar magnitude of coefficients for female. Further, some of specifications seem to pass the weak IV problems. Again, the magnitude is not directly comparable to the main results because it is a composite measure. However, the result supports the previous findings that education causes patience.

5.4 Bound Analysis

In addition to factor analysis, we also examine the effects by applying bound analysis to minimize the measurement error concerns. We calculated boundary estimates from the equations (3)-(5) in the appendix. We present bound analysis results in table A15 -A18. Since the bound estimate is higher for the patient people, the sign of the coefficients should be opposite to the main results to be consistent. As expected, in table A15, we find the significant effects for female. Using all fixed effects including year of birth and year of birth quadratic (column 2-3), we observe the same pattern as the main results. Consistent with the main results and several robustness checks, no significant results have been found for time preference measure A and risk preference measure A for female.

5.5 Reduced Form Effects

In the appendix table A19-A21, we report the estimates of the reduced form effects of individual preferences on the INPRES intensity, an instrument variable for year of schooling. We found statistically significant effects for time preference B but not for time preference A and risk preference A. This is consistent results as we saw in the previous main results. Given the fact that the point estimate in the reduced form equations are arithmetically identical to the point estimates from the traditional second stage IV-2SLS multiplied by the point estimates from the first stage, we can expect the sign and magnitude of the coefficients. For example, we found the significant effect in the second stage for time preference B in table 8 (-0.0640 in column (6)). When we multiply this point estimate by the first stage coefficient (0.3973), we have -0.0254 with significance as identified in column (6) in table A19.

6 Mechanism

In this section, we provide suggestive evidence for the potential mechanism variables that connects education and time preference. Since most mechanism variables are endogenously formed, it is challenging to conclude the interpretation about the mechanisms. Most variables such as cognitive skills (word recall), health status, Community participation, log per capita household-level expenditure, mental health, are possibly affected by education level, including those variables in the regressions to look into mechanisms could be misleading due to misspecification. Unless we have good instrument variables for each potential mechanism variable, deriving the causal interpretation is a tall order. In spite of the difficulties, we try to provide suggestive evidence by utilizing naive approaches. Table 11 and 12 describes the results about the mechanisms. We first run the main specifications by including potential mechanism variable one by one. Then, we compare the change in the main coefficients. In table 11, we observe a little change in coefficients across all specifications but the column 3 with cognition measure, word recall. Since the word recall coefficient is not significant, we only have suspicion about cognition as a potential mechanism variable. In order to further investigate this possibility, we run several regressions in table 12. In this regressions, we include word recall test variable in the all regressions and add other potential variables one by one. In column 5 and 7, we observe quite significant drop in the magnitude of main coefficients (schooling) after including both word recall variable and log PCE or community participation. While significance of log PCE and community participation is reduced, word recall variable stays very significant. This may indicate that the cognitive skills measured by word recall could be a key factor for the mechanism through which education affects time preference. When we include all health measures and word recall, we don't observe changes in the main coefficient, but we observe big changes in column 1 and 2. This also might suggest that health could also play a central role as a mechanism through which education affects time preference.

As we mentioned, this is very suggestive evidences. It is possible that unobservable factors affects both time preference and these potential mechanism variables. According to theories of choice bracketing (Amos Tversky and Daniel Kahneman 1981; Daniel Read, George Loewenstein, and Matthew Rabin 1999), those who have a better cognition are likely to be future-oriented and patient. According to Dohmen et al. (2010), emotions and cognition are both important factors in decision making. When cognition dominates emotions, individual is likely to be more risk neutral and to take a longer-run view. Based on these theoretical backgrounds, the empirical results here could contributed to this topic.

7 Discussion and Summary

We find significant correlation between education and individual preferences. In this paper, consistently, we find statistically significant causal effect of education on time preference. We also exploit several different specifications to support the main results. Factor analysis and bound analysis are also taken into account to minimize the measurement error. The results are pretty much the same across all specifications. If we compare the magnitude of OLS estimates and IV-Probit or IV-2SLS estimates, we realized that the magnitude of IV-Probit marginal effect or the magnitude of IV-2SLS estimates is about ten times greater than the the magnitude of OLS estimates. There may be several reasons for this. We suggest that omitted variables may cause this big difference. Omitted variables could be household wealth level, individual health status, individual's marital status, job status or parent's preference. These omitted variable bias may cause the underestimation of the effect of education on time preference when we run the OLS/FE. Unfortunately, we can't include some of observed omitted variables because they are potentially endogenous and are hugely affected by education level. We also argue that these omitted variables could be the mechanisms for the effect of education on individual preferences as we already examined in the results section. For example, education causes individuals to earn more and again this affects individual's preference as elicited in the model by Becker and Mulligan (1997). In addition, individuals are likely to have more promising job because of higher education, more promising work could provide a room for preparing the future and may affect the formation of individual preferences.

Our contribution to this literature is twofold. First, as far as we know, this is the first paper to document the causal relationship between education and individual preferences using primary schooling level in developing countries. In human capital production function, education is one of the most important factors. As our results suggest, if education causally affects time preference, the effect of education on human capital formation could be over or underestimated. Since time preference may affect a lot of behavior measure or decision makings that are related to human capital. Second, we provide a new evidence on the link between education and health. Individual preferences play a crucial role in determining health behavior, and this is affected by education. By suggesting the link from education to individual preferences, we reveal additional determinants of health capital and this help us understand the structural human capital production function.

For the future work, we are also going to estimate the effect of parent education on children's preferences using our sample cohort and their children. While Hryshko et al. (2011) identify the effect using compulsory education in United States, we found no studies identifying the effect of parent's education on children's preferences in the context of developing countries. Finally, we are going to check any causal link from health to individual preferences to explain the entire mechanisms. With IFLS5 (forthcoming), we could check the changes in individual preferences and find the potential factors for the formation of individual preferences. This is very promising work because only few studies try to look at this topic.

References

Adhvaryu, A., Molina, T., Nyshadham, A., and Tamayo, J. (2014). Recovering from early life trauma: Dynamic substitution between child endowments and investments. *Unpublished Manuscript*.

Aizer, A. and Cunha, F. (2012). The production of human capital: Endowments, investments and fertility. Technical report, National Bureau of Economic Research.

Anderson, L. R. and Mellor, J. M. (2008). Predicting health behaviors with an experimental measure of risk preference. *Journal of Health Economics*, 27(5):1260–1274.

Arrow, K. J. (1965). *Aspects of the theory of risk-bearing*. Yrjö Jahnssonin Säätiö.

Barsky, R. B., Juster, F. T., Kimball, M. S., and Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 112(2):pp. 537–579.

Bartlett, M. S. (1938). Methods of estimating mental factors. *Nature*, 141(3570):609–610.

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

Behrman, J. R., Parker, S. W., and Todd, P. E. (2011). Do conditional cash transfers for schooling generate lasting benefits? a five-year followup of progresa/oportunidades. *Journal of Human Resources*, 46(1):93–122.

Cameron, A. C. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.

Cameron, L. and Shah, M. (2015). Risk-taking behavior in the wake of natural disasters. *Journal of Human Resources*, 50(2):484–515.

Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics*, 3:1801–1863.

Cassar, A., Healy, A., and Von Kessler, C. (2011). Trust, risk, and time preferences after a natural disaster: experimental evidence from thailand. *Unpublished manuscript*.

Chiteji, N. (2010). Time-preference, non-cognitive skills and well-being across the life course: Do non-cognitive skills encourage healthy behavior? *The American Economic Review*, 100(2):200.

Cramer, J. S., Hartog, J., Jonker, N., and Van Praag, C. M. (2002). Low risk aversion encourages the choice for entrepreneurship: an empirical test of a truism. *Journal of Economic Behavior and Organization*, 48(1):29–36.

Cutler, D. M. and Lleras-Muney, A. (2006). Education and health: evaluating theories and evidence. Technical report, National Bureau of Economic Research.

Cutler, D. M. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics*, 29(1):1–28.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *The American Economic Review*, 100(3):pp. 1238–1260.

Dohmen, T., Huffman, D., Schupp, J., Falk, A., Sunde, U., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):pp. 522–550.

Donkers, B., Melenberg, B., and Van Soest, A. (2001). Estimating risk attitudes using lotteries: A large sample approach. *Journal of Risk and Uncertainty*, 22(2):165–195.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *The American Economic Review*, 91(4):pp. 795–813.

Duflo, E. (2004). The medium run effects of educational expansion: Evidence from a large school construction program in indonesia. *Journal of Development Economics*, 74(1):163–197.

Field, E., Robles, O., and Torero, M. (2009). Iodine deficiency and schooling attainment in tanzania. *American Economic Journal: Applied Economics*, pages 140–169.

Frederick, S., Loewenstein, G., and O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, pages 351–401.

Fuchs, V. R. (1980). Time preference and health: an exploratory study.

Grossman, M. (2006). Education and nonmarket outcomes. *Handbook of the Economics of Education*, 1:577–633.

Guiso, L., Sapienza, P., and Zingales, L. (2013). Time varying risk aversion. Technical report, National Bureau of Economic Research.

Hamoudi, A. (2006). Risk preferences in households and families. *Unpublished Manuscript*.

Hanaoka, C., Shigeoka, H., and Watanabe, Y. (2015). Do risk preferences change? evidence from panel data before and after the great east japan earthquake. Working Paper 21400, National Bureau of Economic Research.

Heckman, J. J. (2000). Policies to foster human capital. *Research in economics*, 54(1):3–56.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, 104(33):13250–13255.

Hershberger, S. L. (2005). Factor score estimation. *Encyclopedia of statistics in behavioral science*.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *The American Economic Review*, 92(5):1644–1655.

Hryshko, D., Luengo-Prado, M. J., and Sørensen, B. E. (2011). Childhood determinants of risk aversion: The long shadow of compulsory education. *Quantitative Economics*, 2(1):37–72.

Leigh, J. P. (1986). Accounting for tastes: Correlates of risk and time preferences. *Journal of Post Keynesian Economics*, 9(1):pp. 17–31.

Lewbel, A., Dong, Y., and Yang, T. T. (2012). Comparing features of convenient estimators for binary choice models with endogenous regressors. *Canadian Journal of Economics/Revue canadienne d'économique*, 45(3):809–829.

Maccini, S. and Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *The American Economic Review*, 99(3):pp. 1006–1026.

Meier, S. and Sprenger, C. D. (2015). Temporal stability of time preferences. *Review of Economics and Statistics*, 97(2):273–286.

Miguel, E. and Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217.

Ng, J. (2013). Elicited risk and time preferences: the role of demographics, cognition, and interviewers. *University of Southern California Job Market Paper*.

Oreopoulos, P. and Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *The Journal of Economic Perspectives*, pages 159–184.

Perez-Arce, F. (2011). The effect of education on time preferences. *RAND working paper series WR-844*.

Peters, K. (1990). *Indonesia: strategy for a sustained reduction in poverty*. Washington: World Bank.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica: Journal of the Econometric Society*, pages 122–136.

Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, pages 76–90.

Strauss, J., Witoelar, F., Sikoki, B., and Wattie, A. M. (2009). The fourth wave of the indonesia family life survey: Overview and field report.

Table 3: Summary Statistics

Variable	**Mean**	**Std. Dev.**	**N**
Schooling	7.561	4.306	6329
Gender	0.488	0.5	6843
Age	45.178	7.216	6843
Muslim (dummy)	0.892	0.311	6738
JAWA (dummy)	0.436	0.496	6742
Mother Schooling	5.75	2.675	2292
Father Schooling	6.453	3.039	3079
Never married	0.03	0.171	6843
Smoking	0.35	0.477	6740

NOTE: Muslim indicates the proportion of people who are muslim. JAWA is the most popular ethnicity in Indonesia.

Table 4: Relationship between time preferenceB and education

VARIABLES	(1) timeprefb	(2) timeprefb	(3) timeprefb	(4) timeprefb	(5) timeprefb	(6) timeprefb
Schooling	-0.0052*	-0.0072*	-0.0038**	-0.0051*	-0.0072**	-0.0032*
	(0.0025)	(0.0035)	(0.0017)	(0.0026)	(0.0034)	(0.0019)
Observations	4,443	2,298	2,145	4,443	2,298	2,185
R-squared	0.0860	0.1211	0.1401	0.0879	0.1266	0.1397
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	YES	YES	YES	YES	YES	YES
Religion FE	YES	YES	YES	YES	YES	YES
YOB FE	NO	NO	NO	YES	YES	YES
YOB, YOBSQ	YES	YES	YES	NO	NO	NO
Sample	Pooled	Male	Female	Pooled	Male	Female

NOTE: Timeprefb is a dummy variable, which is equal to 1 if the respondent is the most impatient person, is equal to 0 otherwise. All regressions include a dummy for mother, father education, urban dummy. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Relationship between time preferenceA and education

VARIABLES	(1) timeprefa	(2) timeprefa	(3) timeprefa	(4) timeprefa	(5) timeprefa	(6) timeprefa
Schooling	-0.0095***	-0.0108***	-0.0072***	-0.0094***	-0.0105***	-0.0069**
	(0.0022)	(0.0031)	(0.0025)	(0.0023)	(0.0031)	(0.0025)
Observations	4,490	2,322	2,168	4,490	2,322	2,168
R-squared	0.0920	0.1467	0.1416	0.0965	0.1536	0.1564
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	YES	YES	YES	YES	YES	YES
Religion FE	YES	YES	YES	YES	YES	YES
YOB FE	NO	NO	NO	YES	YES	YES
YOB, YOBSQ	YES	YES	YES	NO	NO	NO
Sample	Pooled	Male	Female	Pooled	Male	Female

NOTE: Timeprefa is a dummy variable, which is equal to 1 if the respondent is the most impatient person, is equal to 0 otherwise. All regressions include a dummy for mother, father education, urban dummy. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table 6: IV-probit model-Time preferenceB

VARIABLES	(1) timeprefb	(2) timeprefb	(3) timeprefb	(4) timeprefb	(5) timeprefb
Schooling	-0.1199	-0.2257***	-0.2067**	-0.2054***	-0.1756**
	(0.0987)	(0.0364)	(0.0908)	(0.0459)	(0.0775)
Marginal Effect	-0.0309	-0.0606***	-0.0554**	-0.0612***	-0.0490*
	(0.0265)	(0.0101)	(0.0253)	(0.0090)	(0.0262)
First stage coefficient on IV	0.2670*	0.4440***	0.2967***	0.4682***	0.4733***
	(0.1465)	(0.1333)	(0.1032)	(0.0849)	(0.0774)
Sample	Pooled	Female	Female	Female	Female
Observations	4,252	1,908	1,908	1,499	1,499
Mean DV	0.8130	0.8008	0.8008	0.7985	0.7985
SEASON FE	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	NO	NO	NO
JAWA	YES	YES	YES	YES	YES
Religion FE	YES	YES	YES	YES	YES
Rain	NO	NO	NO	YES	YES
YOB FE	NO	NO	YES	NO	YES
YOB, YOBSQ	YES	YES	NO	YES	NO
P value of Wald Test	0.365	0.000	0.151	0.007	0.114

NOTE: IV-Probit is not an IV estimator. It is a control function. If we can't reject the wald test of exogeneity (null hypothesis of no endogeneity), there is not sufficient information in the sample to reject the null, so a regular probit regression may be appropriate. JAWA is an indicator for Jawa ethnicity, which comprises about 50% of ethnicity in Indonesia. When we use ethnicity fixed effect, some of MLE estimates does not converge. Standard errors are clustered at province level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7: IV-probit model-Time preferenceA

VARIABLES	(1) timeprefa	(2) timeprefa	(3) timeprefa	(4) timeprefa	(5) timeprefa
Schooling	-0.1284	-0.2078	-0.1896	-0.1114	-0.0459
	(0.1758)	(0.1293)	(0.2281)	(0.1508)	(0.1281)
Marginal Effect	-0.0383	-0.0597*	-0.0538	-0.0565	-0.0327
	(0.0505)	(0.0339)	(0.0645)	(0.0432)	(0.1115)
First stage coefficient on IV	0.2263*	0.3968***	0.2632**	0.4850***	0.4539***
	(0.1222)	(0.1507)	(0.1191)	(0.0826)	(0.0800)
Sample	Pooled	Female	Female	Female	Female
Observations	4,406	2,003	2,003	1,595	1,595
Mean DV	0.7224	0.7239	0.7239	0.7229	0.7229
SEASON FE	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	NO	NO	NO
JAWA	YES	YES	YES	YES	YES
Religion FE	YES	YES	YES	YES	YES
Rain	NO	NO	NO	YES	YES
YOB FE	NO	NO	YES	NO	YES
YOB, YOBSQ	YES	YES	NO	YES	NO
P value of Wald Test	0.613	0.319	0.582	0.585	0.578

NOTE: IV-Probit is not an IV estimator. It is a control function. If we can't reject the wald test of exogeneity (null hypothesis of no endogeneity), there is not sufficient information in the sample to reject the null, so a regular probit regression may be appropriate. When we use ethnic fixed effect, the MLE estimates does not converge. Instead we use JAWA indicator. Standard errors are clustered at province level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: The causal effect of education on time preferenceB - 2SLS

VARIABLES	(1) timeprefb	(2) timeprefb	(3) timeprefb	(4) timeprefb	(5) timeprefb	(6) timeprefb
Schooling	-0.0461**	-0.0319	-0.0628**	-0.0712***	-0.0697***	-0.0640**
	(0.0166)	(0.0226)	(0.0279)	(0.0192)	(0.0158)	(0.0311)
First stage coefficient on IV	0.4900***	0.4298***	0.4269***	0.4591***	0.4868***	0.3973***
	(0.0797)	(0.0974)	(0.0949)	(0.1036)	(0.0860)	(0.1153)
Sample	Female	Female	Female	Female	Female	Female
Observations	1,732	1,732	1,732	1,732	1,732	1,732
IV F-stat	37.76	19.48	20.22	19.62	32.04	11.88
Mean DV			0.8279			
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	YES	NO	YES
Religion FE	YES	YES	YES	NO	YES	NO
Rain	YES	YES	YES	YES	YES	YES
YOB FE	NO	YES	YES	NO	NO	YES
YOB, YOBSQ	YES	NO	NO	YES	YES	NO

NOTE: Instrument variable for year of schooling is INPRES intensity. All regressions include a dummy for mother, father education, urban dummy. Mean of dependent variable is 0.8279. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Relationship between behavior and preferences for female

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	smoking	smoking	smoking	insurance	insurance	insurance	selfemploy	selfemploy	selfemploy
timeprefa	-0.0147**			-0.0364			0.0502*		
	(0.0067)			(0.0222)			(0.0285)		
timeprefb		0.0005			-0.0512**			0.0401	
		(0.0088)			(0.0191)			(0.0493)	
riskavera			-0.0188			0.0022			-0.0648
			(0.0108)			(0.0278)			(0.0429)
Observations	2,400	2,419	1,260	2,400	2,419	1,260	1,692	1,705	900
R-squared	0.1785	0.1747	0.2247	0.1830	0.1789	0.2597	0.1417	0.1396	0.2421

NOTE: All regressions include season FE, a dummy for mother and father education, religion FE, urban dummy, ethnics FE, prokab FE. Robust standard errors are calculated at the province level. All regressions examine female only. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 10: Relationship between behavior and preferences for female

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	saving	saving	saving	logsavingamount	logsavingamount	logsavingamount	comjoin	comjoin	comjoin
timeprefa	-0.0352			-0.2756			-0.0490**		
	(0.0297)			(0.6266)			(0.0217)		
timeprefb		0.0011			-0.6535			-0.0285	
		(0.0433)			(0.4051)			(0.0318)	
riskavera			-0.0017			-0.6098			-0.0577
			(0.0398)			(0.5941)			(0.0332)
Observations	1,556	1,569	835	314	313	180	2,400	2,419	1,260
R-squared	0.2076	0.2100	0.2940	0.5711	0.5708	0.6921	0.2536	0.2500	0.3284

NOTE: All regressions include season FE, a dummy for mother and father education, religion FE, urban dummy, ethnics FE, prokab FE. Robust standard errors are calculated at the province level. All regressions examine female only. *** p<0.01, ** p<0.05, * p<0.1

Table 11: Mechanisms Check (Female)

VARIABLES	(1) timeprefb	(2) timeprefb	(3) timeprefb	(4) timeprefb	(5) timeprefb	(6) timeprefb	(7) timeprefb
Schooling	-0.0685**	-0.0648**	-0.0481***	-0.0613**	-0.0693**	-0.0650**	-0.0647**
	(0.0310)	(0.0294)	(0.0180)	(0.0254)	(0.0338)	(0.0287)	(0.0299)
Self health	0.0596**						
	(0.0248)						
Health compared to others		0.0304					
		(0.0306)					
Number of Word Recall			0.0154				
			(0.0099)				
Community participation				0.0630*			
				(0.0341)			
log PCE					0.0751		
					(0.0466)		
Life Quality (SW)						0.0237**	
						(0.0119)	
Mental health							-0.0227***
							(0.0058)
Original Coefficient	-0.0628**	-0.0628**	-0.0586**	-0.0628**	-0.0628**	-0.0629**	-0.0628**
	(0.0279)	(0.0279)	(0.0241)	(0.0279)	(0.0279)	(0.0279)	(0.0279)
IV F-stat	17.26	19.78	23.18	19.88	13.33	18.90	19.53
Observations	1,732	1,732	1,679	1,732	1,732	1,729	1,732
Mean DV				0.8249			

NOTE: Self health is a dummy variable, self health report. Health compared to others is a dummy variable, self health report compared to others. Number of Word recall is the total correct number of Word recall test. Community participation is a dummy variable for community participation. log PCE is the log amount of per household expenditure. Life Quality is a dummy variable for self assessment of life quality. Mental health is the total number of mental problems. Robust standard errors are clustered at the province level.*** p<0.01, ** p<0.05, * p<0.1

Table 12: Mechanisms Check (Female)- Cognition

VARIABLES	(1) timeprefb	(2) timeprefb	(3) timeprefb	(4) timeprefb	(5) timeprefb	(6) timeprefb	(7) timeprefb
Schooling	-0.0595***	-0.0814***	-0.0683***	-0.0727***	-0.0613***	-0.0669***	-0.0575***
	(0.0129)	(0.0207)	(0.0155)	(0.0159)	(0.0125)	(0.0140)	(0.0110)
Self health		0.0403**	0.0414**				
		(0.0170)	(0.0175)				
Life Quality		0.0285***	0.0204***				
		(0.0091)	(0.0062)				
Mental health		-0.0200***	-0.0190***				
		(0.0077)	(0.0062)				
log PCE				0.0800***	0.0537**		
				(0.0296)	(0.0222)		
Community participation						0.0838***	0.0654**
						(0.0301)	(0.0281)
Number of Word Recall	0.0194***		0.0213***		0.0189***		0.0183***
	(0.0059)		(0.0067)		(0.0061)		(0.0056)
Original Coefficient	-0.0703***	-0.0703***	-0.0703***	-0.0697***	-0.0697***	-0.0697***	-0.0697***
	(0.0171)	(0.0171)	(0.0171)	(0.0168)	(0.0168)	(0.0168)	(0.0168)
IV F-stat	28.62	15.18	26.73	20.04	29.09	17.75	26.99
Observations	1,676	1,676	1,676	1,679	1,679	1,679	1,679

NOTE: All regressions control for season FE, district FE, ethnicity FE, religion FE. We also control for log rainfall and year of birth and quadratic of year of birth. Robust standard errors are clustered at the province level.*** $p<0.01$, ** $p<0.05$, * $p<0.1$

A Relative risk aversion - true risk preference measure

Following Barsky et al. (1997), we use relative risk aversion rate to estimate level of risk aversion with the following constant relative risk aversion (CRRA) utility function, independently initiated by Pratt (1964) and Arrow (1965). The CRRA utility function is:

$$U(C) = \frac{C^{1-\rho} - 1}{1-\rho} \tag{3}$$

Given the CRRA utility function, a relative risk aversion rate can be defined as $\rho = -\frac{U''(C)C}{U'(C)}$. With the CRRA utility function used, it would be expected to assume that the level of risk aversion is constant regardless of amounts of consumption. If we assume the constant risk aversion, then we could have a bound estimate of relative risk aversion for each category. In our context, every individual was catergorized into four different groups, relying on the answer to which option one prefers. As already explained in the previous section, each individual is given the opportunity to choose one of the two alternatives; to receive a certain amount(c) or to take a gamble where one receive a twice larger amount($2c$) or a smaller amount(λc, $0 \leq \lambda < 1$) with probability of 50 percent, respectively. We can refer to the categorical responses as the consequence of maximizing expected utility with the CRRA function. Following Barsky et al. (1997), λ represents how much individuals cut their income with bad luck in a 50-50 gamble. For example, we have the questions asking either respondents choose RP 800 thousand or choose RP 400 thousand/ RP 1600 thousand by 50-50 chance. In this case, λ is going to be 0.5. If one chooses to take the risk , then by expected utility theory, the following inequality must hold.

$$0.5U(2c) + 0.5U(\lambda c) \geq U(c), \tag{4}$$

that is, the expected utility of accepting the 50-50 gamble exceed the one of receiving the constant income. Under CRRA utility function, the equation (3) shows the association between λ and relative risk aversion rate, ρ. Thus with the information on λ, we could calculate relative risk aversion rate(ρ)[4]:

$$\lambda = (2 - 2^{(1-\rho)})^{[1/(1-\rho)]} \tag{5}$$

Table A1: Relative Risk Aversion Rate

	λ	Upper bound	Lower bound	Midpoint	% of respondents
Category 4	-	∞	2.915	-	50.8
Category 3	0.75	2.915	1	1.958	8.5
Category 2	0.5	1	0.306	0.653	13.8
Category 1	0.25	0.306	0	0.153	27.0

Notes: Category 1 indicates the least risk averse or most risk taking, category 4 indicates the most risk averse.

[4] We could use absolute risk aversion rate as suggested by Cramer et al. (2002).

Table 3 shows the lower and upper bounds of the relative risk aversion rate for four different categories. With hypothetical questions and limited categories, we are not able to calculate the upper bound of category 4. As Holt and Laury (2002) documents that the effect of using a different method to measure risk preference is not large, using hypothetical setting represents well the real payment setting. In addition to this study, Dohmen et al. (2011) finds that answers to hypothetical questions are closely related to true measures. We support our analysis by not only using a dummy variable for risk preference (our main outcome variable) but also using risk aversion bound measure (Cramer et al. (2002); Hanaoka et al. (2015)) as well.

B Discounting Rate - True time preference measure

In order to calculate time discounting rate, we use Double Multiple Price Lists (MPL) with CRRA utility function. MPL uses the cutoff point at which respondents change their responses according to different amounts with different periods. For example, if a sooner reward, c_t, and a later reward, c_{t+1}, give the same to an individual, then discounting rate (DR) is calculated as $\delta = (\frac{U(c_t)}{U(c_{t+s})})^{(1/s)}$. It is important to note that we consider a case where an individual discount future value exponentially. Later we deal with how differently an individual discount the future once we introduce hyperbolic discounting rate. However, unlike experimental environment, we don't know the exact value making an individuals receive identical satisfaction (identical utilities) from either option.

Similar to the risk preference, each individual is asked 2 questions and categorized into 4 categories, based on the answer to each question. This again leads us to estimate the upper and lower bounds of discounting rate. In addition, it is worth to note that we need a curvature of the utility function (ρ) in order to estimate discounting rate as long as we use the CRRA utility function, and this is another place where problem lies. We assume that we could calculate the curvature of CRRA utility function from relative risk aversion estimates. However, estimating the boundary on relative risk aversion as we did in the previous section does not provide each individual's curvature of the utility functions. Once we use the category-specific estimates as a curvature, estimated discounting rates might be at odds with the responses to the survey. In order to overcome this potential problem, after we calculate time discounting rate with individual's curvature estimates, we assign the average of respondents discounting rate with respect to each category. In other words, every individual within the category is assigned the same discounting rates. The following equation shows how we estimate the discounting rate.

$$\delta_i = E[(\frac{c_{t,j}^{1-\rho} - 1}{c_{t+s,j}^{1-\rho} - 1})^{(1/s)} | j = i], i, j = 1, 2, 3, 4 \tag{6}$$

where i and j are an indicator of the category for each time preference question, $C_t$is the amount chosen, t represents time period and s indicates the time after t period.

Table 4 shows the upper and lower bounds on the discount rate for each category. We also include the midpoint in the table. Using the midpoint, the category 1 is around three times more patient than the category 4, and the category 2 is two times more patient than category 4. Similar to relative risk aversion rate, one important criticism of the calculating bounds is that the results are also dependent upon the hypothetical questions rather than real payment.

We also present the hyperbolic discounting rate since exponential discounting might be too strong an assumption(Frederick et al. (2002)). Table 4 shows the discounting rate estimated in a hyperbolic way. In other words, the next period (t+1) is discounted at a rate of $\beta\delta$ and s periods later than t (t+s) are discounted at a rate of ($\beta\delta^s$). The hyperbolic discounting reflects the fact that people are likely to discount more in circumstances in which they compare the current and the future than in circumstances in which they compare the future and future. In order to estimate hyperbolic discounting rate, we exploit the result from both the Exponential Time Preference A and B together with the following equation.

$$\delta_i = E[(\frac{c_{t,j}^{1-\rho} - 1}{c_{t+s,j}^{1-\rho} - 1)\beta\delta})^{(1/s)}|j = i], j = 1, 2, 3, 4 \tag{7}$$

where $\beta\delta$ in the equation is replaced by the midpoint of discounting rate estimated with the Exponential Time Preference A. While we estimate the hyperbolic discounting rate, individuals who reveal $\beta > 1$ were dropped for data quality purposes. Estimated hyperbolic discounting rates are ranged between estimation from the Exponential Time Preference A and B, except the category 4.

Table A2: Time Discounting Rate - Exponential and Hyperbolic discounting

	Exponential Time Preference A			Exponential Time Preference B			Hyperbolic Discounting A		
	Upper Bound	Lower Bound	Mid Point	Upper Bound	Lower Bound	Mid Point	Upper Bound	Lower Bound	Mid Point
Category 4	0.668	0	0.334	0.861	0	0.431	0.636	0	0.318
Category 3	0.731	0.665	0.693	0.903	0.851	0.877	0.780	0.726	0.753
Category 2	0.834	0.773	0.803	0.954	0.913	0.934	0.895	0.850	0.873
Category 1	1	0.801	0.900	1	0.943	0.971	1	0.874	0.937

C Factor Analysis - True time preference measure

For time preference measures, we also use factor analysis. We exploit the similar factor analysis model used in Aizer and Cunha (2012). Let $Z_{x,i,j,k}$ denote the k^{th} error-riden measure of individual preference (revealed preference) of a respondent i lived in Kabupaten j. We assume that $Z_{x,i,j,k}$ is comprised of true measure of individual preference $x_{i,j}$, by the following equation:

$$Z_{x,i,j,k} = \alpha_{x,k}x_{i,j} + \epsilon_{x,i,j,k} \tag{8}$$

In IFLS4, we studied several behaviors that might be closely linked to individual preference. We only focus on time preference for this analysis. We exploited five different measures of time preference such as time preference survey questions A and B, saving, arisan participation, and insurance. Thus, we have K=5 dimensions of time preference observations. The parameters $\alpha_{x,k}$ are the factor loadings and the error term $\epsilon_{x,i,j,k}$ represents the measurement error in $Z_{x,i,j,k}$. We assume that $\epsilon_{x,i,j,k}$ is independent from $\epsilon_{x,i,j,l}$, $l \neq k$, and $\epsilon_{x,i,j,k}$ is independent from $x_{i,j}$. Our goal is to estimate $\hat{x_{i,j}}$. We use the Bartlett Method (Bartlett (1938)) since

Bartlett factor scores produce unbiased estimates of true factor scores (Hershberger (2005)). The Bartlett factor score of $\hat{x_{i,j}}$ is defined as:

$$\hat{x_{i,j}} = (\alpha_x' \Theta_x \alpha_x)^{-1} \alpha_x' \Theta_x Z_{x,i,j}$$

Factor loadings and Bartlett score coefficients are presented in the Appendix. We could apply factor analysis to risk preference measure as we do it for time preference. However, since we have some doubts about the response to risk aversion B survey questions, it might be difficult to derive true risk preference if we use factor analysis. Thus, when we use factor analysis for risk preference, we might lose the most significant indicators (risk aversion B survey questions) leading to an incorrect estimation of Bartlett factor scores. This will remain as a future work. We suggest two different factor analysis model. One is using only time preference A and B as factors and another is using behavior measures that are proven to be closely related to time preference A and B, such as insurance, community participation, self employ and smoking.

D Limits

However, we admit that individual preference measures are not perfect. For example, time preference measure could be affected by liquidity issues when the respondents answer the questions.

Table A3: Relationship between risk preferenceA and education

VARIABLES	(1) riskavera	(2) riskavera	(3) riskavera	(4) riskavera	(5) riskavera	(6) riskavera
Schooling	-0.0033	-0.0070	0.0018	-0.0035	-0.0071	0.0003
	(0.0035)	(0.0048)	(0.0043)	(0.0036)	(0.0047)	(0.0043)
Observations	2,667	1,529	1,138	2,667	1,529	1,138
R-squared	0.1702	0.2166	0.2711	0.1752	0.2211	0.2848
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	YES	YES	YES	YES	YES	YES
Religion FE	YES	YES	YES	YES	YES	YES
YOB FE	NO	NO	NO	YES	YES	YES
YOB, YOBSQ	YES	YES	YES	NO	NO	NO
Sample	Pooled	Male	Female	Pooled	Male	Female

NOTE: Riskavera is a dummy variable, which is equal to 1 if the respondent is the most risk averse person, is equal to 0 otherwise. All regressions include a dummy for mother, father education, urban dummy. Robust standard errors are calculated at the province level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A4: IV-probit model-Risk preferenceA

VARIABLES	(1) riskavera	(2) riskavera
Schooling	-0.0181	-0.1037
	(0.1900)	(0.2840)
Marginal Effect	-0.008	-0.009
	(0.0604)	(0.1346)
First stage coefficient on IV	0.3901**	0.6401***
	(0.1894)	(0.1151)
Sample	Female	Female
Observations	1,052	785
Mean DV	0.5494	0.5529
SEASON FE	YES	YES
PROKAB FE	YES	YES
Ethnicity FE	NO	NO
JAWA	YES	YES
Religion FE	YES	YES
Rain	NO	YES
YOB FE	NO	YES
YOB, YOBSQ	YES	NO
P value of Wald Test	0.911	0.752

NOTE: IV-Probit is not an IV estimator. It is a control function. If we can't reject the wald test of exogeneity (null hypothesis of no endogeneity), there is not sufficient information in the sample to reject the null, so a regular probit regression may be appropriate. When we use ethnic fixed effect, the MLE estimates does not converge. Instead we use JAWA indicator. Standard errors are clustered at province level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A5: The causal effect of education on time preferenceA - 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	timeprefa	timeprefa	timeprefa	timeprefa	timeprefa	timeprefa
Schooling	-0.0266	0.0200	-0.0012	-0.0437	-0.0447	0.0032
	(0.0482)	(0.0441)	(0.0404)	(0.0429)	(0.0422)	(0.0428)
Sample	Female	Female	Female	Female	Female	Female
Observations	1,720	1,720	1,720	1,720	1,720	1,720
IV F-stat	35.38	20.12	20.52	17.92	28	12.30
Mean DV	0.7441					
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	YES	YES	YES
Religion FE	YES	YES	YES	NO	NO	NO
Rain	YES	YES	YES	YES	YES	YES
YOB FE	NO	YES	YES	NO	NO	YES
YOB,YOBSQ	YES	NO	NO	YES	YES	NO

NOTE: Instrument variable for year of schooling is INPRES intensity. All regressions include a dummy for mother, father education, urban dummy. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table A6: The causal effect of education on risk preferenceA - 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	riskavera	riskavera	riskavera	riskavera	riskavera	riskavera
Schooling	-0.0056	-0.0381	-0.0538	-0.0587	-0.0165	-0.1047
	(0.0685)	(0.0996)	(0.1066)	(0.0649)	(0.0783)	(0.0884)
Sample	Female	Female	Female	Female	Female	Female
Observations	873	873	873	873	873	873
IV F-stat	10.70	16.44	4.228	2.954	5.312	2.178
Mean DV	0.5521					
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	YES	YES	YES
Religion FE	YES	YES	YES	NO	NO	NO
Rain	YES	YES	YES	YES	YES	YES
YOB FE	NO	YES	YES	NO	NO	YES
YOB,YOBSQ	YES	NO	NO	YES	YES	NO

NOTE: Instrument variable for year of schooling is INPRES intensity. All regressions include a dummy for mother, father education, urban dummy. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table A7: Robustness check for time preferenceB - clustered at prokab level

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	timeprefb	timeprefb	timeprefb	timeprefb	timeprefb	timeprefb
Schooling	-0.1024**	-0.0697**	-0.1041	-0.0628*	-0.1108	-0.0640
	(0.0475)	(0.0333)	(0.0877)	(0.0352)	(0.0961)	(0.0392)
Sample	Female	Female	Female	Female	Female	Female
Observations	2,145	1,732	2,145	1,732	2,145	1,732
IV F-stat	8.300	11.52	2.440	8.774	2.257	7
Mean DV	0.8233	0.8279	0.8233	0.8279	0.8233	0.8279
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	YES	YES	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	NO	YES	YES
Religion FE	YES	YES	YES	YES	NO	NO
Rain	NO	YES	NO	YES	NO	YES
YOB FE	NO	NO	YES	YES	YES	YES
YOB, YOBSQ	YES	YES	NO	NO	NO	NO

NOTE: Instrument variable for year of schooling is INPRES intensity. All regressions include a dummy for mother, father education, urban dummy. Robust standard errors are calculated at the prokab level. *** p<0.01, ** p<0.05, * p<0.1

Table A8: Robustness check for time preferenceA - clustered at prokab level

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	timeprefa	timeprefa	timeprefa	timeprefa	timeprefa	timeprefa
Schooling	-0.0998	-0.0447	-0.1056	-0.0012	-0.1055	0.0032
	(0.0717)	(0.0422)	(0.1397)	(0.0433)	(0.1498)	(0.0469)
Sample	Female	Female	Female	Female	Female	Female
Observations	2,128	1,720	2,128	1,720	2,128	1,720
IV F-stat	6.545	10.74	2.103	8.958	1.991	7.212
Mean DV	0.7392	0.7441	0.7392	0.7441	0.7392	0.7441
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	YES	YES	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	NO	YES	YES
Religion FE	YES	YES	YES	YES	NO	NO
Rain	NO	YES	NO	YES	NO	YES
YOB FE	NO	NO	YES	YES	YES	YES
YOB, YOBSQ	YES	YES	NO	NO	NO	NO

NOTE: Instrument variable for year of schooling is INPRES intensity. All regressions include a dummy for mother, father education, urban dummy. Robust standard errors are calculated at the prokab level. *** p<0.01, ** p<0.05, * p<0.1

Table A9: Robustness check for risk preferenceA - clustered at prokab level

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	riskavera	riskavera	riskavera	riskavera	riskavera	riskavera
Schooling	-0.0198	-0.0165	-0.2201	-0.0538	-0.3017	-0.1047
	(0.0946)	(0.0838)	(0.3284)	(0.1047)	(0.4356)	(0.1135)
Sample	Female	Female	Female	Female	Female	Female
Observations	1,119	873	1,119	873	1,119	873
IV F-stat	2.389	3.114	0.444	2.838	0.388	1.950
Mean DV	0.5514	0.5521	0.5514	0.5521	0.5514	0.5521
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	YES	YES	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	NO	YES	YES
Religion FE	YES	YES	YES	YES	NO	NO
Rain	NO	YES	NO	YES	NO	YES
YOB FE	NO	NO	YES	YES	YES	YES
YOB, YOBSQ	YES	YES	NO	NO	NO	NO

NOTE: Instrument variable for year of schooling is INPRES intensity. All regressions include a dummy for mother, father education, urban dummy. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table A10: The causal effect of education on time preferenceB - 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	timeprefb	timeprefb	timeprefb	timeprefb	timeprefb	timeprefb
Schooling	-0.0480	-0.0357	-0.0795	-0.0885	-0.0884	-0.0788
	(0.0454)	(0.0429)	(0.0503)	(0.0636)	(0.0541)	(0.0586)
First stage coefficient on IV	0.5275***	0.5527***	0.4820**	0.4646*	0.5026**	0.4485*
	(0.1566)	(0.1510)	(0.1915)	(0.2195)	(0.1961)	(0.2136)
Sample	Female	Female	Female	Female	Female	Female
Observations	1,612	1,612	1,612	1,612	1,612	1,612
IV F-stat	13.47	13.39	6.339	4.481	6.567	4.409
Mean DV			0.8294			
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	YES	NO	YES
Religion FE	YES	YES	YES	NO	YES	NO
Rain	YES	YES	YES	YES	YES	YES
YOB FE	NO	YES	YES	NO	NO	YES
YOB, YOBSQ	YES	NO	NO	YES	YES	NO

NOTE: We include the additional controls such as the interaction of treatment dummy and children population in 1971, interaction of treatment dummy and enrollment rate in 1971, interaction of treatment dummy and water sanitation program. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table A11: The causal effect of education on time preferenceA - 2SLS

VARIABLES	(1) timeprefa	(2) timeprefa	(3) timeprefa	(4) timeprefa	(5) timeprefa	(6) timeprefa
Schooling	0.0190	0.0445	0.0183	-0.0034	-0.0104	0.0266
	(0.0474)	(0.0416)	(0.0454)	(0.0510)	(0.0467)	(0.0488)
First stage coefficient on IV	0.5497***	0.5169***	0.4542*	0.4524*	0.4833**	0.4275*
	(0.1709)	(0.1696)	(0.2206)	(0.2431)	(0.2218)	(0.2405)
Sample	Female	Female	Female	Female	Female	Female
Observations	1,601	1,601	1,601	1,601	1,601	1,601
IV F-stat	10.34	9.288	4.240	3.463	4.749	3.161
Mean DV			0.7495			
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	YES	NO	YES
Religion FE	YES	YES	YES	NO	YES	NO
Rain	YES	YES	YES	YES	YES	YES
YOB FE	NO	YES	YES	NO	NO	YES
YOB, YOBSQ	YES	NO	NO	YES	YES	NO

NOTE: We include the additional controls such as the interaction of treatment dummy and children population in 1971, interaction of treatment dummy and enrollment rate in 1971, interaction of treatment dummy and water sanitation program. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table A12: The causal effect of education on risk preferenceA - 2SLS

VARIABLES	(1) riskavera	(2) riskavera	(3) riskavera	(4) riskavera	(5) riskavera	(6) riskavera
Schooling	0.0398	0.0589	0.0778	0.0075	0.0486	0.0411
	(0.0728)	(0.0593)	(0.0899)	(0.0877)	(0.1027)	(0.0729)
First stage coefficient on IV	0.6705*	0.7904*	0.6586	0.5363	0.5517	0.6692
	(0.3765)	(0.3946)	(0.4203)	(0.4695)	(0.4101)	(0.4966)
Sample	Female	Female	Female	Female	Female	Female
Observations	822	822	822	822	822	822
IV F-stat	3.172	4.012	2.455	1.305	1.810	1.816
Mean DV			0.5438			
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	YES	NO	YES
Religion FE	YES	YES	YES	NO	YES	NO
Rain	YES	YES	YES	YES	YES	YES
YOB FE	NO	YES	YES	NO	NO	YES
YOB, YOBSQ	YES	NO	NO	YES	YES	NO

NOTE: We include the additional controls such as the interaction of treatment dummy and children population in 1971, interaction of treatment dummy and enrollment rate in 1971, interaction of treatment dummy and water sanitation program. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table A13: Factor Analysis 1 - time preference

VARIABLES	(1) preference	(2) preferencef	(3) preferencef	(4) preferencef	(5) preferencef
Schooling	0.2149	-0.1501	-0.0660	-0.1203***	-0.0608
	(0.1623)	(0.1049)	(0.0553)	(0.0445)	(0.0542)
Sample	Pooled	Female	Female	Female	Female
Observations	3,612	2,158	1,746	1,746	1,746
SEASON FE	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	NO	YES	YES
Religion FE	YES	YES	YES	YES	YES
Rain	YES	NO	YES	YES	YES
YOB FE	NO	NO	NO	NO	YES
YOB, YOBSQ	YES	YES	YES	YES	NO
IV F-stat	4.200	12.28	28.02	22.58	17.87

NOTE: 'preference' is a composite measure of time preferenceA and time preferenceB by using factor analysis and pooled sample. 'preference' is a composite measure of time preference A and time preference B for female. Robust standrad errors are calculated at the province level.

Table A14: Factor Analysis 2 - time preference and behavior measure

VARIABLES	(1) factorpool	(2) factorfemale	(3) factorfemale	(4) factorfemale	(5) factorfemale	(6) factorfemale
Schooling	0.4215	-0.1352	-0.0539	-0.1183	-0.1100**	-0.0462
	(0.3075)	(0.0831)	(0.0666)	(0.0883)	(0.0559)	(0.1006)
Sample	Pooled	Female	Female	Female	Female	Female
Observations	3,001	1,498	1,214	1,498	1,214	1,214
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	NO	NO	YES	YES
JAWA	NO	YES	YES	NO	NO	NO
Religion FE	YES	YES	YES	YES	YES	YES
Rain	YES	YES	YES	NO	YES	YES
YOB FE	NO	NO	NO	NO	NO	YES
YOB, YOBSQ	YES	YES	YES	YES	YES	NO
IV F-stat	1.644	9.012	7.733	10.39	5.757	2.896

NOTE: ‘factortpool’ and ’factor female’ are a composite measure of time preferenceA and time preferenceB using factor analysis. Insurance, community participation, self employ and smoking is a significant indicator for either time preferenceA or time preferenceB. Using these behavior measures we estimate common factors. Jawa is the dummy variable for the most popular ethnicity in Indonesia. Robust standard errors are clustered at province level.

Table A15: Bound Analysis-Time preferenceB

VARIABLES	(1) lboundtb	(2) lboundtb	(3) lboundtb	(4) lboundtb	(5) lboundtb
Schooling	-0.0275	0.0161**	0.0171**	0.0115	0.0248
	(0.0341)	(0.0079)	(0.0069)	(0.0107)	(0.0164)
Sample	Pooled	Female	Female	Female	Female
Observations	3,648	1,766	1,766	1,766	1,766
SEASON FE	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	NO	NO	YES
JAWA	NO	NO	YES	YES	NO
Religion FE	YES	YES	YES	YES	YES
Rain	YES	YES	YES	YES	YES
YOB FE	NO	NO	NO	YES	YES
YOB, YOBSQ	YES	YES	YES	NO	NO
IV F-stat	3.345	29.20	42.50	37.82	19.32

NOTE: The outcome variable ’lboundtb’ is the lower bound estimate for each category. Lower bound estimate for category1 is 0.971, lower bound estimate for cateogry2 is 0.934, lower bound estimate for category3 is 0.877, lower bound for category4 is 0.431. The bigger number implies more patience. Robust standard errors are calculated at the province level.*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A16: Bound Analysis-Time preferenceA

VARIABLES	(1) lboundta	(2) lboundta	(3) lboundta	(4) lboundta	(5) lboundta
Schooling	-0.0373**	0.0128	0.0150	-0.0074	-0.0004
	(0.0164)	(0.0244)	(0.0217)	(0.0199)	(0.0209)
Sample	Pooled	Female	Female	Female	Female
Observations	3,626	1,754	1,754	1,754	1,754
SEASON FE	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	NO	NO	YES
JAWA	NO	NO	YES	YES	NO
Religion FE	YES	YES	YES	YES	YES
Rain	YES	YES	YES	YES	YES
YOB FE	NO	NO	NO	YES	YES
YOB, YOBSQ	YES	YES	YES	NO	NO
IV F-stat	4.668	29.92	43.20	39.59	19.07

NOTE: The outcome variable 'lboundta' is the lower bound estimate for each category. Lower bound estimate for category1 is 0.9, lower bound estimate for cateogry2 is 0.803, lower bound estimate for category3 is 0.693, lower bound for category4 is 0.334. The bigger number implies more patience. Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table A17: Bound Analysis using hyperbolic function-Time preferenceA

VARIABLES	(1) lboundth	(2) lboundth	(3) lboundth	(4) lboundth	(5) lboundth
Schooling	-0.0417**	0.0145	0.0170	-0.0086	-0.0008
	(0.0186)	(0.0276)	(0.0246)	(0.0222)	(0.0234)
Sample	Pooled	Female	Female	Female	Female
Observations	3,626	1,754	1,754	1,754	1,754
SEASON FE	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	NO	NO	YES
JAWA	NO	NO	YES	YES	NO
Religion FE	YES	YES	YES	YES	YES
Rain	YES	YES	YES	YES	YES
YOB FE	NO	NO	NO	YES	YES
YOB, YOBSQ	YES	YES	YES	NO	NO
IV F-stat	4.668	29.92	43.20	39.59	19.07

NOTE: The outcome variable 'lboundth' is the lower bound estimate for each category. Lower bound estimate for category1 is 0.937, lower bound estimate for cateogry2 is 0.873, lower bound estimate for category3 is 0.753, lower bound for category4 is 0.318. The bigger number implies more patience. Robust standard errors are calculated at the province level.*** p<0.01, ** p<0.05, * p<0.1

Table A18: Bound Analysis-Risk PreferenceA

VARIABLES	(1) lbound	(2) lbound	(3) lbound	(4) lbound	(5) lbound
Schooling	0.3315	-0.0168	-0.0019	-0.1124	-0.1749
	(0.7543)	(0.2007)	(0.1986)	(0.2811)	(0.3183)
Sample	Pooled	Female	Female	Female	Female
Observations	2,110	888	888	888	888
SEASON FE	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	NO	NO	YES
JAWA	NO	NO	YES	YES	NO
Religion FE	YES	YES	YES	YES	YES
Rain	YES	YES	YES	YES	YES
YOB FE	NO	NO	NO	YES	YES
YOB, YOBSQ	YES	YES	YES	NO	NO
IV F-stat	0.390	10.24	13.77	25	4.814

NOTE: The outcome variable 'lbound' is the lower bound estimate for each category. Lower bound estimate for category1 is 0, lower bound estimate for cateogry2 is 0.306, lower bound estimate for category3 is 1, lower bound for category4 is 2.915. he bigger number implies more risk averse. Robust standard errors are calculated at the province level.*** p<0.01, ** p<0.05, * p<0.1

Table A19: Reduced form estimation - time preference B

	(1) timeprefb	(2) timeprefb	(3) timeprefb	(4) timeprefb	(5) timeprefb	(6) timeprefb
INPRES intensity	-0.0226 *	-0.0137	-0.0268 *	-0.0327 **	-0.0339 ***	-0.0254*
	(0.0090)	(0.0110)	(0.0111)	(0.0082)	(0.0076)	(0.0109)
Sample	Female	Female	Female	Female	Female	Female
Observations	1732	1732	1732	1732	1732	1732
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	YES	NO	YES
Religion FE	YES	YES	YES	NO	YES	NO
Rain	YES	YES	YES	YES	YES	YES
YOB FE	NO	YES	YES	NO	NO	YES
YOB, YOBSQ	YES	NO	NO	YES	YES	NO

NOTE: Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table A20: Reduced form estimation - time preference A

	(1) timeprefb	(2) timeprefb	(3) timeprefb	(4) timeprefb	(5) timeprefb	(6) timeprefb
INPRES intensity	-0.0127	0.00862	-0.000528	-0.0195	-0.0210	0.00129
	(0.0253)	(0.0202)	(0.0194)	(0.0230)	(0.0232)	(0.0189)
Sample	Female	Female	Female	Female	Female	Female
Observations	1720	1720	1720	1720	1720	1720
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	YES	YES	YES
Religion FE	YES	YES	YES	NO	NO	NO
Rain	YES	YES	YES	YES	YES	YES
YOB FE	NO	YES	YES	NO	NO	YES
YOB,YOBSQ	YES	NO	NO	YES	YES	NO

NOTE: Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1

Table A21: Reduced form estimation - risk preference A

	(1) riskavera	(2) riskavera	(3) riskavera	(4) riskavera	(5) riskavera	(6) riskavera
treatnin	-0.00330	-0.0213	-0.0263	-0.0266	-0.00845	-0.0466
	(0.0468)	(0.0661)	(0.0707)	(0.0408)	(0.0494)	(0.0588)
Sample	Female	Female	Female	Female	Female	Female
Observations	873	873	873	873	873	873
SEASON FE	YES	YES	YES	YES	YES	YES
PROKAB FE	YES	YES	YES	YES	YES	YES
Ethnicity FE	NO	NO	YES	YES	YES	YES
Muslim Dummy	NO	NO	NO	YES	YES	YES
Religion FE	YES	YES	YES	NO	NO	NO
Rain	YES	YES	YES	YES	YES	YES
YOB FE	NO	YES	YES	NO	NO	YES
YOB,YOBSQ	YES	NO	NO	YES	YES	NO

NOTE: Robust standard errors are calculated at the province level. *** p<0.01, ** p<0.05, * p<0.1