Patience, Risk-Taking, and Human Capital Investment across Countries

Eric A. Hanushek, Lavinia Kinne, Philipp Lergetporer, and Ludger Woessmann

Online Appendix

September 29, 2021

Content

Appendix A: Conceptual Framework

Our analysis of international differences in test scores is motivated by a desire to understand how systematic differences in national preferences contribute to variations in human capital across nations. This appendix provides a conceptual framework that discusses how national preferences related to intertemporal choices enter the human capital production model. We start by depicting educational choices in a human capital investment model with intertemporal preferences, incorporating several prior lines of inquiry into human capital investments (section A.1). We then focus on the production of skills in order to understand how national preferences affect the individual and public choices of inputs into the education production function and the ultimate set of skills (section A.2). Finally, we provide a deeper discussion of how patience and risk-taking enter separately and jointly into intertemporal decision-making (section A.3).

A.1 Education as Intertemporal Choice

Educational decisions are fundamentally an intertemporal choice: initial investments of time, effort, and resources are set against expected future gains. Early human capital models thus directly related educational returns and investments to the rate at which future earnings are discounted (Mincer (1958, 1974); Becker (1964)). Further development of modern human capital theory naturally moved to optimal investment decisions of individuals, focusing on the maximization of lifetime earnings and stressing the time dimension of investments (Ben-Porath (1967, 1970); Heckman (1976); Rosen (1976)). The focus on a representative individual with perfect foresight precluded any deeper consideration of individual differences in intertemporal preferences. Given the intertemporal optimization decision, however, the two preference components related to balancing the present and the future – time and risk preferences – are crucial in understanding individual educational choices.

Surprisingly little explicit attention has been given to individual willingness to postpone gratification captured in patience, even though it is obviously a key element in the educational investment decision and thus in the earnings distribution. Detailed consideration of risk, by contrast, has entered human capital modeling at least since the contributions by Weiss (1972) and Levhari and Weiss (1974).[1]

[1] While the interaction of risk and human capital investment has been previously considered, attention has been confined mostly to labor-market outcomes. In considering different forms of labor-market risks such as variations in wages and employment, the optimal investment literature has produced varying predictions on how risk attitudes

The models of optimal human capital investment almost always focus on decisions about the quantity of education, which becomes the measure of individual skills. This focus has been natural given the availability of data and the consistency with the view of human capital investment as one of time. The perspective has been extraordinarily successful: The basic lifetime earnings model of Mincer (1974) has made years of schooling virtually synonymous with human capital in a wide range of empirical studies. Yet, school quantity is an imperfect measure of the underlying skill development that prescribes the optimality of downstream quantitative decisions in models of skill formation (Cunha, Heckman, and Schennach (2010)) and that has future payoffs on the labor market (Hanushek and Woessmann (2008)).

A.2 Human Capital Investment, Educational Production, and National Preferences

Direct investigation of the production of skills has developed mostly separately from the study of optimal human capital investment (Hanushek (1986)). Research into skill development during the production stage focuses on what is actually learned, generally measured by achievement tests (rather than the time spent in school).[2] This research almost exclusively considers issues of technical efficiency of input usage and of the productivity of different inputs – without relation to human capital investment behavior.[3] But in reality, the education production function depicts how chosen inputs relate to human capital, as the observed proximate inputs to skill development are themselves the result of human capital investment decisions.

Further, even though the canonical human capital production model depicts skills as a function of family and school inputs, it is difficult to presume that these measured outcomes perfectly reflect the optimizing decisions of parents. The process of skill acquisition involves numerous actors – including the students themselves, their peers and friends, families, neighborhoods, teachers, school principals, and so on. Each presumably is optimizing over a different value function that may include different intertemporal preference parameters. Because

may affect human capital investment. Interestingly, because of its focus on labor-market outcomes, the analysis of risk in this literature has largely ignored how risk might also enter into the production process for skills. This separation of optimal individual investments from consideration of the underlying production process leads to considerable distortion in the analysis of the role of intertemporal preferences.

[2] Our focus on achievement scores does not imply a different interest from the school attainment work. We view the intermediate measures of adolescents' achievement as a good index of the ultimate skills of completed human capital investments. As an alternative, the Programme for International Assessment of Adult Competencies (PIAAC) measures the cognitive skills of adults, but analysis is hampered by limited country coverage.

[3] There are exceptions, for example, when the choices of parental investments are related to other inputs in the production function (Kim (2001); Todd and Wolpin (2003)).

of different assessments of the long-run value from human capital investments and different valuation of present versus future costs and payoffs, children may for example choose effort levels according to a preference for playing football or computer games over studying math in a way that diverges from what parents deem optimal in their maximization calculus.[4]

Importantly, many of the relevant educational investment decisions are actually made at the group level. How much to invest in school resources is usually publicly chosen at the municipal, state, or country level. Similarly, the institutional structures of school systems – features such as school accountability, autonomy, and choice which have been shown to matter greatly for student outcomes (Hanushek and Woessmann (2011); Woessmann (2016b)) – are decided upon at the group level, and in most countries at the national level. As a consequence, aggregate societal intertemporal preferences will affect many parts of the education production process, making the set of preferences shared by the group important.

Therefore, in our analysis we change the perspective from individual preferences to group preferences and their relation to national cultures. Guiso, Sapienza, and Zingales (2006), p. 23, define culture as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation." While different theoretical and empirical concepts and definitions exist (Alesina and Giuliano (2015)), relevant cultural values encompass the set of preferences shared by the group – including the intertemporal preferences that we deem important for educational choices.[5]

A key element of the existing cultural analyses is an emphasis on values that are transmitted persistently across generations (Bisin and Verdier (2011); Alesina and Giuliano (2014)).[6] This persistent transmission motivates our empirical strategy below that looks at measures of national preferences in the home countries of migrants. Empirically, looking at migrants living in the same residence country allows us to distinguish cultural factors from other features of the

[4] Children may also be less willing or able to solve the dynamic optimization problem, leading to behavioral biases that prevent them from pursuing their own long-run well-being (Lavecchia, Liu, and Oreopoulos (2016)).

[5] The concept of culture is related to the concepts of values, preferences, and personality traits and sometimes even subsumed in noncognitive skills, and the interrelations and distinctions between the concepts often remain vague. See Almlund et al. (2011) for an extensive discussion of the relationship between personality traits and preferences.

[6] Patience and risk-taking have been shown to correlate consistently between parents and their children at the individual level (e.g., Kosse and Pfeiffer (2012); Alan et al. (2017)).

residence country such as institutions and economies (Carroll, Rhee, and Rhee (1994); Giuliano (2007); Fernández and Fogli (2009); Figlio et al. (2019)).

Recognition of intergenerational transmission also suggests some care in the specification of empirical models because common family factors may reflect cultural features. Analyses of educational production functions – whether within or across countries – commonly include measures of parental education (e.g., Hanushek (1986); Hanushek and Woessmann (2011)). But if national cultures influence human capital investment, parents' realized educational patterns may proxy the culture of their country. As such, they may partially be bad controls in the study of national preferences because they absorb part of the influence of the cultural factors.

More generally, a country's national preferences may affect all inputs in the education production process – on both the family and the school side – as well as the overall productivity with which these inputs are transformed into educational outcomes. This conceptualization implies that analyses of the effect of national preferences on student achievement should use very parsimonious specifications of the vector of control variables contained in the education production function.

A.3 Patience, Risk-Taking, and their Interrelatedness

While national preferences can encompass a wide variety of common traits, our interest in intertemporal decisions related to educational investments leads us to focus on two specific preference components: time preferences and risk preferences.

Time Preferences. The central role of the discount rate in models of optimal investment in human capital implies that time preferences are a key element of choices about whether to invest additional time, effort, and resources in improving educational outcomes. Preferences for payoffs in different time periods are reflected in patience, the trait of having a low rate of time discounting. For example, students must consider whether to give up play time with friends today – the opportunity cost of studying in the afternoon – for higher rewards in the future, such as graduating from school with better grades or the opportunity to receive better-paying jobs.[7]

It is remarkable that empirical studies only recently have begun to link validated measures of time preferences among students directly to educational outcomes. For example, Sutter et al.

[7] As such, patience is closely related to similar concepts employed in the study of traits among children, such as the willingness to defer gratification as measured, e.g., by the famous "marshmallow test" (e.g., Mischel, Shoda, and Rodriguez (1989)), self-control (Moffitt et al. (2011)), or perseverance and grit (e.g., Duckworth et al. (2007)).

(2013) show that experimentally elicited measures of patience among Austrian children are significantly related to field behavior, including reduced violations of schools' code of conduct.[8] Using longitudinal Swedish data, Golsteyn, Grönqvist, and Lindahl (2014) find that adolescents' time preferences are associated with human capital investments and lifetime outcomes. Castillo, Jordan, and Petrie (2019) show that experimentally elicited measures of discount rates among students in a school district in the U.S. State of Georgia are significantly related to high school graduation. Similarly, Castillo et al. (2020) show that preschool children in Chicago who are more patient are less likely to receive disciplinary referrals when they are in school years later. Combining the Hofstede (1991) cultural measure with migrant students in Florida schools as well as with the PISA data, Figlio et al. (2019) show that students from cultures with greater long-term orientation perform better on several measures of educational achievement.[9] At the macro level, Galor and Özak (2016) and Sunde et al. (2021) show that time preferences are importantly related to economic and educational outcomes in the long run.[10]

Risk Preferences. Beginning with the empirical study of occupational choices by Weiss (1972) and the theoretical analysis in Levhari and Weiss (1974), a stream of studies of human capital investments explicitly introduced various components of uncertainty and risk. In a very general way, Levhari and Weiss (1974) consider a range of risky elements in labor-market investment decisions including future supply and demand conditions as well as knowledge of one's own ability, of how time and money convert into human capital, and of the quality of schools along the investment path. They show that it is not possible a priori to determine how risk affects human capital investment incentives, a conclusion reiterated in the extensive review by Benzoni and Chyruk (2015). For example, higher earnings variance in higher-educated jobs may give rise to a positive association between risk-taking preferences and investment in higher education (e.g., Hartog and Diaz-Serrano (2007, 2014)), whereas lower unemployment risk of

[8] Alan and Ertac (2018) and Alan, Boneva, and Ertac (2019) show that measures of patience and grit are malleable to classroom interventions.

[9] Mendez (2015) shows the potential relevance of cultural traits for student achievement using a principal component from eleven different value questions in the World Values Survey (WVS) with migrant students in seven host countries in PISA, but the approach does not delve into specific components of national preferences. Cordero et al. (2018) include WVS measures in efficiency measurement of school systems in PISA.

[10] While formulated from a different perspective, a recent literature suggests that student behavioral differences related to effort, care, motivation, and perseverance may impact country test scores (e.g., Borghans and Schils (2012); Balart, Oosterveen, and Webbink (2018); Akyol, Krishna, and Wang (2018); Gneezy et al. (2019); Zamarro, Hitt, and Mendez (2019)). These behavioral differences may in turn reflect underlying cultural differences. We return to the role of student test-taking effort in robustness analyses below.

higher-educated jobs (e.g., Woessmann (2016a)) may give rise to the opposite association. In contrast to the indeterminate nature of the impact of different forms of labor-market risks, the role of risk-taking is more clear-cut when directly considering student behavior in the human-capital production process: Drawing on insights from the economics of crime, Castillo, Jordan, and Petrie (2019) argue that risk-lovingness may deter educational effort by favoring misbehavior in adolescence if there is uncertainty about getting caught by teachers or parents.

Existing empirical evidence on the association between risk and human capital investment is closely related to the specific components of risk considered in individual studies. Using U.S. data, Brown, Fang, and Gomes (2012) suggest that lower risk-taking leads to more investment in high-school education compared to less than high school but less investment in college compared to high school. Analyzing both wage and employment uncertainty, Groot and Oosterbeek (1992) find different results on returns by type of schooling (vocational or college) in the U.S. and the Netherlands, while Koerselman and Uusitalo (2014) find little differential effect of lifetime income variability on different schooling choices in Finland. Palacios-Huerta (2003) compares human capital risks to financial assets risk and detects wide variation in risk-adjusted rates of return. Using direct measures of children's risk preferences, Sutter et al. (2013) find little evidence of associations with field behavior. By contrast, Castillo, Jordan, and Petrie (2018, 2019) show a negative association of risk-taking preferences with high-school graduation and a positive with disciplinary referrals, in line with the notion that lower risk-taking may keep students out of trouble during the human-capital production process.[11]

The Interrelatedness of Time and Risk Preferences. While much of the prior literature has considered time and risk preferences separately, behavioral economics has emphasized their inherent interrelatedness: since only the present can be certain and the future always contains an element of uncertainty, it is inescapable that the two preference components are intertwined (Halevy (2008); Andreoni and Sprenger (2012)).[12]

An important implication of this interrelatedness is the need to control for the one preference component when studying the effect of the other. In fact, because many of the studies of risk-taking do not control for patience, this interrelationship may help explain the reasons for the

[11] There is also evidence of associations of patience and risk with intelligence among adults, again with mixed evidence on risk (Dohmen et al. (2010, 2018); Potrafke (2019)).

[12] Their particular formulation has been questioned, but the basic concept seems clear. See the exchange in Cheung (2015), Epper and Fehr-Duda (2015), Miao and Zhong (2015), and Andreoni and Sprenger (2015).

divergent empirical effects of risk on investment. Even more, given the a priori indeterminate direction of the effect of risk-taking in the human capital production function, the direction of bias when estimating the effect of patience without considering risk-taking is also unclear.

Appendix B: Additional Information on the Data

The information on the PISA and GPS datasets provided in this appendix complements the basic information contained in section 2 of the main text.

B.1 The Programme for International Student Assessment (PISA)

The target population of the representative random sampling of PISA are 15-year-old students, independent of grade level or educational track attended (OECD (2019)). The sampling in most countries proceeds in two steps. First, a random sample of schools that teach 15-year-old students is drawn using sampling probabilities that assure representativeness. Second, 35 students aged 15 years are randomly sampled in each school.[13] PISA only reports data for countries that meet the OECD's high sampling and data-quality standards.

To create comprehensive measures of students' competencies, PISA has students complete a broad array of tasks of varying difficulty in assessments that last for up to two hours. The testing mode was paper and pencil until 2012 and changed to computer-based testing in 2015. PISA achievement in math, science, and reading were standardized to a mean of 500 test-score points and a standard deviation of 100 test-score points for OECD-country students in wave 2000 (and rescaled on the same metric again in 2003 in math and in 2006 in science). We divide PISA scores by 100 throughout to express achievement in percent of a standard deviation. As a rule of thumb for interpreting PISA scores, about a quarter to a third of a standard deviation corresponds to the learning gains of one year of schooling (Woessmann (2016b)). Table 6 shows descriptive statistics of country-level PISA achievement in the three subjects.

In addition to achievement data, PISA elicits background information on student and family characteristics using student questionnaires, as well as contextual information on school resources and the institutional environment using school questionnaires completed by school principals. From these rich background data, we select core control variables – student gender, age, and migration status (first and second generation) – for our regression analysis. In addition, we derive measures of proximate inputs at the family, school, and institutional level that we use in our channel analysis. At the student level, these are parental education (six categories),

[13] We use the first plausible value (PV) provided by PISA throughout. Our results hold when considering other PVs and when employing estimation procedures that explicitly account for imputation and stratified sampling in the PISA data (results available upon request).

parental occupation (four categories), books at home (four categories), computer for school work at home (dummy), and other language than the test language spoken at home (dummy). At the school level, these are school location (three categories), school size, share of fully certified teachers, and shortage of educational material (dummy). At the country level, these are GDP per capita, share of privately managed schools, share of government funding of schools, central exit exams (dummy), and a school-autonomy index. The share of missing values for these variables is generally very low, averaging 5 percent. We impute missing values using the respective country-by-wave mean and include imputation indicators (one dummy per variable that equals one if the respective variable is missing and zero otherwise) in our regression analysis.[14]

The complex structure of country inclusion in our analyses is explained in greater detail in the note to Table 5. In the migrant analysis, countries can be included as residence countries even if there are no GPS measures for them (as long as they participated in PISA and have migrant children from countries of origin that participated in the GPS) and as countries of origin even if there are no PISA measures for them (as long as they participated in GPS and have "sent" students as migrants to PISA-participating countries).

B.2 The Global Preference Survey (GPS)

The Global Preference Survey (GPS) was conducted within the framework of the 2012 wave of the international Gallup World Poll, an annual survey on social and economic topics. Altogether, the GPS uses twelve survey items to measure preferences in the six domains. In an ex-ante validation exercise, students at the University of Bonn took different incentivized decisions in a controlled laboratory setting and answered numerous survey questions for each preference domain (Falk et al. (2016)). The survey items were then selected based on their ability to predict the incentivized choices. For most preference domains, this exercise led to the selection of a combination of one qualitative survey question and one hypothetical choice scenario (see Falk et al. (2018) for details).[15] For each domain, the selected survey items are then combined into a single preference measure using weights from the validation procedure.

[14] In the few cases where a variable is missing for an entire wave in a given country, we impute by averaging over the country's other PISA waves. Dropping these country-by-wave observations as a robustness check does not affect our results (results available upon request).

[15] Exceptions are trust and negative reciprocity, which are measured using one and three qualitative survey questions, respectively. While the qualitative items elicited on Likert scales may be subject to reference bias in the cross-country setting, this is less likely for the choice scenarios.

For patience, the qualitative survey item, elicited on an 11-point Likert scale, is: "*How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?*" The hypothetical choice scenario for patience entails a series of binary decisions between 100 Euro today or a higher amount in the future: "*Suppose you were given the choice between receiving a payment today or a payment in 12 months. We will now present to you five situations. The payment today is the same in each of these situations. The payment in 12 months is different in every situation. For each of these situations we would like to know which one you would choose. Please assume there is no inflation, i.e., future prices are the same as today's prices. Please consider the following: Would you rather receive amount x today or y in 12 months?*"

For risk-taking, the qualitative 11-point-scale question is: "*In general, how willing are you to take risks?*" The quantitative staircase measure is based on the question: "*Please imagine the following situation. You can choose between a sure payment of a particular amount of money, or a draw, where you would have an equal chance of getting amount x or getting nothing. We will present to you five different situations. What would you prefer: a draw with a 50% chance of receiving amount x, and the same 50% chance of receiving nothing, or the amount of y as a sure payment?*"

The GPS dataset does not provide responses to individual items, so we use the available combined preference measures in our analysis. The GPS dataset contains one *z*-standardized variable for each preference domain. Standardization is conducted at the individual level so that each preference has mean zero and standard deviation one in the individual-level world sample. For the purpose of our analysis, we *z*-standardize each individual preference measure in our respective analytical sample and collapse standardized preference measures at the country level. Table 6 presents descriptive statistics of the resulting data.

In the 49-country sample of our baseline analysis, there is a significant cross-country correlation between patience and risk-taking of 0.358 (depicted in Figure 2). Table 7 shows country-level correlations for all preference measures. While patience is not significantly correlated with the other four GPS preference domains (although there are marginal correlations with negative reciprocity and trust), there is a significant correlation of risk-taking with negative reciprocity.

The GPS has several important advantages over alternative international datasets with proxies for national preferences, because it provides scientifically validated measures of the two preference components underlying intertemporal decision-making from representative samples for a large set of countries. The closest alternatives are the World Values Survey (WVS) and the Hofstede (1991) data, both of which provide survey data on attitudes, beliefs, and personality traits across countries. While the WVS is based on representative samples, the Hofstede data are mainly based on IBM employees and are not representative. Importantly, in contrast to the GPS, the validity of the proxies for patience, risk-taking, and other preference domains from these surveys is unknown.[16] Reinforcing the high quality of the GPS data, Falk et al. (2018) show that the GPS patience measure is more predictive of comparative economic development than the measures of long-term orientation from the other two datasets. Interestingly, the correlations of the GPS measures of patience and risk-taking with their respective proxies in the WVS and Hofstede datasets are limited: The correlations of GPS patience with the WVS and Hofstede long-term orientation measures are -0.060 and 0.247, respectively, and statistically insignificant (Table 7). The correlations of GPS risk-taking with WVS risk-taking and Hofstede uncertainty avoidance are only slightly stronger at 0.239 and -0.302, respectively. For our robustness analyses, however, we investigate the WVS and Hofstede data as alternative measures for patience and risk-taking (see Appendix C and Appendix D).

[16] For instance, the best proxy for patience in the WVS is an item on "long-term orientation" that asks "*Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important?*" and is coded 1 if the respondent selects the item "*thrift, saving money and things*", and 0 otherwise. The Hofstede dataset contains proxies for long-term orientation and uncertainty avoidance that are composed of a collection of four qualitative survey items each, several of which appear somewhat unrelated to the concepts that they mean to measure. For example, long-term orientation includes an item on "*How proud are you to be a citizen of your country?*" and uncertainty avoidance includes an item on "*All in all, how would you describe your health these days?*" (see footnote 7 in Falk et al. (2018) for details).

Appendix C: Additional Robustness Analyses for the Baseline Analysis

This appendix provides additional robustness analyses for the baseline analysis presented in section 3 of the main text.

Restriction to wave 2015 of PISA. Our main analysis pools the achievement data of all seven PISA waves (2000-2018), which is justified because cultural aspects by definition are focused on traits that are fairly unchanged over the long run. Moreover, the vast majority of country variation in PISA scores is between countries rather than over time. Pooling extends the country sample and provides more precise measures of long-run educational achievement. Table 10 shows that results are qualitatively the same when restricting the analysis to the 2015 PISA wave (the first PISA wave after the elicitation of the GPS data in 2012), indicating that the pooled analysis is not affected by the relative timing of the observation of preference and achievement data.

Country subsamples. To test whether our main results differ by level of development, columns 1 and 2 of Table 11 present separate regressions for OECD countries and non-OECD countries, respectively (measured as ever belonged to OECD). The qualitative pattern of our findings is very similar and does not differ significantly between the two subsamples.

In additional subsample analyses, we re-estimated the models in Table 1 (columns 3 and 4) excluding one wave or one country at a time. Qualitative results are insensitive to this alteration. The coefficients on patience and risk-taking remain significant in all these regressions (not shown).

Additional subjects. Our main analysis focuses on math achievement, which is generally conceived to be most readily comparable across countries compared to other subjects such as reading (which by construction is to some extent language specific). Reassuringly, results are very similar for achievement in science and reading (columns 1 and 2 of Table 12). In science, a one s.d. increase in patience (risk-taking) is associated with a test-score increase (decrease) by 1.12 s.d. (1.17 s.d.). In reading, the corresponding coefficient is 1.11 s.d. (1.13 s.d.). Thus, the reported associations are universal in the sense that they do not depend on a particular subject.

Alternative preference measures. Given the rather vague measurement of the underlying intertemporal concepts in the WVS and Hofstede datasets (see Appendix B.2), we are less confident about the validity of these alternative measures. Still, Table 13 shows that the WVS cultural measures yield a similar pattern of a positive association of student achievement with long-term orientation and a negative association with risk-taking (column 1). Using the Hofstede measures, long-term orientation is significantly positively associated with student achievement, whereas uncertainty avoidance is insignificant (column 2).[17]

Oster (2019) analysis. We also perform an analysis of unobservable selection and coefficient stability following Oster (2019). We compare our baseline model (column 3 of Table 1) to a restricted model without the control variables and follow Oster (2019) in setting $R_{max} = 1.3\tilde{R}$. Results in column 1 of Table 14 indicate that, assuming $\delta = 1$, the estimated bias-adjusted treatment effect β^* for patience is 1.358, which is even larger than our baseline estimate because adding the controls to the restricted model increases the coefficient estimate. For risk-taking, β^* is -1.181, only slightly below our baseline estimate. Thus, both estimates remain substantial in the bounding analysis that assumes that selection on unobservables is as strong as selection on observables. In both cases, the value of δ for which β = 0 far exceeds the suggested cutoff of $\delta = 1$: for patience, δ = -18.093, and for risk-taking, δ = 8.224. That is, the degree of selection on unobservables would have to be several times as large as selection on observables to eliminate the main result.

Accounting for uncertainty in GPS estimates. The GPS measures of patience and risk-taking are effectively generated estimators, reflecting estimated sample means based on random samples of around 1,000 respondents in each country (see section 2.2 in the main text). This uncertainty in the generated regressors should be accounted for in the regressions (e.g., Pagan (1984); Murphy and Topel (1985)). We use a two-step bootstrap procedure where the first step draws a random sample of 992 observations with replacement in each country (the smallest number of observations within a country in the GPS data) and computes the sample means of patience and risk-taking for each country. The second step uses these preference values to run our main regression (column 3 of Table 1). Repeating this procedure 1,000 times, we calculate our coefficient estimates as the mean of these repeated estimations and the bootstrapped standard errors of our coefficient estimates as the standard deviation in the sample of 1,000 estimated coefficients. The bootstrapped coefficients on patience and risk-taking are very similar to the

[17] In a specification that includes all preference measures from the GPS, WVS, and Hofstede together, only the GPS measures of patience and risk-taking remain large and statistically significant, whereas the WVS and Hofstede measures lose their statistical significance (not shown). The GPS results are also robust to controlling for WVS trust, which does not enter significantly (not shown).

ones reported in the paper and remain statistically highly significant (results available upon request).

Uncertainty in the earnings returns to education. To scrutinize the relative importance of different potential channels of the relationship between risk-taking and human capital investment discussed in Appendix A.3, we conducted additional regression analyses in which we account for the uncertainty in earnings returns to education (proxied by the R^2 of Mincer-type regressions; see Hanushek et al. (2015, 2017)). The results show that the coefficient on risk-taking becomes less negative as earnings uncertainty increases. At face value, this finding suggests that both (i) a negative effect of risk-taking suggested by the crime literature and – to a lesser extent – (ii) a positive effect of risk-taking suggested by studies highlighting the earnings variance in higher-educated occupations may act simultaneously on student achievement (not shown).

Appendix D: Additional Robustness Analyses for the Migrant Analysis

This appendix provides additional robustness analyses for the migrant analysis presented in section 4 of the main text.

The regressions of the migrant analysis include 180 fixed effects for each residence-country by wave cell. Table 15 shows that results are very similar in specifications that include 48 fixed effects for the respective residence countries and six fixed effects for waves, but not their interactions.

Country subsamples. Results of the migrant analysis also do not differ significantly by the level of development of migrants' countries of origin. Patience enters significantly positively, and risk-taking significantly negatively, in the subsamples of migrant students from both OECD and non-OECD countries of origin (columns 3 and 4 of Table 11). The positive point estimate of patience is somewhat larger in OECD countries, whereas the negative point estimate of risk-taking is somewhat larger (in absolute terms) in non-OECD countries. However, neither difference is statistically significant.

Additional subjects. Columns 3 and 4 of Table 12 present results for student achievement in science and reading, respectively. Results are again very similar to math, although the negative coefficient on risk-taking is not statistically significant in the other two subjects.

Alternative preference measures. The qualitative pattern of results on patience is also confirmed with the alternative measures in the WVS and Hofstede datasets (columns 3 and 4 of Table 13). In both cases, there is a significant positive effect of long-term orientation on student achievement, in line with the results in Figlio et al. (2019). The WVS data also confirm a significant negative effect of risk-taking. By contrast, the Hofstede risk measure points in the opposite direction – a negative effect of uncertainty *avoidance* – which presumably reflects the poor measurement of the underlying concept by the items contained in this proxy (see Appendix B.2).

Different migrant subgroups. In addition to distinguishing by the language spoken at home (see section 4.3 in the main text), another migrant subgroup analysis distinguishes migrants by the time at which the students themselves migrated to the country of residence. Results in columns 1 and 2 of Table 16 show that the effect of patience does not differ significantly between second-generation migrants (born in the residence country after their parents had migrated) and first-generation migrants (born in the country of origin), and the negative effect of risk-taking is actually larger for second-generation migrants.

We can exploit information on the age of migration in our dataset to subdivide the first-generation migrants further by whether they arrived in the residence country before or after age 6, when they would usually start school. Within the first-generation migrants, the effects of the two preference components do not differ significantly by whether students had migrated earlier or later (columns 3 and 4). While these patterns show the robustness of our main findings, they do not support the notion that later migrants hold onto more of their country-of-origin culture.

Alternative migrant definitions. Our main specification uses the country of origin of the students' fathers for reference for second-generation migrant students. Results in the first column of Table 17 show that estimates are virtually identical when the country of origin of the mother is used instead. Column 2 uses the average value of the national preferences of the country of origin of both parents when both are available, and the measure of the respective country of origin of the father or mother if the information is available only for one of them. Again, results hardly change in the slightly larger sample of 83,798 students.

As the PISA data allow us to observe both parents' country of origin, we can also enter the national preferences of fathers' and mothers' country of origin simultaneously (column 3). While this horse-race specification is identified only from children whose parents come from different countries, results still provide a relatively clear pattern. The effect of patience is significantly positive for both parents, although it is twice as large for fathers' compared to mothers' patience. By contrast, the effect of risk-taking is fully driven by fathers, with risk-taking of the mothers' country of origin not entering migrant students' achievement.

In our main specification, we adopt a rather narrow definition of migrants that includes only students whose parents are both born in a different country than the testing country. Alternatively, we can use a wider definition that includes all students with at least one parent born abroad – defining as natives only those with both parents born in the testing country. This wider definition increases the number of observations to over 140,000 migrant students. While, expectedly, point estimates are slightly smaller with this broader measurement, results are in fact very similar to those in the smaller sample with the narrower definition, independent of whether the country of origin is defined based on the mother, the father, or the average (columns 4-6).

Including the national preferences of both parents' countries of origin simultaneously also again yields very similar results (column 7).

In a few cases, the effect of a specific country of origin is identified from only a limited number of student observations, potentially introducing substantial measurement error for these countries of origin. However, if we restrict the analysis to cases where at least 50 students are observed from each country of origin – which reduces the number of countries of origin from 58 to 46 – results remain virtually unaffected (column 8).[18]

Selective migration. In addition to the detailed analysis of selective migration presented in the main text, we also estimated a specification that interacts the preference variables with the raw difference in the preference variables between country of origin and residence country (rather than the absolute difference as in column 5 of Table 3). None of the interactions is significant, indicating that effects do not differ by whether migrants go to countries with higher versus lower preference values (not shown). The results remain unchanged when the controls for geographical and cultural distance are entered together (not shown).

Oster (2019) analysis. An analysis of unobservable selection and coefficient stability following Oster (2019) again suggests that results remain stable. Comparing our baseline model (column 3 of Table 2) to a restricted model without the control variables and setting $R_{max} = 1.3\tilde{R}$, results in column 2 of Table 14 indicate that the estimated bias-adjusted treatment effect β^* (assuming $\delta = 1$) is 0.875 for patience and -0.264 for risk-taking, very close to our baseline estimates. Accordingly, selection on unobservables would have to be 13.6 the selection on observables for patience and 9.8 for risk-taking in order to eliminate the main result.

Accounting for uncertainty in GPS estimates. To account for uncertainty in the generated regressors, we again implement the two-step bootstrap procedure described in Appendix C. Again, point estimates are very similar to our baseline model and estimates remain statistically highly significant (results available upon request).

[18] This robustness analysis drops the following countries of origin: Bangladesh (17 observations), Canada (1), Chile (47), Finland (2), Georgia (3), Indonesia (27), Kazakhstan (34), Lithuania (3), Moldova (11), Nigeria (4), Saudi Arabia (8), and Thailand (20).

Appendix E: Models with Extended Controls

Estimates of education production functions usually contain measures for proximate inputs – family inputs, school resources, and institutional features. To the extent that these proximate inputs are themselves the outcomes of intertemporal choice decisions, they would be bad controls in a model depicting the overall effect of national preferences on student achievement. Including proximate input factors in our model, however, provides a descriptive evaluation of the importance of these input channels and shows the robustness of the preference-achievement association to consideration of variation in input factors that stem from other sources. Therefore, in this appendix we report specifications that include a rich set of proximate input factors as control variables that would generally be included in education production functions:

$$T_{ict} = \beta_1 Patience_c + \beta_2 Risk_c + \alpha_1 B_{ict} + \alpha_2 F_{ict} + \alpha_3 S_{ict} + \alpha_4 I_{ct} + \mu_t + \varepsilon_{ict} \qquad (3)$$

which, in addition to our baseline model (1), includes measures of the inputs from student's families F, schools S, and institutional structures of school systems I.

When adding the proximate inputs as control variables to the model of the baseline analysis, the coefficient estimates on the two preference components remain large and statistically highly significant, but are reduced in size (column 1 of Table 9). The extended set of controls for family, school, and institutional inputs (described in the table notes) are likely bad controls because they too are outcomes of the deeper cultural traits. The reduction of the coefficients on patience by 39 percent and on risk-taking by 33 percent (when comparing column 3 of Table 1 to column 1 of Table 9) in this descriptive analysis indicates that a substantial part of the overall effects of the two preference components may work through the channels of these proximate inputs.

Column 2 of Table 9 provides equivalent results for the migrant analysis. The specification adds the set of extended controls on family and school inputs in the residence country, as well as the country of origin's GDP per capita. This latter control addresses the concern that, for instance, better performance of migrants from high-patience countries merely reflects differences in income (as opposed to genuine effects of cultural traits). As expected, the coefficient on patience is reduced in this specification (because the family and GDP controls may take out some of the total effect of cultural traits), but it remains large and significant (as does the coefficient on risk-taking).[19]

Section 5 in the main text provides a closer analysis of the association of the different input factors with the two preference components.

[19] When added to the model, a negative coefficient on an interaction between patience and risk-taking is significant in the baseline model but loses significance when extended controls are included (not shown).

Appendix F: Details of the Channel Analysis

In order to shed light on potential mechanisms, section 5 in the main text considers how patience and risk-taking relate to the major categories of proximate inputs – family inputs, school inputs, and institutional inputs. The starting point of the channel analysis is developing composite measures of the three categories of proximate inputs. We map the input variables (see Appendix B.1) into the three input vectors as follows: family inputs: gender, age, migration status, parental education, parental occupation, books at home, computer at home, language spoken at home, and GDP per capita (capturing overall economic wellbeing in the country); school inputs: school location, school size, share of fully certified teachers, and shortage of educational material; institutional inputs: share of privately managed schools, share of government funding at school, central exit exams, and school autonomy.

Building on the typical analysis of international educational production functions found in Woessmann (2016b), we run a pooled cross-country regression of PISA math scores on our full set of input variables. We then use the coefficient estimates on the individual variables in the model to aggregate them into family, school, and institutional factors. That is, for each input category, we calculate a linear combination as the sum of the products of the individual variables times their respective coefficients. We finally collapse the three combined input factors, as well as the residual of the achievement regression, to the country level.

This aggregation of the individual proximate input variables uses coefficient estimates from an education production function that may be biased by the omission of the deeper preference variables. Because individual coefficient estimates will be more biased for variables that are more strongly correlated with the preference measures, the estimates based on this aggregation serve as an upper bound for the preference relationships.

Thus, estimating the first step of the aggregation analysis including controls for the two national preferences can serve as a lower bound, as the preference measures take out important parts of the variation in the proximate inputs. By construction, the input coefficients estimated in the lower-bound analysis are unaffected by the cultural factors, and the coefficients and the R^2 for the residual category are zero.

Put together, the results of the upper-bound and lower-bound procedures presented in Table 4 are consistent with different input components of the education production function playing a role as channels through which the two intertemporal preferences affect student achievement.

An interesting aspect is that national preferences have limited association with institutional factors. Prior analyses have highlighted the importance of institutional factors in explaining cross-country achievement differences (Hanushek and Woessmann (2011); Woessmann (2016b)), implying that changing institutions may be a way for nations wishing to improve their schools to break out of cultural constraints.

References

Akyol, Ş. Pelin, Kala Krishna, Jinwen Wang (2018). Taking PISA Seriously: How Accurate Are Low Stakes Exams? NBER Working Paper 24930. Cambridge, MA: National Bureau of Economic Research.

Alan, Sule, Nazli Baydar, Teodora Boneva, Thomas F. Crossley, Seda Ertac (2017). Transmission of Risk Preferences from Mothers to Daughters. *Journal of Economic Behavior & Organization* 134: 60-77.

Alan, Sule, Teodora Boneva, Seda Ertac (2019). Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit. *Quarterly Journal of Economics* 134 (3): 1121-1162.

Alan, Sule, Seda Ertac (2018). Fostering Patience in the Classroom: Results from Randomized Educational Intervention. *Journal of Political Economy* 126 (5): 1865-1911.

Alesina, Alberto, Paola Giuliano (2014). Family Ties. In *Handbook of Economic Growth, Vol. 2*, edited by Philippe Aghion, Steven N. Durlauf. Amsterdam: North Holland: 177-215.

Alesina, Alberto, Paola Giuliano (2015). Culture and Institutions. *Journal of Economic Literature* 53 (4): 898-944.

Almlund, Mathilde, Angela L. Duckworth, James Heckman, Tim Kautz (2011). Personality Psychology and Economics. In *Handbook of the Economics of Education, Vol. 4*, edited by Eric A. Hanushek, Stephen Machin, Ludger Woessmann. Amsterdam: North Holland: 1-181.

Andreoni, James, Charles Sprenger (2012). Risk Preferences Are Not Time Preferences. *American Economic Review* 102 (7): 3357-3376.

Andreoni, James, Charles Sprenger (2015). Risk Preferences Are Not Time Preferences: Reply. *American Economic Review* 105 (7): 2287-2293.

Balart, Pau, Matthijs Oosterveen, Dinand Webbink (2018). Test Scores, Noncognitive Skills and Economic Growth. *Economics of Education Review* 63: 134-153.

Becker, Gary S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York, NY: National Bureau of Economic Research.

Ben-Porath, Yoram (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy* 75 (4): 352-365.

Ben-Porath, Yoram (1970). The Production of Human Capital over Time. In *Education, Income and Human Capital*, edited by W. Lee Hansen. New York: National Bureau of Economic Research: 129-147.

Benzoni, Luca, Olena Chyruk (2015). The Value and Risk of Human Capital. *Annual Review of Financial Economics* 7 (1): 179-200.

Bisin, Alberto, Thierry Verdier (2011). The Economics of Cultural Transmission and Socialization. In *Handbook of Social Economics*, edited by Jess Benhabib, Alberto Bisin, Matthew O. Jackson. Amsterdam: North-Holland: 339-416.

Borghans, Lex, Trudie Schils (2012). The Leaning Tower of Pisa: Decomposing Achievement Test Scores into Cognitive and Noncognitive Components. Mimeo.

Brown, Jeffrey, Chichun Fang, Francisco Gomes (2012). Risk and Returns to Education. NBER Working Paper No. 18300. Cambridge, MA: National Bureau of Economic Research.

Carroll, Christopher D., Byung-Kun Rhee, Changyong Rhee (1994). Are There Cultural Effects on Saving? Some Cross-Sectional Evidence. *Quarterly Journal of Economics* 109 (3): 685-699.

Castillo, Marco, Jeffrey L. Jordan, Ragan Petrie (2018). Children's Rationality, Risk Attitudes and Field Behavior. *European Economic Review* 102: 62-81.

Castillo, Marco, Jeffrey L. Jordan, Ragan Petrie (2019). Discount Rates of Children and High School Graduation. *Economic Journal* 129 (619): 1153-1181.

Castillo, Marco, John A. List, Ragan Petrie, Anya Samek (2020). Detecting Drivers of Behavior at an Early Age: Evidence from a Longitudinal Field Experiment. NBER Working Paper No. 28288. Cambridge, MA: National Bureau of Economic Research.

Cheung, Stephen L. (2015). Comment on "Risk Preferences Are Not Time Preferences": On the Elicitation of Time Preference under Conditions of Risk. *American Economic Review* 105 (7): 2242-2260.

Cordero, Jose M., Cristina Polo, Daniel Santín, Rosa Simancas (2018). Efficiency Measurement and Cross-Country Differences among Schools: A Robust Conditional Nonparametric Analysis. *Economic Modelling* 74: 45-60.

Cunha, Flavio, James J. Heckman, Susanne M. Schennach (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica* 78 (3): 883-931.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review* 100 (3): 1238-1260.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde (2018). On the Relationship between Cognitive Ability and Risk Preference. *Journal of Economic Perspectives* 32 (2): 115-134.

Duckworth, Angela L., Christopher Peterson, Michael D. Matthews, Dennis R. Kelly (2007). Grit: Perseverance and Passion for Long-Term Goals. *Journal of Personality and Social Psychology* 92 (6): 1087-1101.

Epper, Thomas, Helga Fehr-Duda (2015). Comment on "Risk Preferences Are Not Time Preferences": Balancing on a Budget Line. *American Economic Review* 105 (7): 2261-2271.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, Uwe Sunde (2018). Global Evidence on Economic Preferences. *Quarterly Journal of Economics* 133 (4): 1645-1692.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, Uwe Sunde (2016). The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences. IZA Discussion Paper 9674. Bonn: Institute for the Study of Labor.

Fernández, Raquel, Alessandra Fogli (2009). Culture: An Empirical Investigation of Beliefs, Work, and Fertility. *American Economic Journal: Macroeconomics* 1 (1): 146-77.

Figlio, David, Paola Giuliano, Umut Özek, Paola Sapienza (2019). Long-Term Orientation and Educational Performance. *American Economic Journal: Economic Policy* 11 (4): 272-309.

Galor, Oded, Ömer Özak (2016). The Agricultural Origins of Time Preference. *American Economic Review* 106 (10): 3064-3103.

Giuliano, Paola (2007). Living Arrangements in Western Europe: Does Cultural Origin Matter? *Journal of the European Economic Association* 5 (5): 927-952.

Gneezy, Uri, John A. List, Jeffrey A. Livingston, Xiangdong Qin, Sally Sadoff, Yang Xu (2019). Measuring Success in Education: The Role of Effort on the Test Itself. *American Economic Review: Insights* 1 (3): 291-308.

Golsteyn, Bart H.H., Hans Grönqvist, Lena Lindahl (2014). Adolescent Time Preferences Predict Lifetime Outcomes. *Economic Journal* 124 (580): F739-F761.

Groot, Wim, Hessel Oosterbeek (1992). Optimal Investment in Human Capital under Uncertainty *Economics of Education Review* 11 (1): 41-49.

Guiso, Luigi, Paola Sapienza, Luigi Zingales (2006). Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives* 20 (2): 23-48.

Halevy, Yoram (2008). Strotz Meets Allais: Diminishing Impatience and the Certainty Effect. *American Economic Review* 98 (3): 1145-1162.

Hanushek, Eric A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature* 24 (3): 1141-1177.

Hanushek, Eric A., Guido Schwerdt, Simon Wiederhold, Ludger Woessmann (2015). Returns to Skills around the World: Evidence from PIAAC. *European Economic Review* 73: 103-130.

Hanushek, Eric A., Guido Schwerdt, Simon Wiederhold, Ludger Woessmann (2017). Coping with Change: International Differences in the Returns to Skills. *Economics Letters* 153: 15-19.

Hanushek, Eric A., Ludger Woessmann (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature* 46 (3): 607-668.

Hanushek, Eric A., Ludger Woessmann (2011). The Economics of International Differences in Educational Achievement. In *Handbook of the Economics of Education, Vol. 3,* edited by Eric A. Hanushek, Stephen Machin, Ludger Woessmann. Amsterdam: North Holland: 89-200.

Hartog, Joop, Luis Diaz-Serrano (2007). Earnings Risk and Demand for Higher Education: A Cross-Section Test for Spain. *Journal of Applied Economics* 10 (1): 1-28.

Hartog, Joop, Luis Diaz-Serrano (2014). Schooling as a Risky Investment: A Survey of Theory and Evidence. *Foundations and Trends in Microeconomics* 9 (3-4): 159-331.

Heckman, James J. (1976). A Life-Cycle Model of Earnings, Learning, and Consumption. *Journal of Political Economy* 84 (4): S11-S44.

Hofstede, Geert H. (1991). *Cultures and Organizations: Software of the Mind*. London: McGraw-Hill.

Hofstede, Geert H., Gert Jan Hofstede, Michael Minkov (2010). *Cultures and Organizations: Software of the Mind: Intercultural Cooperation and Its Importance for Survival. 3rd Ed*. New York: McGraw-Hill.

Kim, Hong-Kyun (2001). Is There a Crowding-out Effect between School Expenditure and Mother's Child Care Time? *Economics of Education Review* 20 (1): 71-80.

Koerselman, Kristian, Roope Uusitalo (2014). The Risk and Return of Human Capital Investments. *Labour Economics* 30: 154-163.

Kosse, Fabian, Friedhelm Pfeiffer (2012). Impatience among Preschool Children and Their Mothers. *Economics Letters* 115 (3): 493-495.

Lavecchia, Adam M., Heidi Liu, Philip Oreopoulos (2016). Behavioral Economics of Education: Progress and Possibilities. In *Handbook of the Economics of Education, Vol. 5*, edited by Eric A. Hanushek, Stephen Machin, Ludger Woessmann. Amsterdam: North Holland: 1-74.

Levhari, David, Yoram Weiss (1974). The Effect of Risk on the Investment in Human Capital. *American Economic Review* 64 (6): 950-963.

Mendez, Ildefonso (2015). The Effect of the Intergenerational Transmission of Noncognitive Skills on Student Performance. *Economics of Education Review* 46: 78-97.

Miao, Bin, Songfa Zhong (2015). Comment on "Risk Preferences Are Not Time Preferences": Separating Risk and Time Preference. *American Economic Review* 105 (7): 2272-2286.

Mincer, Jacob (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy* 66 (4): 281-302.

Mincer, Jacob (1974). *Schooling, Experience, and Earnings*. New York: NBER.

Mischel, Walter, Yuichi Shoda, Monica L. Rodriguez (1989). Delay of Gratification in Children. *Science* 244 (4907): 933-938.

Moffitt, Terrie E., Louise Arseneault, Daniel Belsky, Nigel Dickson, Robert J. Hancox, HonaLee Harrington, Renate Houts, Richie Poulton, Brent W. Roberts, Stephen Ross, Malcolm R. Sears, W. Murray Thomson, Avshalom Caspi (2011). A Gradient of Childhood Self-Control Predicts Health, Wealth, and Public Safety. *Proceedings of the National Academy of Sciences* 108 (7): 2693-2698.

Murphy, Kevin M., Robert H. Topel (1985). Estimation and Inference in Two-Step Econometric Models. *Journal of Business & Economic Statistics* 3 (4): 370-379.

OECD (2019). *PISA 2018 Results (Volume I)*. Paris: OECD Publishing.

Oster, Emily (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics* 37 (2): 187-204.

Pagan, Adrian (1984). Econometric Issues in the Analysis of Regressions with Generated Regressors. *International Economic Review* 25 (1): 221-247.

Palacios-Huerta, Ignacio (2003). An Empirical Analysis of the Risk Properties of Human Capital Returns. *American Economic Review* 93 (3): 948-964.

Potrafke, Niklas (2019). Risk Aversion, Patience and Intelligence: Evidence Based on Macro Data. *Economics Letters* 178: 116-120.

Rosen, Sherwin (1976). A Theory of Life Earnings. *Journal of Political Economy* 84 (4): S45-S67.

Sunde, Uwe, Thomas Dohmen, Benjamin Enke, Armin Falk, David Huffman, Gerrit Meyerheim (2021). Patience and Comparative Development. *Review of Economic Studies*: forthcoming.

Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rützler, Stefan T. Trautmann (2013). Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior. *American Economic Review* 103 (1): 510-531.

Todd, Petra E., Kenneth I. Wolpin (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *Economic Journal* 113 (485): F3-33.

Weiss, Yoram (1972). The Risk Element in Occupational and Educational Choices. *Journal of Political Economy* 80 (6): 1203-1213.

Woessmann, Ludger (2016a). The Economic Case for Education. *Education Economics* 24 (1): 3-32.

Woessmann, Ludger (2016b). The Importance of School Systems: Evidence from International Differences in Student Achievement. *Journal of Economic Perspectives* 30 (3): 3-32.

Zamarro, Gema, Collin Hitt, Ildefonso Mendez (2019). When Students Don’t Care: Reexamining International Differences in Achievement and Student Effort. *Journal of Human Capital* 13 (4): 519-552.

Appendix Tables

Table 5: Countries in the different analyses

	PISA	GPS	Cross-country analysis	Migrant analysis	
				Residence country	Country of origin
	(1)	(2)	(3)	(4)	(5)
Afghanistan		x			x
Algeria	x	x	x		
Argentina	x	x	x	x	x
Australia	x	x	x	x	x
Austria	x	x	x	x	x
Bangladesh		x			x
Belarus	x			x	
Belgium	x			x	
Bolivia		x			x
Bosnia Herzegovina	x	x	x	x	x
Brazil	x	x	x		x
Brunei Darussalam	x			x	
Canada	x	x	x	x	x
Chile	x	x	x		x
China		x			x
Colombia	x	x	x		x
Costa Rica	x	x	x	x	
Croatia	x	x	x	x	x
Czech Republic	x	x	x	x	x
Denmark	x			x	
Dominican Republic	x			x	
Egypt		x			x
Estonia	x	x	x		x
Finland	x	x	x	x	x
France	x	x	x		x
Georgia	x	x	x	x	x
Germany	x	x	x	x	x
Greece	x	x	x		x
Haiti		x			x
Hong Kong	x			x	
Hungary	x	x	x		x
India		x			x
Indonesia	x	x	x	x	x
Iran		x			x
Iraq		x			x
Ireland	x			x	
Israel	x	x	x	x	
Italy	x	x	x		x
Japan	x	x	x		
Jordan	x	x	x	x	x
Kazakhstan	x	x	x		x
Kyrgyzstan	x			x	
Latvia	x			x	
Liechtenstein	x			x	
Lithuania	x	x	x		x
Luxembourg	x			x	
Macao	x			x	
Mauritius	x			x	
Mexico	x	x	x	x	
Moldova	x	x	x	x	x
Montenegro	x			x	

(continued on next page)

Table 5 (continued)

	PISA	GPS	Cross-country analysis	Migrant analysis	
				Residence country	Country of origin
	(1)	(2)	(3)	(4)	(5)
Morocco	x	x	x	x	x
Netherlands	x	x	x	x	x
New Zealand	x			x	
Nicaragua		x			x
Nigeria		x			x
North Macedonia	x			x	
Norway	x			x	
Pakistan		x			x
Panama	x			x	
Peru	x	x	x		
Philippines	x	x	x	x	x
Poland	x	x	x		x
Portugal	x	x	x	x	x
Qatar	x			x	
Romania	x	x	x		x
Russia	x	x	x		x
Saudi Arabia	x	x	x	x	x
Serbia	x	x	x		x
Slovakia	x			x	
Slovenia	x			x	
South Africa		x			x
South Korea	x	x	x	x	x
Spain	x	x	x		x
Suriname		x			x
Sweden	x	x	x		x
Switzerland	x	x	x	x	x
Thailand	x	x	x		x
Turkey	x	x	x	x	x
Ukraine	x	x	x	x	x
United Arab Emirates	x	x	x		x
United Kingdom	x	x	x	x	x
United States	x	x	x		x
Uruguay	x			x	
Venezuela		x			x
Vietnam	x	x	x		x
Total: 86 countries	71	64	49	48	58

Notes: The structure of country inclusion in the different parts of our analysis is complex. Three countries are included only in the baseline analysis because they participated in PISA (and GPS) but do not have migrant students with country-of-origin information (for which there is GPS data) and no student from these countries is observed as a migrant student in another PISA country. Another three countries are included in the baseline analysis and (only) as residence countries in the migrant analysis because they participated in PISA (and GPS) and have migrant students from countries of origin with GPS data, but no student from these countries is observed as a migrant student in another PISA country. 23 countries are included in the baseline analysis and both as residence countries and as countries of origin in the migrant analysis. There is also the case of 20 countries that are included in the baseline analysis and (only) as countries of origin in the migrant analysis because they participated in PISA (and GPS) but do not have migrant students with country-of-origin information (for which there is GPS data), and students from these countries are observed as migrant students in other PISA countries. 22 countries are not included in the baseline analysis, but only as residence countries in the migrant analysis because they participated in PISA, but there is no GPS data for them; however, there is GPS data for the country of origin of some of the migrant students tested in these countries. Finally, 15 countries are included only as countries of origin in the migrant analysis; these countries did not participate in PISA themselves and therefore cannot be included in the baseline analysis or as residence countries in the migrant analysis, but there is GPS data for them and students originating from these countries are observed as migrant students in residence countries that did participate in PISA.

Table 6: Descriptive statistics at the country level

	Mean	Std. dev.	Min	Max
	(1)	(2)	(3)	(4)
PISA scores				
Math	4.520	0.560	3.524	5.410
Science	4.597	0.531	3.579	5.415
Reading	4.535	0.521	3.395	5.345
Preferences				
Patience	-0.003	0.384	-0.555	0.946
Risk-taking	0.027	0.241	-0.746	0.789
Positive reciprocity	-0.016	0.315	-1.094	0.558
Negative reciprocity	0.025	0.308	-0.510	0.716
Altruism	-0.022	0.346	-0.923	0.679
Trust	-0.016	0.249	-0.575	0.507

Notes: PISA scores: country means, pooled across all PISA waves 2000-2018, weighted by sampling probabilities. Preferences: country means of GPS preference data. Sample: 263 country-by-wave observations (reflecting 49 countries) contained in our baseline analysis of Table 1. Data sources: PISA international student achievement test, 2000-2018; Falk et al. (2018).

Table 7: Country-level correlation of different preference components

	Patience	Risk-taking	Positive reciprocity	Negative reciprocity	Altruism	Trust	WVS long-term orientation	WVS risk-taking	Hofstede long-term orientation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Risk-taking	0.358								
	(0.011)								
Positive reciprocity	-0.154	-0.148							
	(0.291)	*(0.310)*							
Negative reciprocity	0.236	0.334	-0.277						
	(0.103)	*(0.019)*	*(0.054)*						
Altruism	-0.051	0.110	0.699	-0.200					
	(0.728)	*(0.451)*	*(0.000)*	*(0.168)*					
Trust	0.197	0.162	0.259	-0.025	0.207				
	(0.176)	*(0.265)*	*(0.072)*	*(0.864)*	*(0.153)*				
WVS long-term orientation	-0.060	-0.334	-0.195	0.057	-0.163	-0.104			
	(0.700)	*(0.027)*	*(0.204)*	*(0.715)*	*(0.290)*	*(0.500)*			
WVS risk-taking	-0.260	0.239	0.117	0.138	0.269	0.313	-0.079		
	(0.125)	*(0.160)*	*(0.498)*	*(0.423)*	*(0.112)*	*(0.063)*	*(0.646)*		
Hofstede long-term orientation	0.247	-0.219	-0.326	0.321	-0.256	-0.246	0.609	-0.310	
	(0.115)	*(0.164)*	*(0.035)*	*(0.038)*	*(0.101)*	*(0.116)*	*(0.000)*	*(0.084)*	
Hofstede uncertainty avoidance	-0.558	-0.302	-0.055	0.123	-0.185	-0.527	0.006	-0.093	0.024
	(0.000)	*(0.046)*	*(0.721)*	*(0.426)*	*(0.228)*	*(0.000)*	*(0.971)*	*(0.611)*	*(0.881)*

Notes: Correlation coefficients; *p*-values in parentheses. Sample: 49 countries contained in our baseline analysis. Number of country observations: 49 among GPS measures, 44 between GPS and Hofstede uncertainty avoidance or WVS long-term orientation, 42 between GPS and Hofstede uncertainty avoidance and among Hofstede measures, 36 between GPS and WVS risk-taking and among WVS measures, and 32 between WVS and Hofstede measures. Data sources: Falk et al. (2018); World Values Survey (WVS); Hofstede, Hofstede, and Minkov (2010).

Table 8: Results by gender

	Baseline analysis		Migrant analysis	
	Girls	Boys	Girls	Boys
	(1)	(2)	(3)	(4)
Patience	1.208***	1.242***	0.953***	0.912***
	(0.129)	*(0.137)*	*(0.109)*	*(0.125)*
Risk-taking	-1.190***	-1.294***	-0.302***	-0.285**
	(0.184)	*(0.187)*	*(0.111)*	*(0.133)*
Residence-country by wave fixed effects	No	No	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Observations	1,005,770	985,412	39,757	40,634
Countries of origin			58	57
Residence countries	49	49	48	48
R^2	0.194	0.201	0.292	0.264

Notes: Dependent variable: PISA math test score, col. 1-2: waves 2000-2018, col. 3-4: waves 2003-2018. Least squares regressions. Col. 1-2: weighted by students' sampling probability. Col. 3-4: sample: students with both parents not born in the country where the student attends school; indicated preference variables refer to country of origin. Control variables: col. 1-2: student gender, age, and migration status; imputation dummies; and wave fixed effects; col. 3-4: student gender, age, dummy for OECD country of origin, imputation dummies. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2000-2018; World Values Survey (WVS); Falk et al. (2018).

Table 9: Model with extended controls

	Baseline analysis	Migrant analysis
	(1)	(2)
Patience	0.748***	0.667***
	(0.192)	*(0.100)*
Risk-taking	-0.835***	-0.352***
	(0.147)	*(0.092)*
Residence-country by wave fixed effects	No	Yes
Baseline control variables	Yes	Yes
Extended control variables	Yes	Yes
Observations	1,992,276	80,398
Countries of origin		58
Residence countries	49	48
R^2	0.368	0.364

Notes: Dependent variable: PISA math test score, col. 1: waves 2000-2018, col. 2: waves 2003-2018. Least squares regressions. Col. 1: weighted by students' sampling probability. Col. 2: sample: students with both parents not born in the country where the student attends school; indicated preference variables refer to country of origin. Baseline control variables: col. 1: student gender, age, and migration status; imputation dummies; and wave fixed effects; col. 2: student gender, age, dummy for OECD country of origin, imputation dummies. Extended control variables, col. 1: baseline controls plus parental education, parental occupation, books at home, computer at home, language spoken at home; school location, school size, share of fully certified teachers at school, shortage of educational material; country's GDP per capita, share of privately managed schools, share of government funding at school, central exit exams, and school autonomy; col. 2: Extended control variables: baseline controls plus parental education, parental occupation, books at home, computer at home, language spoken at home; school location, school size, share of fully certified teachers at school, shortage of educational material; country-of-origin GDP per capita. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2000-2018; Falk et al. (2018).

Table 10: Baseline cross-country analysis restricted to the PISA 2015 wave

	(1)	(2)	(3)	(4)	(5)
Patience	0.794***		1.090***	1.078***	0.763***
	(0.125)		*(0.129)*	*(0.113)*	*(0.175)*
Risk-taking		-0.361	-1.226***	-1.292***	-0.912***
		(0.340)	*(0.220)*	*(0.209)*	*(0.178)*
Positive reciprocity				0.107	
				(0.261)	
Negative reciprocity				0.289*	
				(0.158)	
Altruism				-0.235	
				(0.186)	
Trust				-0.173	
				(0.159)	
Baseline control variables	Yes	Yes	Yes	Yes	Yes
Extended control variables	No	No	No	No	Yes
Observations	319,997	319,997	319,997	319,997	319,997
Countries	41	41	41	41	41
R^2	0.102	0.013	0.157	0.171	0.329

Notes: Dependent variable: PISA math test score, wave 2015. Least squares regression weighted by students' sampling probability. Baseline control variables: student gender, age, and migration status; imputation dummies; and wave fixed effects. Extended control variables: baseline controls plus parental education, parental occupation, books at home, computer at home, language spoken at home; school location, school size, share of fully certified teachers at school, shortage of educational material; country's GDP per capita, share of privately managed schools, share of government funding at school, central exit exams, and school autonomy. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2015; Falk et al. (2018).

Table 11: Results for country subsamples

	Baseline analysis		Migrant analysis	
	OECD	Non-OECD	OECD	Non-OECD
	(1)	(2)	(3)	(4)
Patience	0.963***	1.165**	1.028***	0.812***
	(0.180)	*(0.516)*	*(0.105)*	*(0.185)*
Risk-taking	-0.996***	-1.141***	-0.289**	-0.454**
	(0.271)	*(0.333)*	*(0.132)*	*(0.177)*
Residence-country by wave fixed effects	No	No	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Observations	1,416,506	575,770	28,519	51,879
Countries of origin			24	34
Residence countries	27	22	31	38
R^2	0.112	0.085	0.176	0.309
Difference between subsamples				
Patience	-0.202		0.216	
	(0.540)		*(0.211)*	
Risk-taking	0.144		0.165	
	(0.424)		*(0.219)*	

Notes: Dependent variable: PISA math test score, col. 1-2: waves 2000-2018, col. 3-4: waves 2003-2018. Least squares regressions. Col. 1-2: weighted by students' sampling probability. Col. 3-4: sample: students with both parents not born in the country where the student attends school; indicated preference variables refer to country of origin. Control variables: col. 1-2: student gender, age, and migration status; imputation dummies; and wave fixed effects; col. 3-4: student gender, age, dummy for OECD country of origin, imputation dummies. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2000-2018; Falk et al. (2018).

Table 12: Results in reading and science

	Baseline analysis		Migrant analysis	
	Science	Reading	Science	Reading
	(1)	(2)	(3)	(4)
Patience	1.121***	1.108***	0.995***	0.844***
	(0.121)	*(0.113)*	*(0.143)*	*(0.144)*
Risk-taking	-1.169***	-1.134***	-0.192	-0.106
	(0.180)	*(0.198)*	*(0.124)*	*(0.133)*
Residence-country by wave fixed effects	No	No	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Observations	1,992,276	1,950,722	80,398	80,398
Countries of origin			58	58
Residence countries	49	49	48	48
R^2	0.179	0.189	0.253	0.239

Notes: Dependent variable: PISA test score in science (col. 1 and 3) and reading (col. 2 and 4), respectively. Col. 1-2: waves 2000-2018, col. 3-4: waves 2003-2018. Least squares regressions. Col. 1-2: weighted by students' sampling probability. Col. 3-4: sample: students with both parents not born in the country where the student attends school; indicated preference variables refer to country of origin. Control variables: col. 1-2: student gender, age, and migration status; imputation dummies; and wave fixed effects; col. 3-4: student gender, age, dummy for OECD country of origin, imputation dummies. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2000-2018; Falk et al. (2018).

Table 13: Alternative WVS and Hofstede measures of national preferences

	Baseline analysis		Migrant analysis	
	WVS	Hofstede	WVS	Hofstede
	(1)	(2)	(3)	(4)
WVS long-term orientation	0.171*		0.176***	
	(0.091)		*(0.030)*	
WVS risk-taking	-0.245***		-0.120***	
	(0.075)		*(0.029)*	
Hofstede long-term orientation		0.339***		0.206***
		(0.054)		*(0.029)*
Hofstede uncertainty avoidance		-0.101		-0.092***
		(0.068)		*(0.031)*
Residence-country by wave fixed effects	No	No	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Observations	1,531,302	1,839,052	62,834	74,892
Countries of origin			40	48
Residence countries	36	42	44	48
R^2	0.109	0.134	0.246	0.250

Notes: Dependent variable: PISA math test score, col. 1-2: waves 2000-2018, col. 3-4: waves 2003-2018. Least squares regressions. Col. 1-2: weighted by students' sampling probability. Col. 3-4: sample: students with both parents not born in the country where the student attends school; indicated preference variables refer to country of origin. WVS and Hofstede measures *z*-standardized at the country level. Control variables: col. 1-2: student gender, age, and migration status; imputation dummies; and wave fixed effects; col. 3-4: student gender, age, dummy for OECD country of origin, imputation dummies. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2000-2018; World Values Survey (WVS); Hofstede, Hofstede, and Minkov (2010).

Table 14: Analysis of unobservable selection and coefficient stability following Oster (2019)

	Baseline analysis		Migrant analysis	
	(1)		(2)	
Restricted model				
Patience	1.209***		0.933***	
	(0.131)		*(0.117)*	
Risk-taking	-1.249***		-0.295**	
	(0.184)		*(0.122)*	
Observations	1,992,276		80,398	
Countries of origin			58	
Residence countries	49		48	
R^2	0.189		0.272	
Baseline model				
Patience	1.226***		0.931***	
	(0.132)		*(0.116)*	
Risk-taking	-1.241***		-0.294**	
	(0.184)		*(0.122)*	
Observations	1,992,276		80,398	
Countries of origin			58	
Residence countries	49		48	
R^2	0.198		0.275	
Oster (2019) diagnostics	Patience	Risk-taking	Patience	Risk-taking
δ to match $\beta_{1,2} = 0$	-18.093	8.224	13.575	9.761
Bound β^* for $\delta = 1$	1.358	-1.181	0.875	-0.264

Notes: Dependent variable: PISA math test score, col. 1: waves 2000-2018, col. 2: waves 2003-2018. Least squares regressions. Col. 1: restricted model includes wave fixed effects; baseline model adds controls for student gender, age, migration status, and imputation dummies. Col. 2: restricted model includes residence-country by wave fixed effects and dummy for OECD country of origin; baseline model adds controls for student gender, age, and imputation dummies. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2000-2018; Falk et al. (2018).

Table 15: Migrant analysis: Models with residence-country and wave fixed effects (but not their interaction)

	(1)	(2)	(3)
Patience (country-of-origin)	0.776***		0.929***
	(0.114)		*(0.117)*
Risk-taking (country-of-origin)		0.188	-0.291**
		(0.202)	*(0.125)*
Residence-country fixed effects	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes
Residence-country by wave fixed effects	No	No	No
Control variables	Yes	Yes	Yes
Observations	80,398	80,398	80,398
Countries of origin	58	58	58
Residence countries	48	48	48
R^2	0.265	0.247	0.267

Notes: Dependent variable: PISA math test score, waves 2003-2018. Least squares regressions, including 48 fixed effects for residence countries and six fixed effects for waves. Sample: students with both parents not born in the country where the student attends school. Control variables: student gender, age, dummy for OECD country of origin, imputation dummies. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2003-2018; Falk et al. (2018).

Table 16: Migrant analysis: Subgroups by age of migration

	Second generation	First generation		
		All	Before age 6	After age 6
	(1)	(2)	(3)	(4)
Patience (country-of-origin)	1.023***	0.955***	1.010***	0.981***
	(0.143)	*(0.120)*	*(0.156)*	*(0.103)*
Risk-taking (country-of-origin)	-0.458***	-0.185	-0.228	-0.153
	(0.127)	*(0.145)*	*(0.145)*	*(0.146)*
Residence-country by wave fixed effects	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Observations	47,369	33,029	14,459	16,835
Countries of origin	56	57	51	55
Residence countries	48	48	47	48
R^2	0.297	0.263	0.298	0.258
Difference between subsamples				
Patience (country-of-origin)	-0.068		-0.029	
	(0.085)		*(0.114)*	
Risk-taking (country-of-origin)	0.273**		0.075	
	(0.122)		*(0.062)*	

Notes: Dependent variable: PISA math test score, waves 2003-2018. Least squares regressions. Sample: students with both parents not born in the country where the student attends school. Second generation: migrant students born in the country of residence. First generation: migrant students born in the country of origin; split between whether they migrated to the country of residence before or after age 6 in col. 3 and 4. Control variables: student gender, age, dummy for OECD country of origin, imputation dummies. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2003-2018; Falk et al. (2018).

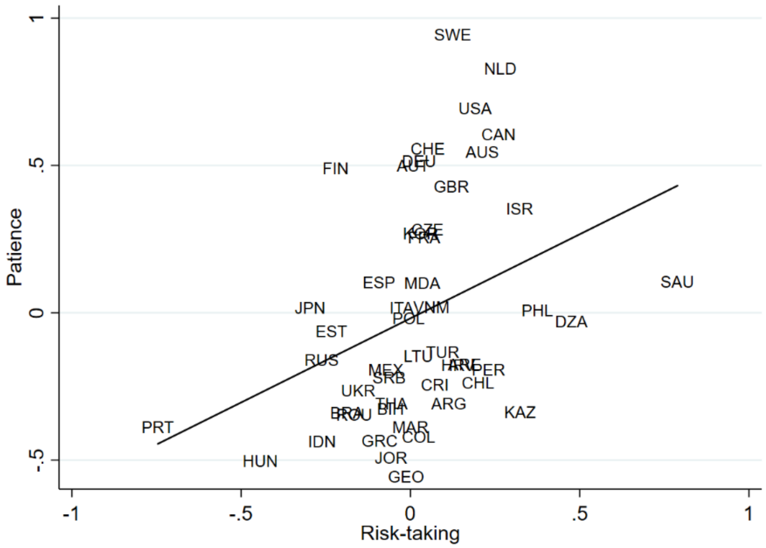
Table 17: Migrant analysis: Different definitions of migrants

	Narrow definition			Wide definition				Dropping countries of origin with <50 observations
	Mother's origin	Parental average	Separate	Mother's origin	Father's origin	Parental average	Separate	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience (mother's country-of-origin)	0.931***		0.343***	0.861***			0.349***	
	(0.109)		*(0.069)*	*(0.109)*			*(0.070)*	
Risk-taking (mother's country-of-origin)	-0.292**		0.032	-0.228*			0.038	
	(0.126)		*(0.086)*	*(0.121)*			*(0.087)*	
Patience (father's country-of-origin)			0.629***		0.858***		0.627***	0.939***
			(0.090)		*(0.112)*		*(0.093)*	*(0.116)*
Risk-taking (father's country-of-origin)			-0.339***		-0.233*		-0.336***	-0.299**
			(0.090)		*(0.119)*		*(0.093)*	*(0.122)*
Patience (average parents' country-of-origin)		0.941***				0.858***		
		(0.115)				*(0.111)*		
Risk-taking (average parents' country-of-origin)		-0.273**				-0.217*		
		(0.124)				*(0.120)*		
Residence-country by wave fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	80,194	83,798	76,796	140,951	141,155	145,506	85,167	80,221
Countries of origin	57	58	58	60	60	60	59	46
Residence countries	48	48	48	48	48	48	48	48
R^2	0.278	0.274	0.280	0.255	0.254	0.254	0.279	0.275

Notes: Dependent variable: PISA math test score, waves 2003-2018. Least squares regressions. Sample: migrant students; see text for narrow and wide definition of migrant status. Control variables: student gender, age, dummy for OECD country of origin, imputation dummies. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2003-2018; Falk et al. (2018).

Appendix Figure

Figure 2: Patience and risk-taking across countries



Notes: Country averages. Data source: Falk et al. (2018).