ECONtribute

Discussion Paper

Patience and Comparative Development

Uwe Sunde Thomas Dohmen Benjamin Enke Armin Falk

David Huffman Gerrit Meyerheim

August 2020

ECONtribute Discussion Paper No. 035

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866 is gratefully acknowledged.

Patience and Comparative Development*

Uwe Sunde Thomas Dohmen Benjamin Enke Armin Falk
David Huffman Gerrit Meyerheim

August 12, 2020

Abstract

This paper studies the relationship between patience and comparative development through a combination of reduced-form analyses and model estimations. Based on a globally representative dataset on time preference in 76 countries, we document two sets of stylized facts. First, patience is strongly correlated with both per capita income and the accumulation of physical capital, human capital and productivity. These correlations hold across countries, subnational regions, and individuals. Second, the quantitative magnitude of the patience elasticity strongly increases in the level of aggregation. To provide an interpretive lens for these patterns, we analyze an OLG model in which savings and education decisions are endogenous to patience, and aggregate production is characterized by capital-skill complementarities. This model reconciles both the correlations between patience and macroeconomic variables as well as the substantial amplification of patience elasticities at higher levels of aggregation. The results of model estimations resemble the reduced-form patterns and suggest that cross-country variation in productivity plays an important role in generating the observed comparative development patterns and aggregation effects.

JEL-classification: D03, D90, O10, O30, O40
Keywords: Time Preference, Comparative Development
Factor Accumulation

*Sunde: University of Munich, Department of Economics; uwe.sunde@lmu.de. Falk: Institute on Behavior and Inequality (briq) and University of Bonn; Armin.Falk@briq-institute.org. Dohmen: University of Bonn, Department of Economics; tdohmen@uni-bonn.de, armin.falk@uni-bonn.de. Enke: Harvard University, Department of Economics; enke@fas.harvard.edu. Huffman: University of Pittsburgh, Department of Economics; huffmand@pitt.edu. Meyerheim: University of Munich, Department of Economics; gerrit.meyerheim@econ.lmu.de.
Uwe Sunde acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) CRC TRR 190 (project number 280092119, Project A07). Armin Falk and Thomas Dohmen acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project A01) and Germany's Excellence Strategy - EXC 2126/1- 390838866.

1 Introduction

A long stream of research in development accounting has documented that both production factors and productivity play an important role in explaining international income differences (Hall and Jones, 1999; Caselli, 2005; Hsieh and Klenow, 2010). By its nature, this line of work does not speak to the reasons why countries or subnational regions exhibit variation in these proximate determinants of comparative development in the first place. According to standard economic theory, the stocks of physical capital, human capital, or research intensity all arise ultimately from an investment process that crucially depends on the same structural parameter of time preference (e.g., Becker, 1962; Ben-Porath, 1967; Romer, 1990; Aghion and Howitt, 1992; Doepke and Zilibotti, 2014). However, perhaps due to a previous lack of reliable and comparable data on time preference on a global scale, the relationship between patience and comparative development is not well-explored. To make progress, this paper utilizes a recently constructed globally representative dataset on patience to present a new set of stylized facts about the relationships between patience, accumulation processes and income at different levels of aggregation. To interpret these stylized facts, we analyze and quantitatively estimate an overlapping generations (OLG) model with cross-national and cross-individual heterogeneity in patience.

Our empirical analysis is based on the Global Preference Survey (GPS), a recently constructed global dataset on economic preferences from representative population samples in 76 countries (Falk et al., 2018). In this survey, patience was measured through a series of structured questions such as hypothetical choices between immediate and delayed monetary rewards. To ensure comparability of preference measures across countries, the survey items underwent an extensive ex ante experimental validation and selection procedure, and the cross-country elicitation followed a standardized protocol that was implemented through the professional infrastructure of the Gallup World Poll. Monetary stakes involved comparable values in terms of purchasing power across countries, and the survey items were culturally neutral and translated using state-of-the-art procedures. Thus, the data provide an ideal basis for the first systematic analysis of the relationship between patience and investment decisions at the micro and macro levels.

Using these data, we present a new set of stylized facts about the relationship between patience, the accumulation of production factors and income at various levels of aggregation. Across countries, average patience statistically explains about 40% of the between-country variation in (log) per capita income (Falk et al., 2018). This reduced-form relationship is shown to be robust across a wide range of empirical specifications, which incorporate controls for many of the deep determinants identified in the comparative development literature, such as geography, climate, the disease environment, anthropological factors, and social capital.

Because canonical macroeconomic models posit that heterogeneity in patience matters for income through its impact on accumulation decisions, we also investigate the correlations between patience and the proximate determinants of development. Here, we find that average patience is also strongly correlated with cross-country variation in capital stocks, savings rates, educational attainment, and total factor productivity (TFP).

While our analyses are correlational in nature, we investigate to what extent the link between patience and cross-country development is likely to be spurious. For instance, measured patience might not reflect actual time preference but instead be confounded by local inflation and interest rates or the quality of the institutional environment. Similarly, patience may be endogenous to education. While controlling for potentially noisy measures is no panacea for omitted variable bias, we gauge the role of these potential confounds for our analysis by controlling for inflation and interest rates, objective and subjective institutional quality, life expectancy, and educational attainment. We find that country-level patience remains strongly correlated with per capita income conditional on these covariates. We also show that the correlations between preferences and macroeconomic variables are specific to patience: none of the other measures from the GPS (such as risk aversion, altruism or trust) are robustly related to income or accumulation.

Next, we leave the realm of cross-country regressions to study subnational and individual heterogeneity in patience, income and accumulation processes. First, akin to the approach taken by Gennaioli et al. (2013), we present estimations that link average regional patience to regional per capita income and educational attainment. While the corresponding regressions investigate the correlates of patience at an aggregate level, as called for by development theories, they also allow us to keep many factors such as the overall institutional environment constant by including country fixed effects. The results reveal robust evidence that, within countries, regions with more patient populations exhibit higher average educational attainment and higher per capita income. Second, we present conceptually analogous analyses across individuals, holding fixed the country or subnational region of residence. Here, again, patience is robustly correlated with a greater propensity to save, higher educational attainment, and higher household income. Taken together, our analysis shows that patience is consistently correlated with income and factor accumulation across levels of aggregation. The within-country and within-region results arguably go a long way towards ruling out that variation in institutional quality or survey interpretation confound the correlation between patience and income.

Another finding that emerges from the analysis at different levels is a quantitatively large amplification effect: the elasticity of the dependent variables with respect to patience strongly increases in the level of aggregation. This is the case in two conceptually related ways. First, restricting attention to across-region (or across-individual) analyses, the patience coefficient in income regressions drops by a factor of 6 – 7 once country fixed effects are included. Second, comparing across-country, across-region and across-individual regressions, the patience coefficient suggests that a one-standard deviation increase in patience is associated with an increase in income per capita of 1.73 log points across countries, of 0.17 log points across regions within countries, and of 0.05 log points across individuals within countries. These large differences in coefficient estimates hold even though our individual-level results are in line with the micro literature on preference heterogeneity. We discuss and simulate measurement error and resulting attenuation bias as a potential driver of aggregation effects, but conclude that attenuation alone is very unlikely to generate the observed patterns. For example, the patience coefficient in individual-level regressions is much smaller in specifications with country fixed effects; this shows a smaller elasticity within country, which is consistent with an amplification but cannot be explained by greater measurement error.

To provide an interpretive lens for this collection of new stylized facts, we analyze a three-period general equilibrium OLG model in which heterogeneity in patience affects individual savings and education decisions. Aggregate production is characterized by capital-skill complementarities. As a result, the accumulation of physical capital and human capital (and, hence, factor incomes) feeds back into individual decisions through general equilibrium effects.

At the level of individual decision makers, the model delivers intuitive predictions, such as that individuals who exhibit higher patience have a higher propensity to become skilled, save more, and have higher lifetime incomes. Analogous qualitative predictions hold when comparing two economies that differ only in their average level of patience. However, as a consequence of general equilibrium effects (pecuniary externalities), the quantitative magnitude of the elasticity of income with respect to average patience can be amplified relative to its individual-level analogue.

We then use the model to evaluate whether the systematic amplification in reduced-form coefficient estimates can plausibly be generated by aggregation effects in general equilibrium. For this purpose, we consider two thought experiments: (i) marginally increasing individual-level patience, holding average patience, aggregate allocations and prices fixed; (ii) marginally increasing average patience, which leads to changes in aggregate allocations and prices. We quantify the model by calibrating the standard parameters based on estimates from the literature and estimate the structural parameters using an indirect inference approach. We estimate the model parameters by matching the empirical patience elasticities that we obtain from the reduced form estimates. This approach allows us to assess the plausibility of the estimated parameters as well as other moments implied by the model.

We analyze two main model specifications. In the first one, total factor productivity is assumed to be fixed at the same level for both economies, so that patience can only matter for the accumulation of physical and human capital. In the second version, TFP implicitly depends on patience through a standard human capital externality specification that we parameterize using estimates from the literature. A helpful way to think about these two model variants is that the first one – in which TFP is fixed – corresponds to empirical estimates across subnational regions, while the second one corresponds to cross-country comparisons in which the broader productivity environment (e.g., national policies) also varies. In both of these quantitative exercises, we match as moments the empirical patience elasticities in our data, and estimate, inter alia, average patience in each of the two economies.

The model estimations deliver sensible parameter values. For example, in both of our main specifications, we estimate average annual discount factors of 0.93 – 0.95. When we simulate the model using the estimated parameter values, the implied patience elasticity is 2.5 times as large at the country relative to the individual level when TFP is fixed at the same value for both countries. While this is a substantial amplification effect, it does not fully account for the empirically observed amplification. However, in our second specification, in which TFP is allowed to vary across economies and implicitly depends on patience through a human capital externality, the estimated amplification effects in the model become substantially larger and increase by a factor of 20 or more.

These results suggest that variation in productivity – induced by patience, perhaps through heterogeneity in factor accumulation as exemplified by human capital externalities – play an important role in generating the observed aggregation effects in the data. At the same time, the quantitative analysis suggests that general equilibrium effects play an important role for aggregation effects through their effects on the accumulation of production factors. These results resonate with the patterns in our reduced-form analysis: the observed amplification going from individual- to regional-level estimates is smaller than the one going from individual- to country-level estimates. Our model offers a potential explanation, showing how differences are magnified at the country level by the dependence of national policies on patience, whereas such policies are roughly constant across sub-national regions.

To benchmark our results on the role of patience, we implement a third model specification that resembles conventional development accounting exercises. Here, we only allow TFP to vary across the two model economies, but hold average patience fixed. Compared to our preferred model specification in which TFP is endogenously determined through a human capital externality (and, hence, patience), this version of the model can account for large income differences across economies but is unable to explain variation in accumulated factors. Moreover, this model variant requires almost 30% more variation in TFP across economies than the implied variation in TFP in the version of the model with heterogeneity in patience and a human capital externality on TFP. This analysis suggests that the assumption of patience-induced variation in TFP fits the data better than imposing exogenous technology differences to explain variation in the stocks of production.

This paper contributes to two lines of research in the literature on comparative development. First, research in development accounting decomposes national income into production factors and productivity (the proximate determinants of development). Second, research on the deep determinants of development focuses on the roles of geography, climate, history, or social capital (e.g., Knack and Keefer, 1997; Olsson and Hibbs Jr, 2005; Spolaore and Wacziarg, 2009; Algan and Cahuc, 2010; Ashraf and Galor, 2013). Our paper relates to the development accounting literature in that it analyzes a potential mechanism that can generate variation in the proximate determinants of development (e.g., Doepke and Zilibotti, 2008, 2014). However, instead of attributing differences in the accumulated factors to exogenous variation in productivity or institutions (Hsieh and Klenow, 2010), our results suggest that variation in cultural factors as reflected by patience can explain heterogeneity in income and in productivity, once one allows for externalities that work through accumulated factors. At the same time, because our paper is largely descriptive and takes patience as given, our work builds on contributions in the deep determinants literature that have pointed to the potential long-run origins of variation in patience (Chen, 2013; Galor and Özak, 2016).

Our paper also contributes to a recent line of work that studies the effects of human capital accumulation on growth (Gennaioli et al., 2013; Squicciarini and Voigtländer, 2015). Several contributions have shown that more realistic representations of the human capital accumulation process account for a considerably higher fraction of income variation than previously thought (see, e.g., Erosa et al., 2010; Caselli and Ciccone, 2013; Manuelli and Seshadri, 2014). Our paper contributes to this literature by providing micro evidence for one hitherto unexplored mechanism (preference heterogeneity) that may generate variation in human capital. Our focus on preference heterogeneity also connects to recent papers on cross-country variation in hours worked (Jones and Klenow, 2016; Bick et al., 2018).

The remainder of the paper proceeds as follows. The data are described in Section 2. Section 3 presents empirical evidence for the reduced-form relationships between patience and development at the individual and aggregate level. Sections 4 and 5 present and estimate the model. Section 6 offers a concluding discussion.

2 Data

Our analysis relies on the Global Preference Survey (GPS), a recently constructed data set on economic preferences from representative population samples in 76 countries. In many countries around the world, the Gallup World Poll regularly surveys representative population samples about social and economic issues. We created the GPS by adding a set of survey items that were explicitly designed to measure a respondent's time preferences, risk preferences, social preferences, and trust, as part of the regular 2012 questionnaire (for details see Falk et al., 2018).

Four features make these data suited for the present study. First, the preference measures were elicited in a comparable way using a standardized protocol across countries.

Second, the data cover representative population samples in each country, which allows for inference about between-country differences in preferences. The median sample size was $N = 1,000$ per country, for a total of $80,000$ respondents worldwide. Respondents were selected through probability sampling and interviewed face-to-face or via telephone by professional interviewers. A third feature of the data is geographical representativeness in terms of the countries being covered. The sample of 76 countries is not restricted to Western industrialized nations, but covers all continents and various levels of development.

Fourth, the preference measures are based on experimentally validated survey items for eliciting preferences. To ensure the behavioral relevance of the measure of patience, the underlying survey items were designed, tested, and selected for the purpose of the GPS through a rigorous ex-ante experimental validation procedure (for details see Falk et al., 2016). In this validation step, subjects participated in choice experiments that measured preferences using real money. They also answered large batteries of survey questions designed to elicit preferences. We then selected those survey items that were (jointly) the best predictors of actual behavior in the experiments, to form the survey module. In order to make these items cross-culturally applicable, (i) all items were translated back and forth by professionals; (ii) monetary values used in the survey were adjusted based on the median household income for each country; and (iii) pretests were conducted in 22 countries of various cultural heritage to ensure comparability. See Appendix A and Falk et al. (2018) for a description of the data set and the data collection procedure.

Patience is derived from the combination of responses to two survey measures, one with a quantitative and the other with a qualitative format. The quantitative survey measure consists of a series of five interdependent hypothetical binary choices between immediate and delayed financial rewards, a format commonly referred to as the "staircase" (or unfolding brackets) procedure. In each of the five questions, participants had to decide between receiving a payment today or a larger payment in twelve months:

> *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?*

The immediate payment x remained constant in all four subsequent questions, but the delayed payment y was increased or decreased depending on previous choices (see Appendix A for an exposition of the entire sequence of binary decisions). In essence, by adjusting the delayed payment according to previous choices, the questions "zoom in" on the respondent's point of indifference between the smaller immediate and the larger delayed payment, which makes efficient use of limited and costly survey time. The sequence of questions has 32 possible ordered outcomes that partition the real line from 100 Euros to 218 Euros into roughly evenly spaced intervals. In the international survey, the monetary amounts x and y were expressed in the respective local currency, scaled relative to the median monthly household income in the given country.

The qualitative measure of patience is given by the respondents' self-assessment of their their willingness to wait on an 11-point Likert scale:

> *We now ask for your willingness to act in a certain way. Please indicate your answer on a scale from* 0 *to* 10*, where* 0 *means you are "completely unwilling to do so" and a* 10 *means you are "very willing to do so". How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?*

Our patience measure is a linear combination of the quantitative and qualitative survey items, using the weights obtained from the experimental validation procedure.[1] As described in detail in Falk et al. (2016), the survey items are strongly and significantly correlated with preference measures obtained from standard incentivized intertemporal choice experiments. Moreover, the measures predict experimental behavior out of sample. The ex-ante validation of the survey items constitutes a methodological advance compared to the often ad-hoc selection of questions for surveys.

A clear advantage of the quantitative staircase measure relative to the qualitative one is that it closely resembles standard experimental procedures of eliciting time preferences and corresponds to how economists typically think about immediate versus delayed rewards. In addition, the measure is context neutral and precisely defined, making it less prone to culture-dependent interpretations. In recent work, Bauer et al. (2020) show that quantitative (staircase-type) survey questions reliably measure preferences also outside the Western world, while this is not necessarily the case for more qualitative questions like subjective self-assessments. Indeed, it turns out that the relationship between patience and comparative development that we identify below is almost entirely driven by the quantitative measure. Still, the analysis relies on the composite patience measure as it was developed in the experimental validation procedure.

The analysis is based on individual-level patience measures that are standardized, i.e., we compute z-scores at the individual level. We then calculate a country's patience by averaging responses using the sampling weights provided by Gallup, see Appendix A. In all figures and regressions, patience is scaled in the same manner, regardless of whether the

[1] Specifically, responses to both items were standardized at the individual level and then aggregated:

$$\text{Patience} = 0.7115185 \times \text{ Staircase measure } + 0.2884815 \times \text{ Qualitative measure } ,$$

with weights being based on OLS estimates of a regression of observed behavior in financially incentivized laboratory experiments on the two survey measures. See Falk et al. (2016, 2018) for details.

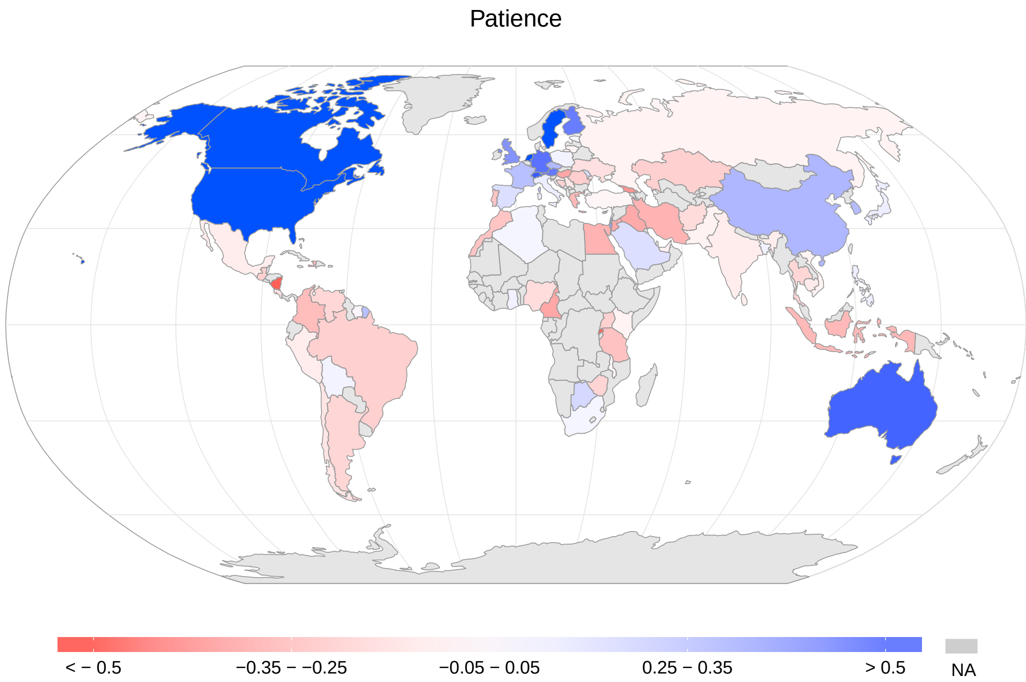


Figure 1: Distribution of patience across countries

level of aggregation is the individual, a subnational region, or a country. Figure 1 depicts the resulting distribution of patience across countries, relative to the world's average individual. Darker red colors and darker blue colors indicate less and more patience, respectively, where differences are measured in terms of standard deviations from the world's average individual, which is colored in white.[2]

All other data used in this paper stem from standard sources such as the World Bank's World Development Indicators or the Penn World Tables. Appendix A describes all variables and their sources.

Summary statistics. Our individual-level data contain 80,377 respondents from 76 countries. Average age in our sample is 41.8 and 54% of all respondents are female. The individual-level patience index is correlated with demographics, as reported in Falk et al. (2018). Women are slightly less patient than men ($\rho = 0.04$), and respondents' subjective self-assessment of their math skills ($0 - 10$) is positively correlated with patience ($\rho = 0.13$). As discussed in Falk et al. (2018), there is a hump-shaped relationship between patience and age. In a joint regression, age, age squared, gender and subjective math skills explain

[2]The variation in patience appears to reflect idiosyncratic variation that is not well-captured by other aspects of cultural variation. For example, the correlations between patience and trust and between patience and risk taking are only $\rho = 0.19$ and $\rho = 0.23$. Moreover, as shown below, the well-known correlation between trust and per capita income vanishes once patience is controlled for.

about 2% of the global individual-level variation in measured patience.

3 Patience and Development: Empirical Evidence

A large body of theoretical work links heterogeneity in patience to the accumulation of production factors, and, hence, income (e.g., Becker, 1962; Ben-Porath, 1967). Motivated by this body of theoretical work, this section presents descriptive evidence on the relationship between patience, the accumulation of productive resources and income at three different levels of aggregation: across countries, across subnational regions, and across individuals.

3.1 Cross-Country Evidence

3.1.1 Patience and Income

Table 1 presents the results of a set of OLS regressions of per capita income on patience. Column (1) documents that a one standard deviation increase in patience is associated with an increase in per capita income of 2.32 log points. The raw correlation between the log of GDP per capita and the patience measure is 0.63, implying that patience alone statistically "explains" about 39% of the variation in log income per capita; also see Falk et al. (2018).[3] Columns (2) through (4) successively add a comprehensive set of geographic and climatic covariates. Column (2) contains controls for world regions.[4] Column (3) contains additional controls for absolute latitude, longitude, the fraction of arable land, and land suitability for agriculture. Column (4) adds average precipitation and temperature as well as the fractions of the population that live in the (sub-) tropics or in areas wherethere exists the risk of contracting malaria. Finally, column (5) additionally controls for trust, and genetic diversity and its square. While the inclusion of this large vector of covariates reduces the coefficient of patience by about 25%, it remains statistically significant and quantitatively large. Interestingly, the evidence indicates that trust, which has previously been identified as a driver of development (Knack and Keefer, 1997; Guiso et al., 2009; Algan and Cahuc, 2010; Tabellini, 2010), adds little to the explanatory power once patience is included in the analysis. Figure 2 illustrates the conditional relationship for the estimates of column (5).

Robustness checks. Appendix D presents two sets of robustness checks. First, Table D.1 documents that additionally controlling for average risk aversion, legal origin dummies, ethnic, religious, and linguistic fractionalization, major religion shares, the fraction of

[3]The coefficient estimate in column (1) slightly differs from the one reported in Falk et al. (2018) because the regressions utilize different GDP data.

[4]Following the World Bank terminology, world regions are defined as North America, Central and South America, Europe and Central Asia, East Asia and Pacific, South Asia, Middle East and North Africa, and South Africa.

Table 1: Patience and national income

	Dependent variable: Log [GDP p/c]				
	(1)	(2)	(3)	(4)	(5)
Patience	2.32***	1.84***	1.60***	1.56***	1.73***
	(0.23)	(0.24)	(0.30)	(0.30)	(0.28)
Distance to equator			0.011	-0.0030	-0.033*
			(0.01)	(0.02)	(0.02)
Longitude			0.0023	0.0055	0.0077
			(0.01)	(0.01)	(0.01)
Percentage of arable land			-0.021*	-0.011	-0.0078
			(0.01)	(0.01)	(0.01)
Land suitability for agriculture			0.38	-0.10	0.15
			(0.66)	(0.48)	(0.44)
Average precipitation				0.0060	0.0019
				(0.00)	(0.00)
Average temperature				0.041*	0.013
				(0.02)	(0.02)
% living in (sub-)tropical zones				-1.29*	-1.18**
				(0.65)	(0.57)
% at risk of malaria				-1.45***	-1.46***
				(0.44)	(0.41)
Predicted genetic diversity					513.2***
					(130.93)
Predicted genetic diversity sqr.					-365.1***
					(96.08)
Trust					-0.076
					(0.42)
Continent FE	No	Yes	Yes	Yes	Yes
Observations	76	76	75	75	74
R^2	0.39	0.69	0.72	0.81	0.84

OLS estimates, robust standard errors in parentheses. * ($p < 0.10$), ** ($p < 0.05$), ***($p < 0.01$)

European descent, the genetic distance to the U.S., and other geographical variables, does not affect the qualitative results. Second, Table D.2 documents that the relationship between patience and per capita income robustly appears in various sub-samples, i.e., within each continent, within OECD or non-OECD countries, or within former colonies and countries that have never been colonized.

Growth extension. Table D.3 in Appendix D presents an extension of the results on cross-national income differences by considering the link between patience and growth rates since World War II. To this end, we compute the (geometric) average annual growth rate in per capita GDP from different base years until 2015. We find that patience is robustly correlated with medium-run growth rates, both in univariate regressions and

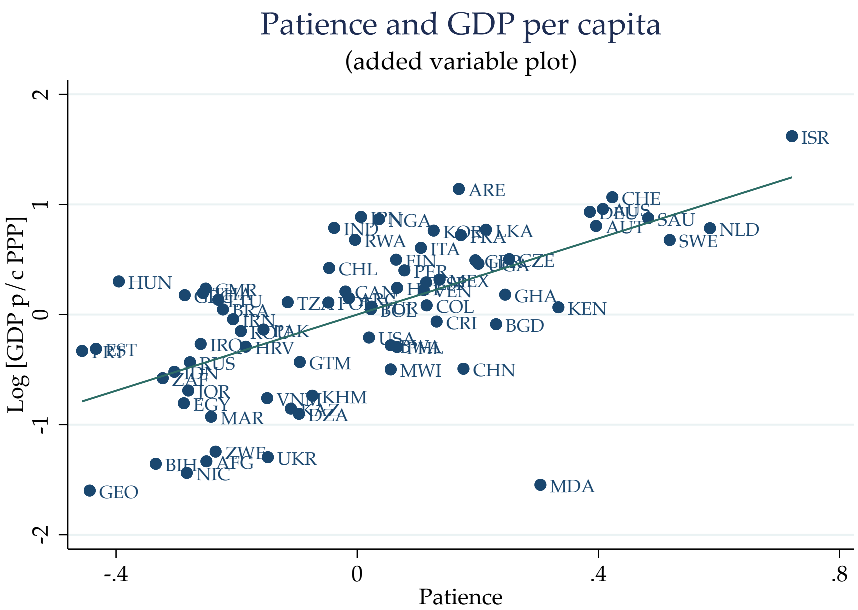


Figure 2: Patience and national income (added variable plot conditional on the full set of covariates in column (5) of Table 1).

when we control for per capita income in the base year and additional covariates.

3.1.2 Patience and Accumulation Processes

In standard textbook models, a reduced-form relationship between patience and development operates through accumulation processes. We therefore investigate whether patience is related to the levels of production factors and productivity as well as the corresponding accumulation flows.

Physical Capital. To understand the relationship between patience and physical capital, we regress the stock of physical capital as well as three separate savings variables on patience. For each dependent variable, Table 2 presents OLS estimates of the unconditional relationship and of the relationship conditional on the full set of covariates from column (5) in Table 1.

Columns (1) and (2) document that patience is strongly correlated with the stock of physical capital, also conditional on controls. Columns (3) to (8) of Table 2 present the results for gross national savings rates, net adjusted national savings rates, and household savings rates as dependent variables. Gross savings rates are given by gross national income net of consumption, plus net transfers, as a share of gross national income. Net adjusted savings rates correspond to gross savings net of depreciation, adding education expenditures and deducting estimates for the depletion of energy, minerals and forests, as well as damages from carbon dioxide emissions. Household savings rates are measured as household savings relative to household disposable income. These data are based on

Table 2: Patience, physical capital, and savings

	Dependent variable:							
	Log [Capital stock p/c]		Gross savings (% of GNI)		Net adj. savings (% of GNI)		HH savings (% of disposable inc.)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience	1.94***	1.17***	7.43***	8.91***	6.08**	7.16*	8.52***	9.80***
	(0.27)	(0.29)	(2.41)	(3.27)	(2.34)	(3.62)	(2.72)	(3.31)
Continent FE	No	Yes	No	Yes	No	Yes	No	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	No
Observations	71	69	75	73	73	71	26	26
R^2	0.32	0.83	0.07	0.36	0.04	0.38	0.15	0.32

OLS estimates, robust standard errors in parentheses. Due to the small number of observations, in column (8), the controls are restricted to continent dummies. See column (5) of Table 1 for a complete list of the additional controls. * $(p < 0.10)$, ** $(p < 0.05)$, *** $(p < 0.01)$

Table 3: Patience and human capital

	Dependent variable:					
	% Skilled			Yrs. of schooling		
	(1)	(2)	(3)	(4)	(5)	(6)
Patience	38.5***	21.9***	20.1***	4.34***	2.58***	2.47***
	(5.45)	(7.15)	(7.20)	(0.58)	(0.71)	(0.86)
Continent FE	No	Yes	Yes	No	Yes	Yes
Additional controls	No	No	Yes	No	No	Yes
Observations	72	72	71	72	72	71
R^2	0.30	0.56	0.73	0.34	0.62	0.76

OLS estimates, robust standard errors in parentheses. The percentage skilled is the percentage of individuals aged 25+ that has at least secondary education (Barro and Lee, 2012). See column (5) of Table 1 for a complete list of the additional controls. * $(p < 0.10)$, ** $(p < 0.05)$, *** $(p < 0.01)$

surveys and are only available for OECD countries. Throughout, the results reveal a significant positive relationship between patience and savings. The finding that variation in patience is related to cross-country variation in household savings rates even within OECD countries is arguably noteworthy, given the similarity of this subset of countries in terms of economic development and other characteristics.

Human Capital. As measures of human capital, we consider proxies for both the quantity and quality of schooling. Our dependent variables are (i) the fraction of the population aged over 25 that has at least secondary education (Barro and Lee, 2012) and (ii) average years of schooling. Columns (1) – (4) of Table 3 report the results. The patience variable is robustly correlated with human capital, and statistically explains between 30% and 34% of the variation in the aforementioned variables.

Productivity. Endogenous growth models highlight the role of patience for the accumulation of ideas and knowledge through research. Relatedly, factor productivity implicitly depends on patience in models that assume human capital externalities. Table 4 documents that patience is strongly correlated with both the TFP measure from the PWT and the number of researchers in research and development (which serves as proxy for human capital externalities). For both dependent variables, the variance explained is again roughly 30%.

3.1.3 Assessing Endogeneity Concerns

While standard models such as the one presented in Section 4 below implicitly presume a causal role of patience for accumulation processes and income, a causal interpretation of our reduced form empirical results is subject to several potential criticisms: (i) the patience variable might not only measure patience but may reflect additional features of the external environment such as institutions, inflation, or interest rates; and (ii) the OLS correlations be driven by omitted variables or reverse causality.

We do not claim that our analysis rules out all potential endogeneity concerns. Rather, we view this paper as a first contribution that studies the systematic relationship between patience, accumulation and income, and documents a novel set of stylized facts. Nonetheless, this section takes a more nuanced look at the data by investigating the extent to which the cross-country correlation between patience and per capita income is likely to be driven by omitted variables, measurement issues, or reverse causality.

Borrowing Constraints. Respondents might be more likely to opt for immediate payments in experimental choice situations if they face upward sloping income profiles and are borrowing constrained. To address this issue, we leverage the idea that borrowing constraints are likely to be less binding for relatively affluent people. We therefore employ the average patience of each country's top income quintile as an explanatory variable. As shown in column (1) of Table 5, the reduced-form relationship between patience and per capita income remains strong and significant using this patience measure.

Inflation and Interest Rates. If some respondents expect higher levels of inflation than others, or live in an environment with higher nominal interest rates, they might appear more impatient in their survey responses, even if they have the same time preference. Note, however, that the quantitative survey question explicitly asked people to imagine that there was zero inflation. Furthermore, we check robustness to this concern empirically by explicitly controlling for inflation (the GDP deflator) and deposit interest rates. We find that the reduced-form coefficient of patience remains quantitatively large and highly statistically significant after controlling for these factors; see column (2) of Table 5.

Table 4: Patience, productivity and R & D

	Dependent variable:					
	TFP			Log [# Researchers in R&D]		
	(1)	(2)	(3)	(4)	(5)	(6)
Patience	0.29***	0.28***	0.17**	2.70***	1.87***	1.49***
	(0.05)	(0.05)	(0.07)	(0.35)	(0.35)	(0.50)
Continent FE	No	Yes	Yes	No	Yes	Yes
Additional controls	No	No	Yes	No	No	Yes
Observations	59	59	58	69	69	68
R^2	0.29	0.52	0.70	0.35	0.71	0.83

OLS estimates, robust standard errors in parentheses. Number of researchers in R & D are per $1,000$ population. Columns $(1) - (3)$ exclude Zimbabwe because it is an extreme *upward* outlier in the TFP data from the Penn World Tables, which is likely due to measurement error. See column (5) of Table 1 for a complete list of the additional controls. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

Subjective Uncertainty. In the quantitative decision tasks between money today and in 12 months, respondents may face subjective uncertainty about whether they would actually receive the (hypothetical) money in the future. Such subjective uncertainty is likely correlated with, or caused by, weak property rights or other institutions. Similarly, respondents may face high subjective uncertainty about receiving future payments if their remaining life expectancy is low. To provide a first pass at assessing the relevance of these considerations, we condition on both objective and subjective measures of the quality of the institutional environment as well as people's life expectancy. First, in column (3) of Table 5 we control for a property rights and a democracy index. Second, in column (4), we make use of the fact that Gallup's background data contain a series of questions that ask respondents to asses their confidence in various aspects of their institutional environment, including the national government, the legal system and courts, the honesty of elections, and the military. In column (5) we control for average life expectancy at birth. The results show that patience continues to be a strong correlate of national income, conditional on objective or subjective institutional quality, or life expectancy.

Cognitive skills and education. Our survey requires respondents to think through abstract choice problems, which might be unfamiliar and cognitively challenging for some participants. This could induce people to decide based on heuristics. Column (6) of Table 5 regresses GDP per capita jointly on patience and average years of schooling, and patience remains highly significant and large in magnitude. Similarly, column (7) shows that patience is significantly correlated with per capita income conditional on a measure of standardized math and science test scores (Hanushek and Woessmann, 2012). Finally,

Table 5: Patience and per capita income: Robustness

	Dependent variable: Log [GDP p/c PPP]							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience of top income quintile	1.60*** (0.19)							
Patience		2.00*** (0.33)	0.77*** (0.27)	1.52*** (0.41)	1.04*** (0.24)	1.17*** (0.24)	1.37*** (0.27)	
GDP deflator		-0.068* (0.03)						
Deposit interest rate		0.037 (0.04)						
Property rights			0.029*** (0.01)					
Democracy			-0.012 (0.05)					
Subj. institutional quality				0.014 (0.01)				
Avg. life expectancy					0.12*** (0.02)			
Avg. years of education						0.24*** (0.05)		
Math and science test scores							0.63** (0.31)	
Patience (binarized staircase)								4.78*** (0.68)
Continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76	59	72	59	76	72	49	76
R^2	0.69	0.64	0.79	0.69	0.81	0.77	0.72	0.66

OLS estimates, robust standard errors in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

column (8) addresses the issue of decision heuristics. In particular, in the quantitative staircase procedure, respondents faced a series of five similar choices. Responses based on a simple heuristic such as "always money today/ in the future" might lead us to overestimate the true variance in patience. We hence generate a binarized individual-level patience index that equals one if the respondent opted for the future payment in the first question and zero otherwise. Even though this measure is much coarser than our composite patience index, it is significantly correlated with per capita income.

Income Effects. It is also conceivable that the correlation between patience and national income is driven by reverse causality, i.e., that higher income causes people to be more patient (or to behave as if they are more patient in our survey tasks). One way of investigating the plausibility of such an account is to examine the relationship between our patience measure and exogenous sources of income, such as oil rents. If it was true

Table 6: Other preference measures

	Dependent variable:						
	Log [GDP p/c]	% skilled	Years schooling	Log [Cap. stock p/c]	Gross savings (% GNI)	TFP	Log [researchers]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Patience	2.27***	37.2***	4.29***	1.80***	6.98**	0.24***	2.71***
	(0.27)	(6.28)	(0.68)	(0.26)	(3.26)	(0.06)	(0.30)
Risk taking	-0.90*	-4.67	-0.82	-0.95*	-2.79	0.050	-1.77***
	(0.45)	(9.75)	(0.94)	(0.49)	(4.76)	(0.08)	(0.65)
Trust	0.91*	7.57	0.34	0.98**	6.14	0.18*	0.39
	(0.49)	(9.97)	(1.02)	(0.46)	(4.82)	(0.10)	(0.59)
Altruism	-0.73	-25.3**	-3.03***	-1.05**	7.61*	-0.036	-0.94
	(0.51)	(10.09)	(1.10)	(0.44)	(4.02)	(0.09)	(0.62)
Pos. reciprocity	0.50	24.7**	2.58**	1.02**	-7.57*	-0.035	1.62**
	(0.51)	(11.74)	(1.15)	(0.51)	(4.39)	(0.12)	(0.65)
Neg. reciprocity	0.38	3.49	0.56	0.65	1.25	0.099	1.07**
	(0.48)	(9.96)	(1.05)	(0.42)	(3.54)	(0.09)	(0.51)
Observations	76	72	72	71	75	59	69
R^2	0.50	0.39	0.43	0.52	0.12	0.37	0.58

OLS estimates, robust standard errors in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

that higher income induces more patience in our procedures, then oil production (which is largely determined by natural resource endowments) should be correlated with patience. The left panel of Figure D.1 in Appendix D plots the raw correlation between log oil production per capita (measured in 2014 Dollars) and patience. The two variables are uncorrelated ($\rho = -0.04$), also conditional on the full set of controls in column (5) of Table 1. While these results do not rule out a causal link between income and patience, they provide an initial piece of evidence that the patience variable picks up variation that is independent of income effects.

3.1.4 Other Preference Measures

The GPS includes information not only about patience but also on risk aversion, trust, altruism, positive reciprocity and negative reciprocity. Table 6 replicates the unconditional analyses from above by including all GPS measures. The results show that patience is always significantly correlated with the outcomes of interest, also conditional on other preferences and trust. Other measures are only inconsistently related with outcomes (ee Falk et al. (2018) for a discussion of the correlation structure among the GPS measures).

3.2 Patience and Development Across Subnational Regions

In a second step of the empirical analysis, we turn to regressions across subnational regions. This is possible since the individual-level patience data in the GPS contain regional identifiers (usually at the state or province level), which allows us to relate the average level of patience in a sub-national region to the level of regional GDP per capita and the average years of education from data constructed by Gennaioli et al. (2013). In total, we were able to match 704 regions from 55 countries.[5]

Our analysis is motivated by a long literature in cultural economics that suggests that psychological variables might vary considerably also within countries. While the regional level of analysis still pertains to an aggregate view on accumulation processes and income, the corresponding regression analyses have the important advantage of allowing us to account for unobserved heterogeneity at the country-level by including country fixed effects. In particular, accounting for country fixed effects relaxes potential concerns about the role of language and institutions for survey responses. Indeed, Gennaioli et al. (2013) provide evidence that while human capital varies considerably even within countries and is strongly correlated with regional income, within-country variation in institutional quality is uncorrelated with regional development.

The benefits of considering regional data naturally come at the cost of losing representativeness, since the sampling scheme was constructed to achieve representativeness at the country level. In some regions, we observe only a relatively small number of respondents. As a consequence, average regional time preference is estimated less precisely than country-level patience. This matters for our analysis because measurement error in regional patience will lead to attenuation bias that makes comparing country- and regional-level results difficult. We pursue two strategies to account for measurement error. First, we exclude all regions with fewer than 15 respondents from the analysis, which leaves us with 648 regions. Second, we apply techniques from the recent social mobility literature (Chetty and Hendren, 2016) and shrink regional patience to the sample mean by its signal-to-noise ratio.[6]

Table 7 reports regression results for average average per capita income and education as dependent variables. We estimate one specification without country fixed effects, one with country fixed effects, and one with additional regional-level covariates (Gennaioli et al., 2013). The results qualitatively mirror those established in the country-level analysis: we find significant relationships between patience and per capita income, and between

[5] See Appendix B for an overview of the number of regions per country.

[6] Specifically, shrunk patience of region j, β_j^s, is computed as a convex combination of observed average patience in region i, β_j, and the mean $\bar{\beta}$ of the region sample averages β_j:

$$\beta_j^s = w_j \beta_j + (1 - w_j)\bar{\beta} \ , \tag{1}$$

where the region-specific weights are given by

$$w_j = \frac{Var(\beta_j) - \mathbb{E}[se_j^2]}{Var(\beta_j) - \mathbb{E}[se_j^2] + se_j^2} \ .$$

Here, $Var(\beta_j)$ is the variance of the regional means and se_j the standard error of β in region j. The results are quantitatively very similar if we do not exclude any regions and implement the shrinkage procedure on the full sample.

Table 7: Regional patience, human capital, and income

	Dependent variable:					
	Log [Regional GDP p/c]			Avg. years of education		
	(1)	(2)	(3)	(4)	(5)	(6)
Patience	1.40***	0.19***	0.17***	3.64***	0.51***	0.47***
	(0.24)	(0.06)	(0.06)	(0.62)	(0.16)	(0.16)
Temperature			-0.025**			-0.055***
			(0.01)			(0.01)
Inverse distance to coast			0.41			0.88
			(0.25)			(0.58)
Log [Oil production p/c]			0.30***			0.044
			(0.07)			(0.06)
# Ethnic groups			-0.10*			-0.25*
			(0.06)			(0.13)
Log [Population density]			0.071**			0.19***
			(0.03)			(0.06)
Country FE	No	Yes	Yes	No	Yes	Yes
Observations	648	648	631	637	637	620
R^2	0.20	0.93	0.94	0.29	0.94	0.95

Regional-level OLS estimates, standard errors (clustered at country level) in parentheses. Patience is shrunk patience, see equation (1). * ($p < 0.10$), ** ($p < 0.05$), ***($p < 0.01$)

patience and human capital, also conditional on country fixed effects.

Moving beyond the observation that patience is significantly correlated with income and education at the subnational level, it is interesting to note the quantitative magnitude of the coefficient estimates. In particular, for both dependent variables, the patience coefficient drops by a factor of seven once country fixed effects are included (columns (2) and (5)). A related observation refers to the across-region coefficient estimates, which are substantially smaller than the corresponding across-country estimates. We will return to this observation below when we discuss the role of aggregation effects.

3.3 Individual-Level Evidence

Finally, we study the relationship between patience, savings, education and income at the individual level using the GPS data. Table 8 presents the results of OLS regressions with three dependent variables: log household income per capita, a binary indicator for whether the respondent saved in the previous year, and a binary indicator for whether the respondent has at least secondary education. For each dependent variable, we report the results of four OLS specifications, one without any covariates, one with country fixed effects, one with regional fixed effects, and one with regional fixed effects and additional individual-level covariates.

The results document that patience is uniformly linked to higher income, a higher probability of saving, and a higher probability of becoming skilled. This pattern holds conditional on a comprehensive vector of individual-level covariates including age, age squared, gender, religion fixed effects, cognitive skills, and three variables that are proxies for the subjectively perceived quality of the institutional environment (these variables are collected and constructed by Gallup, see Appendix E).

For a subset of 13 countries, our dataset contains information on whether the respondent owns a credit card, which we think of as a proxy for access to credit. Table D.4 in Appendix D additionally controls for this binary indicator, with very similar results as in Table 8.

Moving beyond the qualitative patterns, we again see that in income regressions the coefficient estimate of patience drops by a factor of six once country fixed effects are included. This pattern is reminiscent of the results obtained in the regional-level analysis. We now turn to a first discussion of the mechanisms behind these aggregation effects.

3.4 Aggregation Effects: The Role of Measurement Error

Throughout the empirical analysis, the patience variable is expressed as z-score at the individual level, and then aggregated up to the regional or country level. This implies that the point estimates in the income regressions can be directly compared across levels of aggregation. An inspection of the first column in each of the corresponding tables reveals a country-level patience coefficient of 2.32, a regional level coefficient of 1.40, and an individual-level coefficient of 0.34. A different way to look at this pattern is that – in both the regional- and individual-level regressions – the patience coefficient drops by a factor of roughly seven once country fixed effects are included.[7]

A potential reason behind the large variation in coefficient estimates across levels of aggregation is measurement error and resulting attenuation bias. The relationship between individual income and patience should be more attenuated if individual patience is measured with more noise than country-level patience (as is likely the case). Similarly, it is almost certainly true that regional patience is measured with more error than country patience because of the smaller number of respondents. Thus, *some part* of the difference in patience coefficients between country-, regions- and individual-level analysis is likely to be driven by measurement error.

However, note that this argument does not explain why – within individual- or region-

[7]Our individual-level coefficient estimates are broadly in line with those obtained using other medium-scale micro datasets in the literature that focus on particular countries. While direct quantitative comparisons are complicated by the usage of different patience measures and income variables, the few benchmarks that we have reveal encouraging similarities. In the nationally representative German sample of Dohmen et al. (2010), the coefficient of (the z-score of) patience in a regression with log per capita income as outcome variable is 0.09. In a sample of U.S. respondents in the Health and Retirement Study (aged 70+), the same coefficient is 0.23 (Huffman et al., 2017), though the sample is clearly more special than ours.

Table 8: Individual patience, savings, human capital, and income

	Dependent variable:											
	Log [HH income p/c]				Saved last year				1 if at least secondary educ.			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Patience	0.34***	0.056***	0.049***	0.040***	0.051***	0.038***	0.038***	0.032***	0.061***	0.035***	0.033***	0.012***
	(0.05)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
Age				57.9***				-5.91				20.1
				(19.90)				(31.63)				(24.00)
Age squared				-3755.3				-559.6				-9428.7***
				(2307.05)				(3004.87)				(2243.20)
1 if female				-0.086***				-0.0057				-0.028***
				(0.02)				(0.01)				(0.01)
Subj. math skills				0.035***				0.017***				0.028***
				(0.00)				(0.00)				(0.00)
Subjective institutional quality				-0.042*				0.046				-0.062***
				(0.02)				(0.03)				(0.01)
Confidence in financial institutions				4.22***				5.15***				0.76
				(1.17)				(1.24)				(0.67)
Subjective law and order index				0.058**				0.012				0.00018
				(0.02)				(0.03)				(0.01)
Country FE	No	Yes	No	No	No	Yes	No	No	No	Yes	No	No
Regional FE	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Religion FE	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Observations	79245	79245	78271	46383	15260	15260	15260	10438	79357	79357	78403	46550
R^2	0.05	0.61	0.64	0.64	0.01	0.07	0.13	0.14	0.02	0.18	0.23	0.29

Individual-level OLS estimates, standard errors (clustered at country level) in parentheses. The dependent variable in (1)–(4) is ln household income per capita; the dependent variable in (5) – (8) is a binary indicator for whether the individual saved in the previous year; and the dependent variable in (9)–(12) is 1 if the individual has at least secondary education. Age is divided by 100. All results in columns (5) – (12) are robust to estimating probit models. See Appendix E for a detailed description of all dependent variables. * $(p < 0.10)$, ** $(p < 0.05)$, *** $(p < 0.01)$

level analyses – the coefficient drops by a factor of about seven once country fixed effects are included. After all, these regressions all rely on the same measure of patience. Instead, it appears that moving from a partly cross-country to a purely subnational comparison per se reduces the magnitude of the patience coefficient.

Still, to assess the potential quantitative relevance of a measurement error explanation, we conduct simulations that provide an estimate of the magnitude of measurement error that would be required to generate the observed variation in coefficient estimates across different levels of aggregation. Suppose that observed patience p_o is given by $p_o = p_t + \alpha \times \epsilon$, where p_t is the respondent's true patience, α a scaling parameter and $\epsilon \sim \mathcal{N}(0,1)$ a noise term (recall that observed patience is also normalized to have a mean of zero and a standard deviation of one). The simulations, described in Appendix C, show that $\alpha = 6$ is required to explain the observed variation in coefficients. To see that this is unreasonable, note that the test-retest correlation of preference parameters is estimated to be slightly below 0.6 (Beauchamp et al., 2011), yet $\alpha = 6$ would imply a test-retest correlation of only $\rho = 0.02$.[8] While there is reason to believe that the test-retest correlation in heterogeneous large-scale survey samples would be lower than with student subject pools, an implied test-retest correlation of 0.02 appears too low to be reasonable. We conclude from our analysis that some other deeper mechanism must be at play that generates the seeming aggregation effects.[9] To shed further light on this issue, we now turn to presenting and estimating a formal model that features individual- and country-level heterogeneity in patience.

4 A Conceptual Framework

4.1 Setup

We present a deliberately simple model that builds upon previous contributions on the role of patience for the accumulation of physical capital (Ramsey-Cass-Koopmans), human capital (Becker, 1962; Ben-Porath, 1967), and potentially human capital externalities on productivity (Lucas, 1988).[10] Consider an economy of overlapping generations of individuals that live for three periods. Each generation has unit mass and each period lasts for one unit of time. Individuals derive utility from consumption and are heterogeneous

[8] To generate a test-retest correlation close to 0.6, α would have to be approximately 0.75. However, with $\alpha = 0.75$, the coefficient of patience obtained at the country level would be only about twice as large as the individual-level coefficient, again at odds with the data.

[9] An additional measurement-related issue that could in principle generate differences in coefficient estimates between individual and country-level regressions is expansion bias resulting from a left-censoring of the patience variable. Indeed, in our data, about 56% of respondents always opt for the immediate payment in the quantitative staircase procedure, so that we can only identify an upper bound for their patience. Appendix C.2 discusses this issue in detail.

[10] See also Acemoglu (2008) for a comprehensive overview of the role of time preferences for growth and Doepke and Zilibotti (2014) for the role of patience in an education-based growth model.

with respect to their patience. When young, all individuals work as unskilled workers in production and decide whether to become educated, which is analogous to becoming a skilled worker in the second period. Becoming skilled requires young individuals to spend a fraction $(1-\psi)$ of their time on the acquisition of human capital. During the second period of life, individuals work either as unskilled or skilled workers, depending on their previous education choice, and make savings decisions. During the third period of life, individuals retire from the labor market and finance consumption from their savings. At the aggregate level, saved income is transformed one-to-one into physical capital that can be used for production during the following period. The capital accumulated by one generation during their second period of life fully depreciates at the end of the third period of life.

Let generations be indexed by the period during which they are young. The preferences of individual i are represented by

$$U(i) = \ln c_t + \beta(i) \ln c_{t+1} + (\beta(i))^2 \ln c_{t+2} \; , \tag{2}$$

where $\beta(i) \in (0,1)$ is the discount factor of individual i, which corresponds to this individual's level of patience. For analytical convenience, $\beta(i)$ is modeled as a draw from a uniform distribution $\beta \sim U[\chi-\varepsilon;\; \chi+\varepsilon]$, where $\chi > 0$ reflects the average level of patience in the population and where the density is $\frac{1}{2\varepsilon}$ (with $\varepsilon > 0$, $\chi > 0$ and $0 < \chi - \varepsilon < \chi + \varepsilon < 1$). In the analysis below, variation in $\beta(i)$ conditional on χ captures individual-level heterogeneity within an economy, whereas variation in χ reflects comparisons across model economies.

4.2 Optimal Individual Accumulation Decisions

Human Capital Acquisition. Becoming a skilled worker requires time for human capital acquisition corresponding to a fraction $(1-\psi)$ of the first period of life. We assume that the stock of human capital increases with the time spent on education according to a standard Mincerian specification, with the stock of human capital corresponding to $h = e^{\rho(1-\psi)}$, with $\rho > 0$ as parameter for the return. For analytical simplicity, we restrict attention to a binary education choice.

Budget Constraints. Denote the wage of unskilled workers by w^L, the earnings of a skilled worker as w^H, the savings rates of unskilled and skilled workers as s^L and s^H, and the return on capital as R. We assume that individuals cannot save or borrow when young.[11] The respective budget constraints are then

[11] This assumption ensures a role of patience for education choices by preventing consumption smoothing through savings, see, e.g., Doepke and Zilibotti (2014) for a similar setup.

$$\text{unskilled:} \quad c_t^y = w_t^L \ , \qquad c_{t+1}^m = w_{t+1}^L \cdot (1 - s_{t+1}^L) \ , \qquad c_{t+2}^o = w_{t+1}^L \cdot s_{t+1}^L \cdot R_{t+2} \ , \tag{3}$$

$$\text{skilled:} \quad c_t^y = w_t^L \psi \ , \quad c_{t+1}^m = w_{t+1}^H h \cdot (1 - s_{t+1}^H) \ , \quad c_{t+2}^o = w_{t+1}^H h \cdot s_{t+1}^H \cdot R_{t+2} \ . \tag{4}$$

Individuals take wages and capital returns as given.

Optimal Individual Decisions. The optimal savings decision in the second period of life for an unskilled worker i of generation t is determined by maximizing (2) subject to (3). Analogously, the optimal savings decision for individual i conditional on becoming a skilled worker is determined by maximizing (2) subject to (4). Solving the individual decision problems delivers the optimal savings rate as

$$s_{t+1}^L = s_{t+1}^H = \frac{\beta(i)}{1 + \beta(i)} \ , \tag{5}$$

which is strictly increasing in individual i's patience $\beta(i)$. Due to log utility, the savings rate does not depend on the return to capital nor on the education status of the individual.

The choice to become a skilled worker involves a comparison of (indirect) lifetime utilities. The condition for becoming skilled is determined by whether the return for becoming skilled, which is given by the wage ratio $\eta_{t+1} = \frac{w_{t+1}^H}{w_{t+1}^L}$, matches the compensation that an individual requires for being willing to spend a fraction $(1 - \psi)$ of the first period of life on acquiring human capital. After cancelling common terms (wages), substituting from the optimal savings decision and simplifying, the condition for a preference for becoming skilled given by

$$\eta_{t+1} > \eta(i) = \frac{\psi^{\frac{-1}{\beta(i)(1+\beta(i))}}}{h} \ , \tag{6}$$

with $\eta(i)$ denoting the minimum compensation that is required for making the individual with patience $\beta(i)$ indifferent between becoming skilled or remaining unskilled. This minimum compensation is decreasing in patience $\beta(i)$ since a higher $\beta(i)$ implies a greater utility weight on the earnings premium that is associated with becoming skilled, thus implying a lower requirement for market compensation. Intuitively, the earnings premium from becoming skilled accrues during the second period of life and, through savings, also benefits the individual during the third period of life. Hence, the market premium that compensates an individual for the opportunity cost of time foregone for education in the first period of life is smaller the more patient the individual. For a given wage ratio $\eta_{t+1} = \frac{w_{t+1}^H}{w_{t+1}^L}$, condition (6) therefore implicitly determines a threshold level for patience, $\tilde{\beta}_t$, that determines the population share of skilled individuals.

The model has straightforward predictions about how savings, education, and income respond to variation in patience at the individual level. Taking the aggregate allocation as given, a higher level of patience $\beta(i)$ is associated with a higher individual propensity to save as consequence of (5). Likewise, more patient individuals have a higher propensity to become skilled due to (6). As a result of these two mechanisms, lifetime income also increases in individual patience.

4.3 Aggregate Equilibrium

Production. The production of final output Y during period t combines the available stocks of physical capital, skilled labor and unskilled labor. In light of the empirical evidence regarding capital-skill complementarities (Duffy et al., 2004), we assume that the production function takes the form

$$Y_t = A_t \left[\left(K_t^\theta + H_t^\theta \right)^{\frac{\sigma}{\theta}} + L_t^\sigma \right]^{\frac{1}{\sigma}} \ , \tag{7}$$

with the aggregate capital stock in period t being denoted by K_t, the stock of unskilled labor denoted by L_t, the effective stock of skilled labor denoted by H_t and A_t denoting total factor productivity (TFP). Consistent with empirical estimates, we assume $\sigma > \theta > 0$. Markets for capital, unskilled workers and skilled workers are competitive and factors are paid their marginal products. Income can be used for consumption or capital accumulation; saved income is transformed one-to-one into physical capital. From the determination of wages, it follows that during the second period of their lives, skilled workers supply their human capital and enjoy an earnings premium $\eta_{t+1}h = \frac{w_{t+1}^H h}{w_{t+1}^L} = e^{\rho(1-\psi)} \cdot \left[K_{t+1}^\theta + H_{t+1}^\theta \right]^{\frac{\sigma-\theta}{\theta}} \frac{L_{t+1}^{1-\sigma}}{H_{t+1}^{1-\theta}}$.

Factor Market Clearing. In a given generation, only individuals with $\beta(i) > \tilde{\beta}_t$ optimally decide to become skilled. Since unskilled labor is supplied by workers of two adjacent generations (during the first period of life and those that remain unskilled during the second period of life), the stock of unskilled labor is given by

$$L_t = \frac{1}{2\varepsilon} \left[\int_{\chi-\varepsilon}^{\tilde{\beta}_{t-1}} 1 \, d\beta + \int_{\chi-\varepsilon}^{\tilde{\beta}_t} 1 \, d\beta + \int_{\tilde{\beta}_t}^{\chi+\varepsilon} \psi \, d\beta \right] \ , \tag{8}$$

where $\tilde{\beta}_{t-1}$ and $\tilde{\beta}_t$ correspond to the patience thresholds that determine the stock of skilled workers of generations $t-1$ and t, respectively. The stock of skilled workers in a given period is given by

$$H_t = \frac{1}{2\varepsilon} \int_{\tilde{\beta}_{t-1}}^{\chi+\varepsilon} e^{\rho(1-\psi)} \, d\beta \ . \tag{9}$$

Since individual savings differ across education groups as consequence of different labor incomes, the information about the population composition allows for the determination of aggregate capital accumulation, with capital supply given by

$$K_{t+1} = \frac{1}{2\varepsilon}\left[\int_{\chi-\varepsilon}^{\tilde{\beta}_{t-1}} \frac{\beta(i)}{1+\beta(i)}(w_t^L \cdot 1)\, d\beta(i) + \int_{\tilde{\beta}_{t-1}}^{\chi+\varepsilon} \frac{\beta(i)}{1+\beta(i)}(w_t^H \cdot h)\, d\beta(i)\right] . \tag{10}$$

Extension: Human Capital Externalities. In its most basic form, the model does not feature an effect of patience on factor productivity. In a model variant, we consider a simplified human capital externality of the stock of skilled workers on productivity (e.g., Lucas, 1988):

$$A_t = \bar{A} \cdot H_t^{\gamma} , \tag{11}$$

where $\gamma \geq 0$.

Equilibrium. The remaining analysis focuses on the steady state equilibrium. In steady state, wages and the share of skilled individuals are constant, such that $\eta_{t+1} = \eta$ and $\lambda_t = \lambda$. This follows from the one-to-one mapping between λ and $\tilde{\beta}$ and solving the condition for becoming unskilled vs. skilled (6) at the point of indifference, which determines the threshold level for patience as

$$\tilde{\beta} = \frac{1}{2}\left[-1 + \sqrt{1 - 4 \cdot \frac{\ln \psi}{\ln(\eta h)}}\right] .$$

Under the assumption that $\beta(i)$ is distributed uniformly, this mapping is $\lambda = \frac{\chi+\varepsilon-\tilde{\beta}}{2\varepsilon} \Leftrightarrow \tilde{\beta} = \chi + \varepsilon - 2\varepsilon\lambda$.[12]

The equilibrium is characterized by a skill share λ and the aggregate allocations of skilled and unskilled labor and capital, as well as the corresponding competitive prices such that all individual decisions are consistent with the prices and the aggregate allocation.[13] The key condition for equilibrium is the consistency of the indifference condition for education (6) with the earnings premium that emerges from the relative supply of skilled labor, and the corresponding capital supply and demand.

In terms of comparative statics, a key result for the subsequent analysis is that the equilibrium share of skilled individuals is unambiguously higher in a more patient population. In particular, comparing across equilibria, the following conditions hold regarding the effect of an increase in χ: $0 < \frac{d\tilde{\beta}}{d\chi} < 1$, and $\frac{d\lambda}{d\chi} = \frac{1}{2\varepsilon}\left(1 - \frac{d\tilde{\beta}}{d\chi}\right) > 0$.[14] These conditions imply that the threshold in terms of individual patience for becoming skilled is higher in a country with a higher average level of patience.

[12] The average level of patience of unskilled workers is then given by $\underline{\beta} = \chi - \varepsilon\lambda$. Equivalently, the average patience of skilled workers is $\bar{\beta} = \chi + \varepsilon(1-\lambda)$.

[13] See Appendix F.1 for a formal definition of the equilibrium and the corresponding proof of existence and uniqueness.

[14] See Appendix F.2 for the derivations.

5 Bringing the Model to the Data

5.1 Testable Implications

Approach. We will be interested primarily in evaluating the effect of an increase in individual or average patience on education, savings, and income, as well as how this effect varies with the level of aggregation. To construct model analogues for the empirical estimates of the elasticities of these variables with respect to patience, we analyze two thought experiments. At the individual level, the thought experiment is to determine the average elasticity with respect to patience $\beta(i)$ in the cross-section of individuals at a particular point in time, holding the aggregate allocation (reflected by the share skilled, λ, and the associated threshold for patience, $\tilde{\beta}$) fixed. This thought experiment matches the results of individual-level regressions with country fixed effects, where the country fixed effects absorb the aggregate allocation and prices.

At the aggregate level, the thought experiment assesses the consequences of a shift in average patience, χ, on the steady state equilibrium. Conceptually, this reflects the effect of variation in the patience distribution across economies that are identical otherwise. This shift leads to general equilibrium effects that need to be taken into consideration and quantified since the factor allocation and prices change. This thought experiment corresponds to the cross-country or cross-regions regressions above.

To fix ideas and for expositional clarity, we consider a country in which patience is distributed uniformly with mean $\mu = \chi$ and standard deviation $sd = \frac{2\varepsilon}{\sqrt{12}}$. For comparisons across steady states we consider a shift in average patience that corresponds to one standard deviation, i.e., we compare the benchmark allocation of country 1 to that in a second country with $\chi_2 = \chi_1 + sd > \chi_1$, all else equal.

Education. First, consider the effect of patience on an individual's decision to become skilled. It is clear from (6) that the propensity to become skilled can be expressed as a binary choice problem with the compensation that an individual requires at least for becoming skilled, $\eta(i)$, as the latent variable. If the market compensation, η^*, is greater than this minimum compensation, the individual becomes skilled. Since in reality, other unobserved and idiosyncratic factors might influence the education choice, the empirical analogue of the decision to become skilled is a linear probability model in which the decision to become skilled is determined by (6) with an additive, symmetrically distributed noise term, $skilled(i) = (\eta^* - \eta(\beta(i))) + \epsilon(i)$. Consequently the marginal effect of an increase in patience on the propensity to become skilled is given by[15]

$$\frac{\partial skilled(i)}{\partial \beta(i)} = \frac{1}{2\varepsilon} \cdot \left| \frac{d\,\beta(i)}{d\,\eta(i)} \right| .$$

[15]See Appendix G for details.

Notice that the effect of patience on the individual propensity to become skilled depends on the level of patience (and is larger for more patient individuals). The empirical estimate of the elasticity of individual education with respect to patience obtained with a probability model as estimated in the reduced form analysis corresponds to the average of these individual marginal effects. For the quantitative analysis, the expression will be evaluated at the threshold $\tilde{\beta}$.

At the aggregate level, the effect of a shift in the distribution of patience on the share of skilled individuals can be expressed as

$$\frac{d\lambda}{d\chi} = \frac{1}{2\varepsilon}\left(1 - \frac{d\tilde{\beta}}{d\chi}\right) > 0 \, .$$

Since human capital is given by $H = e^{\rho(1-\psi)} \cdot \lambda$, this expression also proportional to the (semi-)elasticity of aggregate human capital with respect to patience with $\frac{dH}{d\chi} = e^{\rho(1-\psi)} \cdot \frac{d\lambda}{d\chi}$.

A comparison of the size of the effects at the individual and at the aggregate level requires additional assumptions. First, since the individual effect increases with patience, the size of the effect depends on $\beta(i)$ at which the effect is evaluated; it is largest when evaluated at $\tilde{\beta}$. Since the patience threshold $\tilde{\beta}$ is higher in countries with a greater average patience (i.e., $\tilde{\beta}(\chi_2) > \tilde{\beta}(\chi_1)$) the individual effect is amplified in countries with greater average patience. In addition, as consequence of the capital-skill complementarity, greater average patience induces general equilibrium effects that affect the aggregate skill share. This implies that the model is capable of generating an amplification of the elasticity of education with respect to variation in patience on the aggregate level compared to the individual level under some parametric restrictions. We discuss these restrictions in Appendix G.

Savings and Capital. In the model, savings are a continuous variable, while in our individual-level data we only observe a binary indicator for whether a respondent saved. For this reason, the quantitative analysis below will not use the elasticity of the individual savings rate with respect to patience as an empirical moment to be matched.

With the individual savings rate given as in (5), the average marginal effect of an increase in patience on individual savings is given by

$$\frac{\partial \bar{S}}{\partial \beta(i)} = \frac{1}{1+\tilde{\beta}}\left[\frac{(1-\lambda)\cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda \cdot w^H h}{1+\chi+\varepsilon}\right] \, , \tag{12}$$

where $\bar{S}$ denotes average individual savings.[16] This implies that the average effect of patience on individual savings is given by the corresponding weighted average effect on individual savings rates, with population shares and respective labor earnings as weights. These weights are fixed when considering the perspective of individual regressions, but

[16]See Appendix G for details.

they vary when comparing across steady states.

At the aggregate level, savings are given by the sum of total savings of unskilled and skilled workers whose per capita savings differ due to the difference in average patience across both groups, with skilled workers saving a higher share of their (higher) income. Thus, when estimating the elasticity of average savings (or capital) with respect to variation in patience across economies, the corresponding differences in the allocation in terms of the share skilled, λ, and wages also imply variation in the corresponding weights of the savings expression. Concretely, the effect of patience on log capital per capita is given by

$$\frac{d\,K}{d\,\chi} = \underbrace{\frac{\partial \bar{S}}{\partial \beta(i)} \frac{\partial \beta(i)}{\partial \chi}}_{\text{individual effect}} + \underbrace{S^L \cdot \frac{d\,w^L}{d\,\chi} + S^H \cdot \frac{d\,w^H h}{d\,\chi} + \frac{\tilde{\beta}}{1+\tilde{\beta}}(w^H h - w^L) \cdot \frac{d\,\lambda}{d\,\chi}}_{\text{GE effects}} \,,$$

where S^L and S^H denote the weighted savings rates among the groups of unskilled and skilled individuals, respectively. The first term captures the average increase in aggregate savings that results from higher individual savings rates in a country with a more patient population. The other terms capture the variation in aggregate savings due to general equilibrium effects that affect earnings.

For an amplification of the effect of patience on the aggregate level it is therefore necessary that the general equilibrium effect is positive. Notice that in a country with greater average patience the share of skilled is unambiguously larger. This implies that the wage of unskilled workers will be larger. With a sufficiently large capital-skill complementarity (as consequence of $\sigma > \theta$), the decline in the wage of skilled workers and in the skill premium is small enough such that the general equilibrium effect is positive, giving rise to an amplification of the effect on the aggregate level.

Income. Finally, consider the effect of patience on income. Average individual income in the cross-section of individuals at a given point in time is given by the average of the per capita income of each of the three generations alive at this point in time. The marginal effect of variation in patience on individual income is then given by

$$\frac{\partial \bar{y}}{\partial \beta(i)} = \frac{\partial \overline{Pr}(\eta(i) < \eta^*)}{\partial \beta(i)} \left[w^H h - (2-\psi) w^L \right] + R \cdot \frac{\partial \bar{S}}{\partial \beta(i)} \,,$$

where for simplicity bars over variables denote population averages. As before with savings, this corresponds to a weighted average of the effects of patience on the propensity to become educated and to save, with weights given by the aggregate allocation in terms of skill composition λ and the corresponding prices.

Turning to the effect of patience on income per capita when considering cross-country variation, the resulting changes in the aggregate allocation imply variation in the corresponding weights of the income expression. Concretely,

$$\frac{dY}{d\chi} = \underbrace{\frac{d\lambda}{d\chi}\left[w^H h - (2-\psi)w^L\right] + R \cdot \frac{dK}{d\chi}}_{\text{direct effects}} + \underbrace{\left[2(1-\lambda)+\psi\right]\frac{dw^L}{d\chi} + \lambda\frac{dw^H h}{d\chi} + K \cdot \frac{dR}{d\chi}}_{\text{GE effects}} .$$

As before, the effect obtained from cross-country variation in patience is amplified compared to the effect from variation on the individual level if the general equilibrium effects are positive. Moreover, it becomes clear that even if the direct effects on education and savings are amplified on the aggregate level, this is not necessarily also the case for income if the general equilibrium effects are negative. Again, a sufficiently large capital-skill complementarity in production makes it more likely that the general equilibrium effects are positive.[17]

5.2 Estimation Approach and Parameter Calibration

We estimate the model through an indirect inference approach. In particular, we estimate the structural model parameters by matching the empirical patience elasticities of the variables of interest from the reduced form evidence (the procedure is described in detail in Appendix H). This approach provides a direct assessment of whether a quantitative version of our model is able to deliver comparable amplification effects in the patience coefficients as in the empirical results presented above.

The baseline model presented so far is based on seven parameters. Many of these parameters are standard, and we calibrate them using conventional estimates from the literature. First, we calibrate the CES elasticities σ and θ based on empirical estimates by Duffy et al. (2004).[18] Second, we set the time requirement for becoming a skilled worker in terms of the fraction of the first period of life, $(1 - \psi = 0.2)$, which is equivalent to five years with the length of a generation being 25 years. Third, we assume a Mincerian return of 10%, which is in line with empirical estimates (e.g., Belzil and Hansen, 2002; Acemoglu and Angrist, 2000).[19] This implies setting $\rho = 2.5$. Finally, in the model extension with a human capital externality, we need to calibrate γ. The literature has not settled on how large human capital externalities are in the data (e.g., Ciccone and Peri, 2006; Acemoglu and Autor, 2012; Thönnessen and Gundlach, 2013). We follow the evidence in Acemoglu and Angrist (2000) and calibrate γ to be consistent with an aggregate Mincerian return of 6.8%. Panel A of Table 9 summarizes the parameter calibrations.

[17] See Appendix G for details.

[18] In particular, we use the average of their estimates for high skilled workers defined as workers with completed secondary education or college attainment.

[19] Regressing household income on years of schooling in our global individual-level data delivers an average Mincerian return of approximately 6.5%. This is likely to be misspecified, however, since years of education are measured with considerable error, the regression involves household income rather than individual labor earnings, and since household income does not necessarily reflect the respondent's labor earnings.

The remaining three parameters are the distributional parameters of the patience distribution, χ and ϵ, and the baseline level of TFP, A. These free parameters are estimated by matching four empirical moments: the patience elasticities of education at the individual and at the aggregate level, and the patience elasticities of income at the individual and at the aggregate level.[20] To keep this analysis directly comparable to the reduced-form patterns, the simulated moments of the model correspond to shifts of individual patience by one standard deviation (as in the individual-level OLS regressions, in which patience was standardized into a z-score). Likewise, we consider a shift in average patience by one standard deviation, which again directly corresponds to the OLS point estimates. Estimation is based on a Wald-type minimization of the vector of quadratic differences of the standardized elasticities.[21]

5.3 Model Specifications

To be able to shed light on the mechanisms behind the observed amplification effects in the data, we estimate several model variants. In the baseline version, there is no human capital externality on TFP, and the two model economies only differ in their patience distribution. Thus, patience only affects the accumulation of human and physical capital. We think of this specification as conceptual analogue to the within-country across-regions regions reported in Section 3.2. Here, patience might affect the formation of physical and human capital, but the broader productivity environment is held constant. For instance, national institutions or policies plausibly affect the productivity environment or the supply of educational resources, yet these are fixed when comparing subnational regions.

Second, we estimate a variant with a human capital externality, where we calibrate $\gamma = 1.7$ based on Acemoglu and Angrist (2000). In this version, TFP implicitly also depends on patience. We think of this variant as analogue of the cross-country regressions, where the broader productivity environment also varies, and patience could implicitly affect the supply of national policies or other productivity shifters. Third, we estimate a variant in which we don't calibrate the externality parameter γ but instead estimate it as additional free parameter. Here, again, potential amplification effects can depend on patience-induced variation in TFP.

We view the aforementioned model variants as primary objects of interest. To benchmark the results on the role of patience for amplification effects, we perform a fourth estimation that resembles development accounting exercises. Here, we fix the patience parameters across the two model economies, and estimate two different TFP levels. That is, in these estimations, all differences in aggregate income are driven by differences in TFP. This analysis will serve as a benchmark to investigate whether exogenous variation

[20] As noted above, we do not match the elasticities for savings due to the conceptual discrepancy that arises since the empirical data only contain binary information on whether a household saved or not.

[21] See Appendix H for details on the estimation procedure.

Table 9: Calibrated and estimated parameters: Baseline Specification

(*A*) Calibrated Parameters		
Parameter	Value	Calibration Details
$1-\psi$	0.2	Fraction of young age required to become skilled (five additional years) (Caselli, 2017)
ρ	2.5^a	Corresponds to a (private) Mincerian return of 10% (Belzil and Hansen, 2002; Acemoglu and Angrist, 2000)
σ	0.62	CES (inverse): labor/capital compound (Duffy et al., 2004)
θ	0.38	CES (inverse): physical/human capital (Duffy et al., 2004)
γ	1.7^a	Corresponds to an external Mincerian return of 6.8% (Acemoglu and Angrist, 2000)

(*B*) Estimated Parameters				
	Baseline	Extensions		
		TFP-Calibrated	TFP-Estimated	Dev. Acc.
$\chi_1,\ \chi_2$	$0.15,\ [0.24]^b$	$0.15,\ [0.23]^b$	$0.2,\ [0.24]^b$	$[0.15,\ 0.15]^e$
ϵ	0.15	0.14	0.06	$[0.15]^e$
$\bar{A}$	0.05	0.12	0.89	
$A_1,\ A_2$		$[0.05,\ 0.1]^c$	$[0.9,\ 1.64]^c$	$0.03,\ 0.07$
γ		$[1.7]^d$	2.38	

Panel (*A*): Calibrated parameters. [a] With the Mincerian human capital production function, a return of $x = 0.1$ for five years of schooling during a 25-year period of youth corresponds to $\rho = \frac{\ln\left(e^{0.1\cdot 5}\right)}{0.2} = 2.5$; γ is calibrated analogously.

Panel (*B*): Estimated parameters. Parameters in brackets [ ] are derived from estimated parameters. [b] Level of χ_2 as implied by the estimated values of χ_1 and ε. [c] Effective levels of A when incorporating the human capital externality. [d] Value calibrated as described in Panel (*A*). [e] Values fixed as in baseline model.

in TFP alone is sufficient to rationalize the patterns in the data, in particular for whether exogenous technology differences could contribute to the observed amplification of patience elasticities.

5.4 Estimation Results

Table 9(A) presents an overview of the calibrated parameters and Table 9(B) shows the results of the estimation of the different model specifications. Throughout the different specifications, the estimation delivers reasonable parameter values for patience. In particular, noticing that the estimates of χ correspond to the country-average of a 25-year discount factor, the estimates are equivalent to an average annual average discount factor of 0.93 to 0.95 (or equivalently a discount rate of $5-7\%$).

Table 10 reproduces the reduced-form estimates of the elasticities of the various variables of interest with respect to patience in our data, and compares them with estimated model quantities. In particular, the first column replicates the empirical elasticities in the different dimensions obtained with reduced form estimates from the most preferred specifications reported in Section 3. These elasticities constitute the empirical targets for the model

Table 10: Quantified model vs. data

	Effect of one SD increase of patience				TFP variation
	Data	Model			
	(Baseline Controls)	Baseline	Extensions		Dev. Acc.
			TFP-Calibrated	TFP-Estimated	
	(1)	(2)	(3)	(4)	(5)
			Individual level		
Income	0.05	0.05	0.05	0.05	0.05
Education	0.03	0.03	0.03	0.03	0.03
			Country level		
Income	1.73	0.12	1.49	1.77	1.78
Fraction skilled	0.20	0.17	0.18	0.17	0.01

The effect sizes in the simulated model are obtained after estimating the parameters through indirect inference as reported in Table 9. In the baseline, estimated parameters are χ, ϵ, and A by matching as moments the effects of patience on individual income, individual propensity to become skilled, aggregate income and aggregate skill share. In the extensions, the parameter γ is either calibrated to the external Mincerian return or estimated by matching the same moments as in the baseline estimation.
The effect sizes from estimates at the individual level correspond to specifications with regional fixed effects as in column (3) of Table 8. The effect sizes from estimates at the country level correspond to the OLS coefficients with all controls from column (4) in Table 1, respectively. In the country-level analysis, the fraction of skilled workers is the share with completed secondary education or higher. In the corresponding individual-level analysis, the fraction of skilled workers refers to the probability of having at least secondary education.

estimation. The other columns present the corresponding estimates of the elasticities from estimating the model using the indirect inference approach.

We start by estimating the baseline version of the model without a human capital externality ($\gamma = 0$). This is the most restrictive version of the model in terms of explaining amplification effects. The results for this baseline specification – in which TFP is fixed across economies – are shown in column (2) of Table 10. The elasticities of income and education with respect to variation in patience at the individual level obtained with the estimated version of the model resemble the reduced form estimates, as shown in the upper panel. The bottom panel of the table shows the corresponding elasticities for variation in patience across countries. The baseline version of the model delivers only a moderate amplification of the elasticity in income and education by a factor of about two. Interestingly, this magnitude of amplification corresponds to the patterns observed in the reduced-form regressions across subnational regions. The fact that the observed amplification is much larger at the country level – and that this cannot be reproduced by our baseline specification – suggests that variation in TFP might correlate with patience differences across countries. Indeed, researchers have argued that many barriers to increasing educational quality are not primarily financial or technological but instead political in nature (Duflo, 2001; Acemoglu and Autor, 2012). Since these factors likely respond to national policies, there is reason to believe that regional levels of development may respond to national factors. To the extent that national policies respond to national patience, this would explain why the observed amplification is considerably smaller at the regional level.

To account for these national productivity factors in a parsimonious way, we estimate the model variant in which the human capital externality is exogenously parameterized. The estimates for the parameters resulting from this extension are similar to those for the baseline model (see Table 9). As documented by the estimation results in column (3) of Table 10, this model delivers patience elasticities of income and education at the individual level that closely resemble the empirical estimates. At the aggregate level, however, this model delivers a large amplification of the country-level elasticity, particularly for income. The estimated coefficient increases by a factor of 30 relative to its individual-level analogue. This suggests that productivity differences play an important role for the observed amplification effects.

To further explore this issue, we present a second TFP extension in which the externality parameter γ is not parameterized but estimated as additional free parameter. The estimates deliver a slightly larger elasticity than in the calibrated version (see Table 9). The results of the estimation are shown in column (4) of Table 10. Again, the individual-level patience elasticities are matched, while this extended version of the model also delivers a substantial amplification of the elasticities at the aggregate level that is very similar to the pattern observed in the reduced form estimates. In fact, the model is able to account for all of the amplification in the income elasticity and for most of the amplification in the education elasticity.

These results raise the immediate question about why TFP differences contribute so much to the observed amplification. A first possibility is that patience-induced variation in TFP is responsible. A second possibility is that even exogenous TFP differences (that are orthogonal to patience) can account for the observed amplification. To investigate this, we estimate a final specification in which we hold patience fixed across both model economies, yet we separately estimate TFP for each country (rich and poor). Column (5) of Table 10 presents the results. For the patience elasticity at the individual level, the patterns are similar as for the other version of the model. Regarding the country-level elasticities, this version of the model also matches the amplification for income. At the same time, this version of the model is unable to generate substantial effects for education, and fails to account for the amplification of the patience elasticity in this dimension.

Overall, the estimation results illustrate two themes. First, without a human capital externality, the model generates a substantial amplification effect in coefficient estimates

Table 11: Simulated Moments

	Data	Model			
	(Baseline Controls)	Baseline	Extensions		
			TFP-Calibrated	TFP-Estimated	Dev. Acc.
	(A) Elasticities: Effect of 1 SD increase of patience on				
Aggregate capital:	1.17	0.5	2.26	1.71	1.73
HH savings (% Income):	9.8	6.18	5.91	2.26	0
TFP	0.17	0	1.0	0.82	1.52
Years of Schooling	2.77	0.83	0.9	0.87	0.07
	(B) Moments and Comparative Differences				
Skill shares ($\lambda_1,\ \lambda_2$)	0.59, 0.82	0.38, 0.55	0.36, 0.54	0.61, 0.78	0.37, 0.39
Ratio skill shares ($\frac{\lambda_2}{\lambda_1}$)	1.39	1.43	1.5	1.28	1.04
Capital ratio ($\frac{K_2}{K_1}$)	3.23	1.5	3.26	2.71	2.73
Output ratio ($\frac{Y_2}{Y_1}$)	2.98	1.12	2.49	2.77	2.78

Panel (A): Elasticities of variables with respect to patience that are not targeted in the estimation. The effect sizes in the simulated model are obtained after estimating the parameters through indirect inference as reported in Table 9.
[a] In the theoretical model, the propensity to save is a continuous variable; due to the assumption of log utility, all individuals save a strictly positive amount, so that only the intensive margin of savings is relevant. The data only contain information about savings propensities at the extensive margin, i.e., whether a household saved or not. Individual and aggregate savings are therefore not directly comparable and hence not used as target moments for the estimation.

Panel (B): Moments that are not targeted in the estimation. Empirical moments for the baseline economy 1 correspond to the respective averages among countries with patience close to the country mean (within 25th and 75th percentile of the distribution of country means of patience); empirical moments for economy 2 correspond to the respective averages among countries with patience around one standard deviation above the country mean (within 75th and 95th percentile of the distribution of country means of patience). Skill shares correspond to the population share aged 25+ with completed secondary education.

going from individual to the aggregate level. This suggests that general equilibrium effects play a significant role for amplification. At the same time, this amplification is considerably smaller than what we observe in the data. Second, once human capital is allowed to exert an externality on TFP, the estimated coefficients get considerably closer to what we observe in the data. These effects primarily appear when TFP varies endogenously as a function of patience rather than when it varies exogenously. This suggests that productivity differences (such as national policies or supply of schooling) that are endogenous to average patience contribute to the observed amplification patterns.[22]

To further asses the model performance and the plausibility of the estimated structural parameters, Table 11 provides a comparison of other, non-targeted data moments and the corresponding moments obtained from the model estimation. In particular, column (1) shows moments obtained from reduced form estimates or raw data. Panel (A) shows

[22] In this respect, the results relate to the literature on aggregation and aggregation bias that has focused on heterogeneity of tastes and non-linearities in shocks (Blundell and Stoker, 2005) and that has pointed to potential biases in coefficient estimates due to the neglect of variation in aggregate conditions (Hanushek et al., 1996).

elasticities of capital accumulation with respect to patience. These moments have not been targeted in the estimation and thus allow for an assessment of the fit of the different specifications of the model. In addition, the table presents patience elasticities of TFP and years of schooling. Panel (B) presents other moments that are relevant from the perspective of comparative development.

By and large, the different versions of the model reveal similar patterns as for the targeted moments. In particular, all versions of the model deliver a reasonable fit of the patience elasticities, yet the model extensions match the country-level elasticities considerably better. A similar comment applies to the estimates of other moments, as shown by the results in Panel (B). In particular, we compare moments of countries with a mean patience that is close to the average of all countries to moments of countries with a mean patience that is substantially (approximately one standard deviation) above the average. In the data, this corresponds to a comparison of the country averages of countries within the 25th and 75th percentile of the cross-country distribution in patience, and country averages for countries within the 75th to 95th percentile. The comparison involves skill shares in terms of the fraction of the population aged 25 and older with at least completed secondary education, and the ratios of skill shares, capital endowments per worker, and output per capita, of the two countries. By and large, the moments implied by the estimates of the different specifications of the model closely resemble the empirical moments. In comparison, the model extensions with a human capital externality on TFP again provide the best overall fit.

Discussion. Overall, the results from the estimation of the model document that significant aggregation effects might arise purely through general equilibrium effects. At the same time, allowing for additional externalities of patience on TFP that work through accumulated factors like human capital imply a substantial increase in the amplification. These externalities reflect that more efficient institutions or technologies complement factor accumulation. For example, if patient populations opted for institutions designed to foster long-term growth as opposed to short-term rent extraction, these institutions may imply additional positive effects on factor accumulation and income. This argument complements evidence that productivity differences might have indirect effects through their influence on factor accumulation (Hsieh and Klenow, 2010; Manuelli and Seshadri, 2014).

The results also speak to the question whether and to which extent heterogeneity in patience in an otherwise neoclassical model can account for observed comparative development patterns. The conventional way to account for development differences is to investigate to which extent external factors that are reflected in TFP are required to account for income differences. As documented in the literature, the neoclassical model typically requires large TFP differences between countries to account for these differences (see, e.g., Hall and Jones, 1999; Bils and Klenow, 2000; Caselli, 2005). Several recent papers have argued for TFP differences interfering with quality-adjusted human capital accumulation or early childhood investments in education, showing that this reduces the variation in unexplained TFP that is required to explain the income gap.[23] The estimation results of the different model extensions provide an alternative perspective on this issue. In particular, the comparison of the our last model extension corresponds to the logic of conventional development accounting by keeping the distribution of patience fixed and instead allowing only TFP to vary across economies. While the estimation results suggest that the specification with only TFP differences is able to match the individual elasticities for education and income, it delivers too little variation in aggregate skill shares compared to the data ($\frac{\lambda_2}{\lambda_1} = 1.04$ compared to $\frac{\lambda_2}{\lambda_1} = 1.39$ in the data). Similarly, this model is unable to match the patience elasticity of education at the aggregate level (0.01 compared to an empirical estimate of 0.2).

On the other hand, allowing for variation in patience across economies improves the fit of the model in all dimensions, as documented by the results from the model extensions that account for human capital externalities on TFP. Heterogeneity in time preferences that maps into the accumulation of factors, which in turn affect productivity differences through externalities, contributes to accounting for the observed patterns in the data without the need for exogenous variation in TFP. In addition, the implied differences in TFP are more than 20% smaller in the version of the model with heterogeneity in patience than in the version that resembles the usual development accounting approach.

6 Concluding Remarks

Time preference is attracting increased attention in microeconomic development studies that employ RCTs. A recurring theme in this literature is that individuals may lack self-control and patience and hence not take up profitable fertilizers (Duflo et al., 2011), fail to save (Ashraf et al., 2006), procrastinate at work (Kaur et al., 2015), or engage in excessive alcohol consumption (Schilbach, 2015). In addition, several papers have shown how differences in patience account for behavioral differences (for a summary, see Table A1 in Falk et al. (2020). For example, among children and adolescents, impatience is associated with a higher likelihood of drinking alcohol and smoking, a higher body mass index, a lower propensity to save, worse grades, more disciplinary conduct violations at school and a lower likelihood to complete high school in time (Castillo et al., 2011, 2019). While these micro studies point to a nexus between patience and individual outcomes, little has been known about the broader macro relationship between development and

[23]For instance, Hsieh and Klenow (2010) argue that TFP differences are amplified through their influence on the accumulation of factors. Erosa et al. (2010) and Manuelli and Seshadri (2014) find that accounting for human capital accumulation differences in human capital substantially amplifies TFP differences across countries. Schoellman (2012) makes a related point based on a novel methodology designed to measure differences in human capital quality.

heterogeneity in time preference.

In this paper, we have documented two key patterns. First, across levels of aggregation, patience is systematically linked to the accumulation of human capital, physical capital, and the stock of knowledge. Second, the data reveal strong aggregation effects with respect to patience. The analysis of a stylized equilibrium model that allows for heterogeneity in patience within and across countries has shown that both patterns are consistent with standard micro- and macroeconomic theories of intertemporal choice. A quantitative analysis based on a structural estimation of our model suggests that the difference in magnitude of coefficients across levels of aggregation can be rationalized by general equilibrium effects in the model. This suggests that heterogeneity in time preference potentially constitutes a relevant factor behind comparative development differences.

Our paper has only provided a first step towards understanding the relationship between patience and development, in particular given that our analyses are correlational in nature. Ultimately, we cannot (and do not intend to) confidently rule out that heterogeneity in patience reflects environmental conditions such as institutional quality or education. At the same time, even if a variable such as institutional quality was the ultimate driver of the results in this paper, the mechanism would likely partly operate through patience.

Still, an important question concerns the ultimate origins of variation in patience. Among the few candidate determinants that have been proposed are religion (Weber, 1930), cultural legacy as manifested in very old linguistic features (Chen, 2013), historical agricultural productivity and crop yield (Galor and Özak, 2014), mortality (Falk et al., 2019), as well as migratory movements of our very early ancestors (Becker et al., 2020). Future research might be able to disentangle the causal mechanisms that are at play here, perhaps along the lines of theoretical contributions that emphasize the two-way links between patience and education or income (Becker and Mulligan, 1997; Doepke and Zilibotti, 2008).

References

Acemoglu, Daron, *Introduction to Modern Economic Growth*, Princeton University Press, 2008.

_ **and David Autor**, "What does human capital do? A review of Goldin and Katz's The race between education and technology," *Journal of Economic Literature*, 2012, *50* (2), 426–63.

_ **and Joshua Angrist**, "How large are human-capital externalities? Evidence from compulsory schooling laws," *NBER Macroeconomics Annual*, 2000, *15*, 9–59.

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–351.

Algan, Yann and Pierre Cahuc, "Inherited Trust and Growth," *American Economic Review*, 2010, pp. 2060–2092.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines," *The Quarterly Journal of Economics*, 2006, *121* (2), 635–672.

Ashraf, Quamrul and Oded Galor, "The Out of Africa Hypothesis, Human Genetic Diversity and Comparative Economic Development," *American Economic Review*, 2013, *103* (1), 1–46.

Barro, Robert J. and Jong-Wha Lee, "A New Data Set of Educational Attainment in the World, 1950–2010," *Journal of Development Economics*, 2012.

Bauer, Michal, Julie Chytilová, and Edward Miguel, "Using Survey Questions to Measure Preferences: Lessons from an Experimental Validation in Kenya," *European Economic Review*, 2020, p. 103493.

Beauchamp, Jonathan, David Cesarini, and Magnus Johannesson, "The Psychometric Properties of Measures of Economic Risk Preferences," *Working Paper*, 2011.

Becker, Anke, Benjamin Enke, and Armin Falk, "Socio-economic status and inequalities in childrens IQ and economic preferences," *AEA Papers and Proceedings*, 2020, *110* (May 2020), 319–323.

Becker, Gary S., "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 1962, *70* (5), 9–49.

_ **and Casey B. Mulligan**, "The Endogenous Determination of Time Preference," *Quarterly Journal of Economics*, 1997, pp. 729–758.

Belzil, Christian and Jörgen Hansen, "Unobserved Ability and the Return to Schooling," *Econometrica*, 2002, *70* (5), 2075–2091.

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

Bick, Alexander, Nicola Fuchs-Schündeln, and David Lagakos, "How do Hours Worked Vary with Income? Cross-Country Evidence and Implications," *American Economic Review*, 2018, *108* (1), 170–99.

Bils, Mark and Peter J. Klenow, "Does Schooling Cause Growth?," *American Economic Review*, 2000, *90* (5), 1160–1183.

Blundell, Richard and Thomas M. Stoker, "Heterogeneity and Aggregation," *Journal of Economic Literature*, 2005, *43* (3), 347–391.

Caselli, Francesco, "Accounting for Cross-Country Income Differences," *Handbook of Economic Growth*, 2005, *1*, 679–741.

_ , *Technology Differences Over Space and Time*, Princeton, N.J.: Princeton University Press, 2017.

_ **and Antonio Ciccone**, "The Contribution of Schooling in Development Accounting: Results from a Nonparametric Upper Bound," *Journal of Development Economics*, 2013, *104*, 199–211.

Castillo, Marco, Paul J. Ferraro, Jeffrey L. Jordan, and Ragan Petrie, "The today and tomorrow of kids: Time preferences and educational outcomes of children," *Journal of Public Economics*, 2011, *95* (11), 1377–1385.

_ **,** _ **,** _ **, and** _ , "Discount Rates of Children and High School Graduation," *Economic Journal*, 2019, *129* (619), 1153–1181.

Chen, M. Keith, "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors and Retirement Assets," *American Economic Review*, 2013, *103* (2), 690–731.

Chetty, Raj and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates," *Working Paper*, 2016.

Ciccone, Antonio and Giovanni Peri, "Identifying human-capital externalities: Theory with applications," *The Review of Economic Studies*, 2006, *73* (2), 381–412.

Doepke, Matthias and Fabrizio Zilibotti, "Occupational Choice and the Spirit of Capitalism," *Quarterly Journal of Economics*, 2008, *123* (2), 747–793.

— **and —**, "Culture, Entrepreneurship and Growth," in Philippe Aghion and Steven Durlauf, eds., *Handbbook of Economic Growth*, Vol. 2A, Elsevier, Amsterdam, 2014.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "Are risk aversion and impatience related to cognitive ability?," *American Economic Review*, 2010, *100* (3), 1238–60.

Duffy, John, Papageorgiou, and Fidel Perez-Sebastian, "Capital-Skill Complementarity? Evidence from a Panel of Countries," *Review of Economics and Statistics*, 2004, *86* (1), 327–344.

Duflo, Esther, "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment," *American economic review*, 2001, *91* (4), 795–813.

—, Michael Kremer, and Jonathan Robinson, "Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya," *American Economic Review*, 2011, *101* (6), 2350–90.

Erosa, Andres, Tatyana Koreshkova, and Diego Restuccia, "How Important Is Human Capital? A Quantitative Theory Assessment of World Income Inequality," *Review of Economic Studies*, 2010, *77*, 1421–1449.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global evidence on economic preferences," *The Quarterly Journal of Economics*, 2018, *133* (4), 1645–1692.

—, —, —, David Huffman, and Uwe Sunde, "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time and Social Preferences," *IZA Discussion Paper*, 2016, *9674.*

—, Fabian Kosse, Pia Pinger, Hanna Schildberg-Hörisch, and Thomas Deckers, "Socio-economic status and inequalities in childrens IQ and economic preferences," *Journal of Political Economy, forthcoming*, 2020.

—, Johannes Hermle, and Uwe Sunde, "Longevity and Patience," *CRC TRR190 Discussion Paper*, 2019, *201.*

Galor, Oded and Ömer Özak, "The Agricultural Origins of Time Preference," *Working Paper*, 2014.

— and —, "The Agricultural Origins of Time Preference," *American Economic Review*, 2016, *106* (10), 3064–3103.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Human Capital and Regional Development," *Quarterly Journal of Economics*, 2013, *105*, 164.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Cultural Biases in Economic Exchange?," *Quarterly Journal of Economics*, 2009, *124* (3), 1095–1131.

Hall, Robert E. and Charles Jones, "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 1999, *114* (1), 83–116.

Hanushek, Eric A and Ludger Woessmann, "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes and Causation," *Journal of Economic Growth*, 2012, *17* (4), 267–321.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor, "Aggregation and the Estimated Effects of School Resources," *Review of Economics and Statistics*, 1996, *78* (4), 611–627.

Hsieh, Chang-Tai and Peter J. Klenow, "Development Accounting," *American Economic Journal: Macroeconomics*, 2010, *2* (1), 207–223.

Huffman, David, Raimond Maurer, and Olivia S Mitchell, "Time discounting and economic decision-making in the older population," *The Journal of the Economics of Ageing*, 2017, *14*, 100121.

Jones, Charles I and Peter J Klenow, "Beyond GDP? Welfare across countries and time," *American Economic Review*, 2016, *106* (9), 2426–57.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Self-control at work," *Journal of Political Economy*, 2015, *123* (6), 1227–1277.

Knack, Stephen and Philip Keefer, "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *Quarterly Journal of Economics*, 1997, pp. 1251–1288.

Lucas, Robert E., "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 1988, *22* (1), 3–42.

Manuelli, Rodolfo E. and Ananth Seshadri, "Human Capital and the Wealth of Nations," *American Economic Review*, 2014, *104* (9), 2736–2762.

Olsson, Ola and Douglas A. Hibbs Jr, "Biogeography and Long-Run Economic Development," *European Economic Review*, 2005, *49* (4), 909–938.

Romer, Paul M., "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), S71–S102.

Schilbach, Frank, "Alcohol and self-control: A field experiment in India," *Unpublished manuscript*, 2015.

Schoellman, Todd, "Education Quality and Development Accounting," *Review of Economic Studies*, 2012, *79*, 388–417.

Spolaore, Enrico and Romain Wacziarg, "The Diffusion of Development," *Quarterly Journal of Economics*, 2009, *124* (2), 469–529.

Squicciarini, Mara P and Nico Voigtländer, "Human capital and Industrialization: Evidence from the Age of Enlightenment," *Quarterly Journal of Economics*, 2015, *130* (4), 1825–1883.

Tabellini, Guido, "Culture and Institutions: Economic Development in the Regions of Europe," *Journal of the European Economic Association*, 2010, *8* (4), 677–716.

Thönnessen, Rasmus and Erich Gundlach, "The size of human capital externalities: cross-country evidence," *Public Choice*, 2013, *157* (3/4), 671–689.

Weber, Max, *The Protestant Ethic and the "Spirit" of Capitalism*, New York: Charles Scribner's Sons., 1930.

APPENDIX

A Details on Data Collection and Patience Measure

The description of the dataset builds on Falk et al. (2018).

A.1 Overview

The cross-country dataset including risk aversion, patience, positive and negative reciprocity, altruism, and trust, was collected through the professional infrastructure of the Gallup World Poll 2012. The data collection process essentially consisted of three steps. First, we conducted an experimental validation procedure to select the survey items. Second, Gallup conducted a pre-test in a variety of countries to ensure the implementability of our items in a culturally diverse sample. Third, the final data set was collected through the regular professional framework of the World Poll 2012.

A.2 Experimental Validation

To ensure the behavioral relevance of our preference measures, all underlying survey items were selected through an experimental validation procedure. To this end, a sample of 409 German undergraduates completed standard state-of-the-art financially incentivized laboratory experiments designed to measure risk aversion, patience, positive and negative reciprocity, altruism, and trust. The same sample of subjects then completed a large battery of potential survey items. In a final step, for each preference, those survey items were selected which jointly performed best in predicting the behavior under real incentives measured in choice experiments. See Falk et al. (2018) for details.

A.3 Pre-Test

Prior to including the preference module in the Gallup World Poll 2012, it was tested in the field as part of the World Poll 2012 pre-test, which was conducted at the end of 2011 in 22 countries. The main goal of the pre-test was to receive feedback and comments on each item from various cultural backgrounds in order to assess potential difficulties in understanding and differences in the respondents' interpretation of items. Based on respondents' feedback and suggestions, minor modifications were made to the wordings of some items before running the survey as part of the World Poll (2012).

The pre-test was run in 10 countries in central Asia (Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Russia, Tajikistan, Turkmenistan, and Uzbekistan), 2 countries in South-East Asia (Bangladesh and Cambodia), 5 countries in Southern and Eastern Europe (Croatia, Hungary, Poland, Romania, Turkey), 4 countries in the Middle East and North Africa (Algeria, Lebanon, Jordan, and Saudi-Arabia), and 1 country in Eastern Africa (Kenya). In each country, the sample size was 10 to 15 people. Overall, more than 220 interviews were conducted. In most countries, the sample was mixed in terms of gender, age, educational background, and area of residence (urban / rural).

Participants in the pre-test were asked to state any difficulties in understanding the items and to rephrase the meaning of items in their own words. If they encountered difficulties in understanding or interpreting items, respondents were asked to make suggestions on how to modify the wording of the item in order to attain the desired meaning.

Overall, the understanding of both the qualitative items and the quantitative items was good. In particular, no interviewer received any complaints regarding difficulties in assessing the quantitative questions. When asked for rephrasing the qualitative patience item in their own words, most participants seemed to have understood the item in exactly the way that was intended.

However, when being confronted with hypothetical choices between monetary amounts today versus larger amounts one year later, some participants, especially in countries with current or relatively recent phases of volatile and high inflation rates, stated that their answer would depend on the rate of inflation, or said that they would always take the immediate payment due to uncertainty with respect to future inflation. Therefore, we decided to adjust the wording, relative to the "original" experimentally validated item, by adding the phrase "Please assume there is no inflation, i.e., future prices are the same as today's prices" to each question involving hypothetical choices between immediate and future monetary amounts.

A.4 Selection of Countries

Our goal when selecting countries was to ensure representativeness for the global population. Thus, we chose countries from each continent and each region within continents. In addition, we aimed at maximizing variation with respect to observables, such as GDP per capita, language, historical and political characteristics, or geographical location and climatic conditions. Accordingly, we favored non-neighboring and culturally dissimilar countries. This procedure resulted in the following sample of 76 countries:

East Asia and Pacific: Australia, Cambodia, China, Indonesia, Japan, Philippines, South Korea, Thailand, Vietnam.

Europe and Central Asia: Austria, Bosnia and Herzegovina, Croatia, Czech Republic, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Italy, Kazakhstan, Lithuania, Moldova, Netherlands, Poland, Portugal, Romania, Russia, Serbia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom.

Latin America and Caribbean: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Guatemala, Haiti, Mexico, Nicaragua, Peru, Suriname, Venezuela.

Middle East and North Africa: Algeria, Egypt, Iran, Iraq, Israel, Jordan, Morocco, Saudi Arabia, United Arab Emirates.

North America: United States, Canada.

South Asia: Afghanistan, Bangladesh, India, Pakistan, Sri Lanka.

Sub-Saharan Africa: Botswana, Cameroon, Ghana, Kenya, Malawi, Nigeria, Rwanda, South Africa, Tanzania, Uganda, Zimbabwe.

A.5 Sampling and Survey Implementation

A.5.1 Background

Since 2005, the international polling company Gallup has conducted an annual World Poll, in which it surveys representative population samples in almost every country around the world on, e.g., economic, social, political, and environmental issues. The collection of our preference data was embedded into the regular World Poll (2012) and hence made use of the pre-existing polling infrastructure of one of the largest professional polling institutes in the world.[24]

A.5.2 Survey Mode

Interviews were conducted via telephone and face-to-face. Gallup uses telephone surveys in countries where there is telephone coverage of at least 80% of the population or where this is the customary survey methodology. In countries where telephone interviewing is employed, Gallup uses a random-digit-dial method or a nationally representative list of phone numbers. In countries where face-to-face interviews are conducted, households are randomly selected in an area-frame-design.

A.5.3 Sample Composition

In most countries, samples are nationally representative of the resident population aged 15 and older. Gallup's sampling process is as follows.

Selecting Primary Sampling Units In countries where face-to-face interviews are conducted, the first stage of sampling is the identification of primary sampling units (PSUs), consisting of clusters of households. PSUs are stratified by population size and / or geography and clustering is achieved through one or more stages of sampling. Where

[24] Compare
`http://www.gallup.com/strategicconsulting/156923/worldwide-research-methodology.aspx`

population information is available, sample selection is based on probabilities proportional to population size. If population information is not available, Gallup uses simple random sampling.

In countries where telephone interviews are conducted, Gallup uses a random-digit-dialing method or a nationally representative list of phone numbers. In countries where mobile phone penetration is high, Gallup uses a dual sampling frame.

Selecting Households and Respondents Gallup uses random route procedures to select sampled households. Unless an outright refusal to participate occurs, interviewers make up to three attempts to survey the sampled household. To increase the probability of contact and completion, interviewers make attempts at different times of the day, and when possible, on different days. If the interviewer cannot obtain an interview at the initially sampled household, he or she uses a simple substitution method.

In face-to-face and telephone methodologies, random respondent selection is achieved by using either the latest birthday or else the Kish grid method.[25] In a few Middle East and Asian countries, gender-matched interviewing is required, and probability sampling with quotas is implemented during the final stage of selection. Gallup implements quality control procedures to validate the selection of correct samples and that the correct person is randomly selected in each household.

Sampling Weights Ex post, data weighting is used to ensure a nationally representative sample for each country and is intended to be used for calculations within a country. First, base sampling weights are constructed to account for geographic oversamples, household size, and other selection probabilities. Second, post-stratification weights are constructed. Population statistics are used to weight the data by gender, age, and, where reliable data are available, education or socioeconomic status.

A.5.4 Translation of Items

The preference module items were translated into the major languages of each target country. The translation process involved three steps. As a first step, a translator suggested an English, Spanish or French version of a German item, depending on the region. A second translator, being proficient in both the target language and in English, French, or Spanish, then translated the item into the target language. Finally, a third translator

[25] The latest birthday method means that the person living in the household whose birthday among all persons in the household was the most recent (and who is older than 15) is selected for interviewing. With the Kish grid method, the interviewer selects the participants within a household by using a table of random numbers. The interviewer will determine which random number to use by looking at, e.g., how many households he or she has contacted so far (e.g., household no. 8) and how many people live in the household (e.g., 3 people, aged 17, 34, and 36). For instance, if the corresponding number in the table is 7, he or she will interview the person aged 17.

would review the item in the target language and translate it back into the original language. If semantic differences between the original item and the back-translated item occurred, the process was adjusted and repeated until all translators agreed on a final version.

A.5.5 Adjustment of Monetary Amounts in Quantitative Items

All items involving monetary amounts were adjusted to each country in terms of their real value, i.e., all monetary amounts were calculated to represent the same share of the country's median income in local currency as the share of the amount in Euro of the German median income since the validation study had been conducted in Germany. Monetary amounts used in the validation study with the German sample were round numbers in order to facilitate easy calculations and to allow for easy comparisons (e.g., 100 Euro today versus $107,50$ in 12 months). In order to proceed in a similar way in all countries, we rounded all monetary amounts to the next "round" number. While this necessarily resulted in some (very minor) variation in the real stake size between countries, it minimized cross-country differences in understanding the quantitative items due to difficulties in assessing the involved monetary amounts.

A.5.6 Staircase procedure

The sequence of survey questions that form the basis for the quantitative patience measure is given by the "tree" logic depicted in Figure 3 for the benchmark of the German questionnaire. Each respondent faced five interdependent choices between receiving 100 euros today or varying amounts of money in 12 months. The values in the tree denote the amounts of money to be received in 12 months. The rightmost level of the tree (5th decision) contains 16 distinct monetary amounts, so that responses can be classified into 32 categories which are ordered in the sense that the (visually) lowest path / endpoint indicates the highest level of patience. As in the experimental validation procedure in Falk et al. (2016), we assign values 1 - 32 to these endpoints, with 32 denoting the highest level of patience.

A.6 Computation of Preference Measures

A.6.1 Cleaning and Imputation of Missings

In order to make maximal use of the available information in our data, missing survey items were imputed based on the following procedure:

If one survey item was missing, then the missing item was predicted using the responses to the other item. The procedure was as follows:

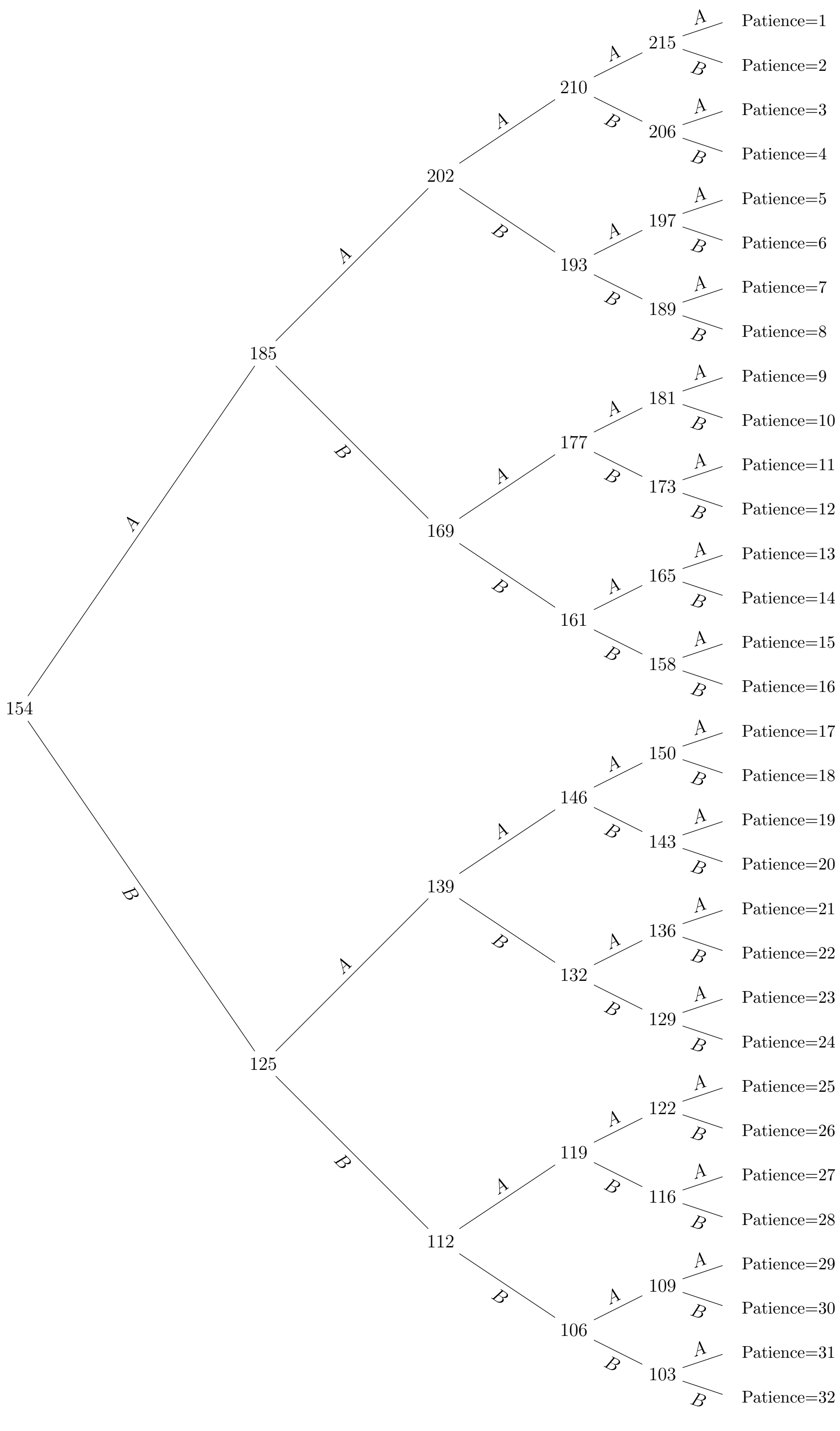


Figure 3: Tree for the staircase time task as implemented in Germany (numbers = payment in 12 months, A = choice of "100 euros today", B = choice of "x euros in 12 months"). First, each respondent was asked whether they would prefer to receive 100 euros today or 154 euros in 12 months from now (leftmost decision node). In case the respondent opted for the payment today ("A"), in the second question the payment in 12 months was adjusted upwards to 185 euros. If, on the other hand, the respondent chose the payment in 12 months, the corresponding payment was adjusted down to 125 euros. Working further through the tree follows the same logic.

- Qualitative question missing: We regress all available survey responses to the qualitative question on responses to the staircase task, and then use these coefficients to predict the missing qualitative items using the available staircase items.

- Staircase item missing: The imputation procedure was similar, but made additional use of the informational content of the responses of participants who started but did not finish the sequence of the five questions. If the respondent did not even start the staircase procedure, then imputation was done by predicting the staircase measure based on answers to the qualitative survey measure using the methodology described above. On the other hand, if the respondent answered at least one of the staircase questions, the final staircase outcome was based on the predicted path through the staircase procedure. Suppose the respondent answered four items such that his final staircase outcome would have to be either x or y. We then predict the expected choice between x and y based on a probit of the "x vs. y" decision on the qualitative item. If the respondent answered three (or less) questions, the same procedure was applied, the only difference being that in this case the obtained predicted probabilities were applied to the expected values of the staircase outcome conditional on reaching the respective node. Put differently, the procedure outlined above was applied recursively by working backwards through the "tree" logic of the staircase procedure.

In total, for about 8% of all respondents, one of the two patience measures was imputed.

A.6.2 Computation of Preference Indices at Individual Level

We compute an individual-level index of patience by (i) computing the z-scores of each survey item at the individual level and (ii) weighing these z-scores using the weights resulting from the experimental validation procedure of Falk et al. (2018). Formally, these weights are given by the coefficients of an OLS regression of observed behavior on responses to the respective survey items, such that the coefficients sum to one. These weights are given by (see above for the precise survey items):

$$\text{Patience} = 0.7115185 \times \text{ Staircase measure } + 0.2884815 \times \text{ Qualitative item } .$$

A.6.3 Computation of Country Averages

In order to compute country-level averages, we weigh the individual-level data with the sampling weights provided by Gallup, see above.

B Details for Regional-Level Analysis

Our regional-level data contain 704 regions (typically states or provinces) from the following countries: Argentina (16), Australia (8), Austria (9), Bolivia (8), Brazil (24), Cambodia (14), Cameroon (10), Canada (10), Chile (12), China (23), Colombia (23), Czech Republic (7), Egypt (3), Germany (16), Finland (4), France (22), Georgia (10), Ghana (10), Great Britain (12), Greece (13), Hungary (7), India (24), Indonesia (17), Iran (30), Israel (6), Italy (17), Jordan (6), Kazakhstan (6), Kenya (8), Lithuania (10), Macedonia (3), Malawi (3), Mexico (28), Morocco (13), Nigeria (6), Nicaragua (17), Netherlands (12), Pakistan (4), Poland (16), Portugal (7), Romania (8), Russia (27), Serbia (2), Spain (19), Sri Lanka (9), Sweden (8), Tanzania (20), Thailand (5), Turkey (4), Uganda (4), Ukraine (27), United Arab Emirates (7), USA (51), South Africa (9), Zimbabwe (10)

C Aggregation Effects

The main text has shown that the coefficient on patience becomes successively larger as one moves to higher levels of aggregation. This Appendix discusses two mechanical reasons that might drive this pattern, i.e., measurement error and resulting attenuation bias and censoring of the patience variable.

C.1 Measurement Error

If individual-level patience is measured with noise, then our country-level average patience measure will be a less noisy estimate of true country-level patience than our individual-level patience estimate. This difference in measurement error would lead to stronger attenuation at subnational levels and hence generate differences in coefficient magnitudes across levels of aggregation.

While we do not question that our data are affected by measurement error, this section investigates how large this measurement error would have to be to generate the observed differences in coefficients. To this end, we generate a synthetic patience measure for which we know the true relationship between income and patience at all levels of aggregation. We then subject this synthetic measure to noise and investigate how much noise we need to inflict on the synthetic measure to generate differences in coefficient sizes across aggregation levels that mirror those observed in our data. To this end, we focus on the comparison between (i) an OLS regression of log household income on patience, conditional on country fixed effects, and (ii) an OLS regression of log GDP p/c on average patience at the country level. Specifically, we conduct the following exercise:

1. To gauge the magnitude of amplification between individual-level and country-level regressions, we develop the following benchmark:
    - Regress household income on patience and country fixed effects.
    - Compute average household income (as proxy for GDP p/c) and average patience by country. Regress average household income on average patience.
    - Compute the ratio of patience coefficients in these two regressions. In our data, this ratio equals 39.
2. Generate a synthetic "true" patience measure which equals log household (respondent) income per capita. By construction, this "true" patience variable has a coefficient of one both when regressing household income on individual patience, and when regression average household income (as proxy for GDP p/c) on average patience.
3. From this "true" patience, we generate a synthetic "observed" patience variable, which equals $p_o = p_t + \alpha \times \epsilon$, where p_t is the respondent's true patience, α a

scaling parameter and $\epsilon \sim \mathcal{N}(0, 1)$ a noise term (recall that observed patience is also normalized to have mean zero and standard deviation of one, so the noise term has the same magnitude as patience).

4. We regress household income on this "observed" patience variable. We also aggregate household income and "observed" patience at the country level and regress average household income on average "observed" patience. We then scale α so that the ratio of observed coefficients equals 39 as in our actual data. It turns out that this requires $\alpha \approx 6$.

5. We evaluate whether $\alpha = 6$ is reasonable. To do so, we generate two separate synthetic "observed" patience variables from the same underlying synthetic "true" patience (this can be thought of as eliciting patience from the same respondent twice). For each individual, we take $\alpha = 6$ and the noise terms ϵ are independent across individuals and "observed" patience variables. The correlation between these two synthetic "observed" patience variables (conditional on country fixed effects) is $\rho = 0.02$.

Given that experimental studies report test-retest correlations for preferences in the ballpark of 0.6 (Beauchamp et al., 2011), we conclude that $\alpha = 6$ is much too large to be meaningful. Indeed, to generate a correlation of 0.6 between the two synthetic "observed" patience measures, we would need to assume $\alpha = 0.75$. But with such a low α, the ratio of coefficients across levels of aggregation equals just 1.7, which is much lower than the observed ratio of 39.

C.2 Censoring

The patience variable is subject to left-censoring because we can only estimate an upper bound on patience for those respondents who always opt for the immediate payment in the quantitative staircase task. In our data, this is true for 56% of respondents. This left-censoring can lead to expansion bias. If such expansion bias was stronger at the country level than at the individual level, this could drive the observed pattern of coefficient magnitudes.

To check whether this is the case, we conduct two separate exercises. The first one is very similar to the thought experiment regarding measurement error above: we generate a synthetic patience measure, censor the measure and then investigate how this effects the OLS coefficients at the individual and country level:

1. Generate a synthetic "true" patience measure which equals log household (respondent) income per capita. By construction, this "true" patience variable has a coefficient of one both when regressing household income on individual patience, and when regression average household income (as proxy for GDP p/c) on average patience.

2. From this “true” patience, we generate a synthetic “observed” patience variable, which is censored: we replace all patience values below the median with the median “true” patience.

3. We regress household income on this “observed” patience variable. We also aggregate household income and “observed” patience at the country level and regress average household income on average “observed” patience. The ratio of observed coefficients equals 1.4, much lower than in the actual data.

As a second – conceptually different – exercise, we remove all censored individuals from the sample and then re-run the OLS regressions of (i) household income on patience, conditional on country fixed effects, and (ii) of average household income on average patience. If censoring drove the difference in coefficients, then this exercise might considerably reduce the coefficient ratio. However, in these regressions, the ratio of patience coefficients is 41, i.e., almost exactly identical to the coefficient ratio when we employ the full sample of respondents. We conclude from these two exercises that censoring is unlikely to drive the amplified patience coefficient at higher levels of aggregation.

D Additional Tables and Figures

Table D.1: Patience and national income: additional control variables

	Dependent variable: Log [GDP p/c PPP]					
	(1)	(2)	(3)	(4)	(5)	(6)
Patience	2.03***	1.99***	1.90***	1.54**	1.57**	1.50**
	(0.30)	(0.33)	(0.38)	(0.60)	(0.62)	(0.71)
Trust	0.0098	0.19	0.042	0.034	-0.091	-0.13
	(0.38)	(0.35)	(0.41)	(0.50)	(0.50)	(0.55)
Risk taking	-0.77**	-0.62*	-0.44	-0.66	-0.41	-0.41
	(0.32)	(0.34)	(0.36)	(0.43)	(0.51)	(0.53)
Mean elevation		-0.89*	-1.33**	-0.94*	-1.17*	-1.09
		(0.51)	(0.50)	(0.53)	(0.66)	(0.80)
Standard deviation of elevation		-0.28	0.13	0.17	0.25	0.24
		(0.48)	(0.46)	(0.42)	(0.48)	(0.52)
Terrain roughness		3.13***	3.07***	1.56	2.45	2.22
		(1.03)	(1.11)	(1.40)	(1.93)	(2.12)
Mean distance to nearest waterway		-0.57*	-0.73**	-0.80**	-0.66*	-0.68*
		(0.29)	(0.32)	(0.36)	(0.33)	(0.39)
1 if landlocked		0.29	0.41	0.26	0.17	0.18
		(0.33)	(0.34)	(0.40)	(0.38)	(0.43)
Log [Area]		0.15	0.18	0.15	0.12	0.11
		(0.11)	(0.11)	(0.12)	(0.11)	(0.13)
Linguistic fractionalization			0.043	0.14	-0.24	-0.19
			(0.41)	(0.47)	(0.53)	(0.57)
Religious fractionalization			-0.22	-0.62	-0.62	-0.63
			(0.46)	(0.54)	(0.60)	(0.77)
Ethnic fractionalization				0.063	0.47	0.41
				(0.75)	(0.69)	(0.69)
% of European descent						0.065
						(0.76)
Genetic distance to the U.S. (weighted)						-0.014
						(0.06)
Additional controls	Yes	Yes	Yes	Yes	Yes	Yes
Legal origin FE	No	No	No	Yes	Yes	Yes
Major religion shares	No	No	No	No	Yes	Yes
Observations	74	74	72	72	72	71
R^2	0.86	0.88	0.88	0.91	0.93	0.93

OLS estimates, robust standard errors in parentheses. Major religion shares include the share of Protestants, Catholics, Muslims, Buddhists, Hinduists, and Atheists. See column (4) of Table 1 for a complete list of the additional controls. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

Table D.2: Patience and national income in sub-samples

	Dependent variable: Log [GDP p/c PPP] in...							
	Africa & Middle East	Europe & C. Asia	SE Asia & Pacific	Americas	OECD	Non-OECD	Colonized	Not colonized
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience	2.49***	1.55***	3.28***	2.16***	1.03***	1.32**	2.18***	2.00***
	(0.70)	(0.24)	(0.96)	(0.34)	(0.16)	(0.56)	(0.32)	(0.46)
Observations	20	27	14	15	22	54	54	22
R^2	0.28	0.48	0.40	0.53	0.62	0.07	0.30	0.46

OLS estimates, robust standard errors in parentheses. In the first column, the sample includes Africa and the Middle East, in the second column Europe and Central Asia, in the third South-East Asia and Pacific, in the fourth the Americas, in the fifth (sixth) all (non-) OECD members, and the seventh (eigth) all formerly colonized (never colonized) countries. The regional groups follow the World Bank definitions. * ($p < 0.10$), ** ($p < 0.05$), ***($p < 0.01$)

Table D.3: Patience and Economic growth

	Dependent variable: Annual growth rate in GDP p/c (in %) since...					
	1950			1975		
	(1)	(2)	(3)	(4)	(5)	(6)
Patience	0.83**	1.00**	1.25***	0.75*	1.42***	1.96***
	(0.32)	(0.41)	(0.30)	(0.41)	(0.40)	(0.43)
Log [GDP p/c base year]		-0.81***	-1.18***		-1.02***	-1.65***
		(0.22)	(0.19)		(0.19)	(0.27)
Continent FE	No	Yes	Yes	No	Yes	Yes
Additionals controls	No	No	Yes	No	No	Yes
Observations	62	62	62	68	68	67
R^2	0.09	0.54	0.81	0.04	0.57	0.75

OLS estimates, robust standard errors in parentheses. See column (5) of Table 1 for a complete list of the additional controls. * ($p < 0.10$), ** ($p < 0.05$), ***($p < 0.01$)

E Description and Sources of Main Variables

E.1 Country-Level Outcome Variables

GDP per capita. Most recent available data point starting at 2016. Source: World Bank Development Indicators.

Annual growth rates. Computed from Maddison dataset.

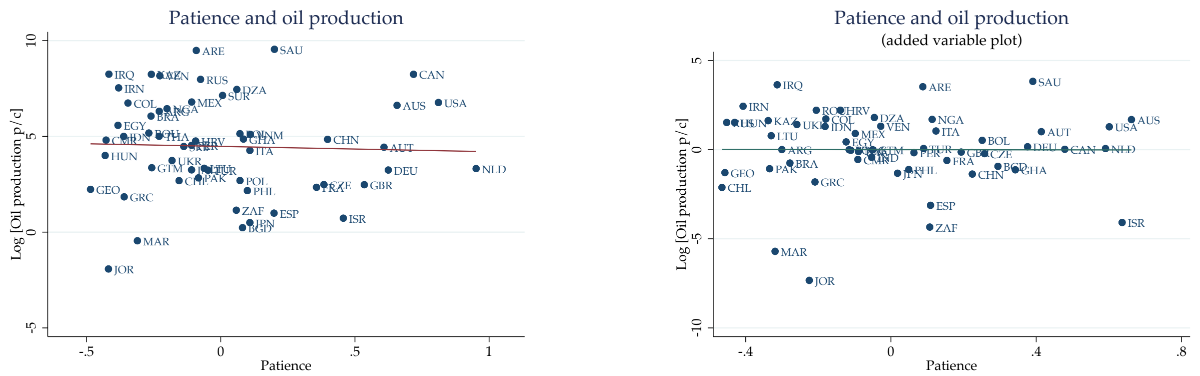
Average years of schooling. Barro and Lee (2012), data in 2010 for population aged 25 and over.

Enrolment rates. Barro and Lee (2012), year 2010.

Table D.4: Individual-level evidence: Controlling for access to credit

	Dependent variable:							
	Log [HH income p/c]				1 if at least secondary educ.			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience	0.069*	0.041***	0.037***	0.029**	0.065***	0.059***	0.058***	0.022**
	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
1 if owns credit card	0.94***	0.60***	0.56***	0.49***	0.23***	0.25***	0.23***	0.17***
	(0.14)	(0.08)	(0.08)	(0.07)	(0.03)	(0.02)	(0.02)	(0.01)
Age				28.3				-61.5*
				(28.15)				(30.72)
Age squared				2034.1				-3721.1
				(3785.22)				(3018.84)
1 if female				-0.13***				-0.014
				(0.02)				(0.01)
Subj. math skills				0.037***				0.033***
				(0.00)				(0.00)
Subjective institutional quality				-0.12***				-0.11***
				(0.04)				(0.03)
Confidence in financial institutions				7.90**				1.99
				(2.74)				(1.30)
Subjective law and order index				0.049				-0.013
				(0.05)				(0.02)
Country FE	No	Yes	No	No	No	Yes	No	No
Regional FE	No	No	Yes	Yes	No	No	Yes	Yes
Religion FE	No	No	No	Yes	No	No	No	Yes
Observations	11374	11374	11374	10248	11556	11556	11556	10406
R^2	0.07	0.32	0.37	0.38	0.04	0.09	0.13	0.31

Individual-level OLS estimates, standard errors (clustered at country level) in parentheses. The dependent variable in (1)–(4) is ln household income per capita; the dependent variable in (5) – (8) is 1 if the individual has at least secondary education. Age is divided by 100. We do not report savings regressions because the credit card question was only asked in countries for which savings information is not elicited in the World Poll.
* $(p < 0.10)$, ** $(p < 0.05)$, *** $(p < 0.01)$



(a) Raw correlation between log oil production per capita and patience ($\rho = -0.04$)

(b) Correlation between log oil production per capita and patience conditional on the full set of baseline covariates in column (5) of Table 1.

Figure D.1: Patience and oil production per capita (in 2014 Dollars)

Math and science test scores. Measure of cognitive skills derived from a series of standardized tests across countries, see Hanushek and Woessmann (2012).

Education expenditure. Most recent available data point starting at 2016. Source: World Bank Development Indicators.

Capital stock. Data taken from the Penn World Tables.

National savings. Gross savings are calculated as gross national income less total consumption, plus net transfers. Net national savings are equal to gross national savings less the value of consumption of fixed capital. Adjusted net savings are equal to net national savings plus education expenditure and minus energy depletion, mineral depletion, net forest depletion, and carbon dioxide and particulate emissions damage. Most recent available data point starting at 2016. Source: World Bank Development Indicators.

Household savings rate. The household saving rate is calculated as the ratio of household saving to household disposable income (plus the change in net equity of households in pension funds). Source: QOG database.

Total factor productivity. TFP level at current PPPs (USA = 1), taken from QOG dataset.

R & D expenditure. Expenditures for research and development are current and capital expenditures (both public and private) on creative work undertaken systematically to increase knowledge, including knowledge of humanity, culture, and society, and the use of knowledge for new applications. R & D covers basic research, applied research, and experimental development. Most recent available data point starting at 2016. Source: World Bank Development Indicators.

Number of researchers in R & D. Researchers in R & D are professionals engaged in the conception or creation of new knowledge, products, processes, methods, or systems and in the management of the projects concerned. Most recent available data point starting at 2016. Source: World Bank Development Indicators.

Democracy index. Index that quantities the extent of institutionalized democracy, as reported in the Polity IV dataset. Taken from QOG dataset.

Property rights. This factor scores the degree to which a country's laws protect private property rights and the degree to which its government enforces those laws. It also accounts for the possibility that private property will be expropriated. In addition, it analyzes the independence of the judiciary, the existence of corruption within the judiciary, and the ability of individuals and businesses to enforce contracts. Average 2003 – (2012), taken from the QOG dataset.

Oil production per capita. Oil production per capita in 2014 Dollars. Taken from Quality of Government dataset.

E.2 Regional-Level Data

Except for the patience measures and a region's size (area), all regional-level data are taken from Gennaioli et al. (2013). The area data were collected by research assistants from various sources on the web.

E.3 Individual-Level Data

Household income per capita. Included in Gallup's background data. To calculate income, respondents are asked to report their household income in local currency. Those respondents who have difficulty answering the question are presented a set of ranges in local currency and are asked which group they fall into. Income variables are created by converting local currency to International Dollars (ID) using purchasing power parity (PPP) ratios. Log household income is computed as log (1+ household income).

Education level. Included in Gallup's background data. Level 1: Completed elementary education or less (up to 8 years of basic education). Level 2: Secondary - 3 year tertiary education and some education beyond secondary education (9 – 15 years of education).

Level 3: Completed four years of education beyond high school and / or received a 4-year college degree.

Subjective self-assessment of math skills. Included in Gallup's background data. *How well do the following statements describe you as a person? Please indicate your answer on a scale from* 0 *to* 10*. A 0 means "does not describe me at all" and a* 10 *means "describes me perfectly". You can also use any numbers between* 0 *and* 10 *to indicate where you fall on the scale, like* 0*,* 1*,* 2*,* 3*,* 4*,* 5*,* 6*,* 7*,* 8*,* 9*,* 10*. I am good at math.*

Saved last year. Binary variable capturing whether the respondent saved any money in the previous year. Included in Gallup's background data.

Confidence in financial institutions. Included in Gallup's background data. Binary response to the question "In this country, do you have confidence in each of the following, or not? How about financial institutions or banks?"

Subjective institutional quality. Included in Gallup's background data. This variable consists of a perceived institutional quality index as it is provided by Gallup. This index combines binary questions (yes / no) about whether people have confidence in the military, the judicial system and courts, the national government, and the honesty of elections. The index is then constructed by averaging the yes / no answers across items.

Subjective law and order index. Included in Gallup's Background data. Derived from responses to three questions: "In the city or area where you live, do you have confidence in the local police force?"; "Do you feel safe walking alone at night in the city or area where you live?"; "Within the last 12 months, have you had money or property stolen from you or another household member?".

F Characterization of the Equilibrium

F.1 Equilibrium: Definition and Existence

Definition 1. *For a given distribution of patience, i.e. for given χ and ε, a steady-state equilibrium is a skill share λ with an associated patience threshold $\tilde{\beta}$, and positive real numbers (w^H, w^L, R, H, L, K) such that:*

a) The prices (w^H, w^L, R) are determined competitively and satisfy $w^H = \frac{\partial Y}{\partial H}$, $w^L = \frac{\partial Y}{\partial L}$, and $R = 1 + r = \frac{\partial Y}{\partial K}$.

b) The threshold value $\tilde{\beta}$ separates households into skilled and unskilled workers, such that $\eta^s = \eta^d$ at the threshold.

c) The factor markets clear with aggregate amounts of skilled labor, unskilled labor and capital given by $H = \lambda e^{\rho(1-\psi)}$, $L = 2(1-\lambda) + \psi\lambda$, and (16), respectively.

Proposition 1. *The steady-state equilibrium exists and is unique.*

Proof. Consider the indifference condition for education (6). On the aggregate, this condition characterizes the threshold level of patience, $\tilde{\beta}$, as function of the emerging earnings premium. Rewriting this condition delivers an implicit expression for the (relative) supply of skilled labor,

$$\eta^s = \frac{\psi^{\frac{-1}{\tilde{\beta}(1+\tilde{\beta})}}}{h} \ . \tag{13}$$

which implies that η^s is a decreasing, convex function in the patience threshold $\tilde{\beta}$.

The relative demand for skilled labor is obtained from the condition for the skill premium that emerges from firm optimization,

$$\eta^d = \frac{w^H}{w^L} = \left[K^\theta + H^\theta\right]^{\frac{\sigma-\theta}{\theta}} \frac{L^{1-\sigma}}{H^{1-\theta}} = \left[\left(\frac{K}{H}\right)^\theta + 1\right]^{\frac{\sigma-\theta}{\theta}} \left(\frac{L}{H}\right)^{1-\sigma} \ , \tag{14}$$

which is an increasing function in the patience threshold $\tilde{\beta}$ conditional on K. Moreover, the skill premium η is an increasing function in K, with capital demand given by

$$K^d = \left[\eta^{\frac{\theta}{\sigma-\theta}} \cdot \left(\frac{H}{L}\right)^{\frac{\theta(1-\sigma)}{\sigma-\theta}} - 1\right]^{\frac{1}{\theta}} \cdot H \ , \tag{15}$$

while in steady state the supply of capital given by (10) simplifies to

$$K^s = \frac{1}{2\varepsilon} w^L \left\{ \left[2\varepsilon(1-\lambda) - \ln\left(1 + \frac{2\varepsilon(1-\lambda)}{1+\chi-\varepsilon}\right)\right] + \eta h \left[2\varepsilon\lambda + \ln\left(1 - \frac{2\varepsilon\lambda}{1+\chi+\varepsilon}\right)\right]\right\} \ , \tag{16}$$

which is an increasing function in η.

Now note that *(a)* is satisfied via (14) and the fact that competitive markets for capital determine R, while constant returns to scale of the production function ensure that factor rents exhaust all production such that the market for capital clears by Walras's Law. Additionally, combining (16) with (14) and using the expressions for L and H in *(c)* gives

$$K = \left\{2\varepsilon(1-\lambda) - \ln\left(1 + \frac{2\varepsilon(1-\lambda)}{1+\chi-\varepsilon}\right) + \eta^d e^{\rho(1-\psi)}\left[2\varepsilon\lambda + \ln\left(1 - \frac{2\varepsilon\lambda}{1+\chi+\varepsilon}\right)\right]\right\}$$

$$\cdot \frac{A\left[\left(\frac{H}{L}\right)^{\frac{\sigma(1-\theta)}{\sigma-\theta}} \cdot \left(\eta^d\right)^{\frac{\sigma}{\sigma-\theta}} + 1\right]^{\frac{1-\sigma}{\sigma}}}{2\varepsilon}, \tag{17}$$

so that any solution to *(b)* will also fulfill *(c)*. Existence and uniqueness then follows since (15) and (17) determine a unique tuple of η and K, which in turn via (14) and (13) determines a unique tuple of η and $\tilde{\beta}$ for a given (unique) K. To see this, note that from (15) K^d is smaller than zero for $\eta = 0$ and increasing, convex in η if $\sigma \leq 2\theta$ while from (17) K^s is greater than zero for $\eta = 0$ and monotonically increasing in η. In addition, for $\sigma > \theta$ $\frac{K^d}{K^s}$ is strictly monotonically increasing for $\eta > 1$ and $\lim_{\eta\to\infty} \frac{K^d}{K^s} = \infty$ so that K^s and K^d intersect exactly once. As a result there exists a unique tuple (η, K) for which $K^s = K^d$. This also pins down η^d. It follows that there exists a unique $\tilde{\beta}$ such that $\eta^s = \eta^d$ as (13) is a decreasing, convex function in the patience threshold $\tilde{\beta}$ that approaches infinity from below as $\tilde{\beta} \to 0$, while (14) is increasing in the threshold. The resulting skill share λ uniquely determines K and the remaining variables of the steady-state equilibrium. This establishes a unique equilibrium allocation for which relative demand for skills and skill supply are consistent with the demand and supply for capital, and a population share of skilled workers λ and the corresponding aggregate allocation and factor prices that is consistent with a cut-off level of patience, $\tilde{\beta}$, that corresponds to the optimal individual decisions to remain unskilled workers (individuals with a $\beta(i) < \tilde{\beta}$) and skilled workers (individuals with a $\beta(i) > \tilde{\beta}$) at the prevailing factor prices.

□

F.2 Comparing Steady States: The Effect of χ on λ

The share of skilled individuals is given by

$$\lambda = \frac{\chi + \varepsilon - \tilde{\beta}}{2\varepsilon},$$

with the total derivative given by

$$d\,\lambda - \frac{1}{2\varepsilon}\cdot d\,\chi + \frac{1}{2\varepsilon}\cdot d\,\tilde{\beta} = 0\ .$$

Therefore

$$\frac{d\,\lambda}{d\,\chi} = \frac{1}{2\varepsilon}\left(1 - \frac{d\,\tilde{\beta}}{d\,\chi}\right)\ .$$

Equating (13) and (14), inserting (16), $H = \lambda e^{\rho(1-\psi)}$, $L = 2(1-\lambda) + \psi\lambda$, and computing the total derivative yields

$$\tilde{\psi}\cdot\frac{d\,\tilde{\beta}}{d\,\chi} = (\sigma-\theta)\left[\tilde{K}\cdot\rho_K + \tilde{H}\cdot\rho_H\right] + (1-\sigma)\rho_L - (1-\theta)\rho_H\ ,$$

where

$$\tilde{\psi} = \frac{1+2\tilde{\beta}}{\left[\tilde{\beta}(1+\tilde{\beta})\right]^2}\cdot\ln\psi \quad < 0\ ,$$

$$\tilde{K} = \frac{K^\theta}{K^\theta + H^\theta} \quad \in (0,1)\ ,$$

$$\tilde{H} = \frac{H^\theta}{K^\theta + H^\theta} \quad \in (0,1)\ ,$$

$$\rho_K = \frac{1}{K}\frac{d\,K}{d\,\chi} = \frac{1-\sigma}{\sigma-\theta}\left[\tilde{\psi}\cdot\frac{d\,\tilde{\beta}}{d\,\chi} + (1-\theta)(\rho_H - \rho_L)\right]\tilde{x} + \tilde{\psi}\frac{\bar{K}_{w^H}}{K} + \frac{\Delta K}{K}\left(1 - \frac{d\,\tilde{\beta}}{d\,\chi}\right)\ ,$$

$$\rho_H = \frac{1}{H}\frac{d\,H}{d\,\chi} = \frac{1}{\chi+\varepsilon-\tilde{\beta}}\left(1 - \frac{d\,\tilde{\beta}}{d\,\chi}\right)\ ,$$

$$\rho_L = \frac{1}{L}\frac{d\,L}{d\,\chi} = -\frac{2-\psi}{(2-\psi)(\tilde{\beta}-\chi) + (2+\psi)\varepsilon}\left(1 - \frac{d\,\tilde{\beta}}{d\,\chi}\right)\ ,$$

$$\tilde{x} = \frac{\left(K^\theta + H^\theta\right)^{\frac{\sigma}{\theta}}}{(K^\theta + H^\theta)^{\frac{\sigma}{\theta}} + L^\sigma} \quad \in (0,1)\ ,$$

$$\frac{\bar{K}_{w^H}}{K} = \frac{1}{K}\cdot w^H h\left[\lambda - \frac{1}{2\varepsilon}\ln\left(\frac{1+\chi+\varepsilon}{1+\tilde{\beta}}\right)\right] \quad \in (0,1)\ ,$$

$$\frac{\Delta K}{K} = \frac{1}{K}\left[\frac{(1-\lambda)\cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda\cdot w^H h}{1+\chi+\varepsilon}\right] \quad \in (0,1)\ .$$

Re-arranging and solving for $\frac{d\tilde{\beta}}{d\chi}$ gives

$$\frac{d\tilde{\beta}}{d\chi} = \frac{\tilde{z} + \tilde{\psi}(\sigma-\theta)\tilde{K}\frac{\bar{K}_{wH}}{K}}{\tilde{z} + \tilde{\psi}\left[1-(1-\sigma)\tilde{x}\tilde{K}\right]} \,,$$

where

$$\tilde{z} = \frac{(1-\theta)\left[(1-\sigma)\tilde{x}\tilde{K}-1\right] + (\sigma-\theta)\left[\tilde{H} + (\chi+\varepsilon-\tilde{\beta})\tilde{K}\frac{\Delta K}{K}\right]}{\chi+\varepsilon-\tilde{\beta}} + \underbrace{\frac{(2-\psi)(1-\sigma)\left[(1-\theta)\tilde{x}\tilde{K}-1\right]}{(2-\psi)(\tilde{\beta}-\chi)+(2-\psi)\varepsilon}}_{<0 \text{ as } (1-\theta)\tilde{x}\tilde{K}<1} \,.$$

Note that $\tilde{z} < 0$ if

$$\theta > \frac{(1-\sigma)\tilde{x}\tilde{K} - 1 + \sigma\left[\tilde{H} + (\chi+\varepsilon-\tilde{\beta})\tilde{K}\frac{\Delta K}{K}\right]}{(1-\sigma)\tilde{x}\tilde{K} - 1 + \left[\tilde{H} + (\chi+\varepsilon-\tilde{\beta})\tilde{K}\frac{\Delta K}{K}\right]} \quad \in (0,1) \,.$$

Therefore, as $\tilde{\psi} < 0$, $\frac{d\tilde{\beta}}{d\chi} > 0$ and $\frac{d\tilde{\beta}}{d\chi} < 1$ if

$$\sigma < \underbrace{\frac{1}{\tilde{K}\left(\frac{\bar{K}_{wH}}{K} - \tilde{x}\right)}}_{>1} + \underbrace{\frac{\tilde{K}\left(\theta\frac{\bar{K}_{wH}}{K} - \tilde{x}\right)}{\tilde{K}\left(\frac{\bar{K}_{wH}}{K} - \tilde{x}\right)}}_{\in(0,1)} \,,$$

which holds for $\sigma < 1$. As a result, if $0 < \theta < \sigma < 1$

$$\frac{d\tilde{\beta}}{d\chi} > 0 \,, \quad \frac{d\tilde{\beta}}{d\chi} < 1 \,, \quad \frac{d\lambda}{d\chi} > 0 \,, \quad \frac{dL}{d\chi} < 0 \,, \quad \frac{dH}{d\chi} > 0 \,, \quad \frac{dK}{d\chi} > 0 \,.$$

G Testable Predictions: Derivation of Expressions

Set-up: Consider a country with a uniform distribution of patience between $\chi - \epsilon$ and $\chi + \epsilon$. The mean and standard deviation are thus given by

$$\mu = \chi \quad \text{and} \quad sd = \frac{2\varepsilon}{\sqrt{12}} .$$

G.1 The Effect of Patience on Education

Effect of a one standard deviation increase in patience on the propensity to become skilled at the individual level: The decision to become skilled for an individual i is determined by the occupation choice trade-off given by (6),

$$\eta^* > \eta(i) = \frac{\psi^{\frac{-1}{\beta(i)(1+\beta(i))}}}{h} ,$$

where η^* denotes the equilibrium wage premium for skilled workers. Hence, the decision to become skilled corresponds to a binary choice problem where the value of the latent variable $\eta(i)$ relative to the market (equilibrium) earnings premium, η^*, determines the propensity of becoming skilled as skilled$(i) = (\eta^* - \eta(i)) + \epsilon(i)$, where the expression in parentheses corresponds to the economic considerations related to the decision to become skilled, and ϵ is an additive noise term that reflects idiosyncratic factors influencing this decision and is drawn from a symmetric distribution. From a population perspective, it then holds that the predicted share of skilled individuals is given by

$$\widehat{Pr}(\text{skilled}) = \widehat{Pr}(\eta(i) < \eta^*) = \int_{\chi-\varepsilon}^{\chi+\varepsilon} \text{skilled}(i) f(\beta) d\,\beta = \lambda .$$

Notice that since $\beta(i)$ is uniformly distributed and $\eta(i)$ is a monotonic function of $\beta(i)$, the distribution (and pdf) of $\eta(i)$ can be derived accordingly by transformation. Since the decision of individuals that are skilled is insensitive to increases in patience, the marginal effect of an increase in patience only affects individuals with $\eta(i) > \eta^*$.[26] The marginal effect of an increase in patience on the propensity of becoming skilled can then be expressed as

$$\frac{\partial skilled(i)}{\partial \beta(i)} = f\left(\eta(i)\right) \cdot \left| \frac{d\,\beta(i)}{d\,\eta(i)} \right|$$

$$\text{with} \quad f\left(\eta(i)\right) \cdot \left| \frac{d\,\beta(i)}{d\,\eta(i)} \right| = -\frac{1}{2\varepsilon} \cdot \frac{\left[\beta(i)\left(1+\beta(i)\right)\right]^2}{\frac{\psi^{\frac{-1}{\beta(i)(1+\beta(i))}}}{h} \ln \psi \left(1 + 2\beta(i)\right)} ,$$

[26] Also note that since $\eta(i)$ decreases in $\beta(i)$, the cdf of η is inverse to the cdf in β.

where $f(\cdot)$ is the pdf of the distribution of β and the second line follows by the fact that $\ln\psi < 0$. As a result the marginal effect of a one standard deviation increase in patience evaluated at the threshold on the propensity to become skilled is given by

$$sd \cdot \frac{\partial skilled(i)}{\partial \beta(i)}\bigg|_{\beta(i)=\tilde{\beta}} = sd \cdot \left(-\frac{1}{2\varepsilon} \cdot \frac{\left[\tilde{\beta}\left(1+\tilde{\beta}\right)\right]^2}{\frac{\psi^{\frac{-1}{\tilde{\beta}(1+\tilde{\beta})}}}{h} \ln\psi \left(1+2\tilde{\beta}\right)} \right)$$

$$= -\frac{1}{\sqrt{12}} \cdot \frac{\left[\tilde{\beta}\left(1+\tilde{\beta}\right)\right]^2}{\frac{\psi^{\frac{-1}{\tilde{\beta}(1+\tilde{\beta})}}}{h} \ln\psi \left(1+2\tilde{\beta}\right)} \ . \tag{18}$$

Notice that the OLS estimates in Table 8 correspond to the population average of a linearized version of this effect under the assumption that $\eta(i)$ is affected by an additive and symmetrically distributed disturbance term. From an economic perspective, as individuals around the threshold should be most sensitive in a deterministic model, evaluating the elasticity at the threshold appears most appropriate. Moreover, using this expression delivers a conservative benchmark for the numerical model analogue to the empirical moment in the context of demonstrating the model's ability to generate an amplification of the patience elasticity of education at the aggregate level.

Effect of a one standard deviation increase in patience on the share of skilled at the aggregate level: Consider a one standard deviation shift in patience at the mean. For concreteness, consider the comparison of a second country with $\chi_2 = \chi_1 + sd$. Notice that the results in Section F.2 imply that

$$\frac{d\lambda}{d\chi} = \frac{1}{2\varepsilon}\left(1 - \frac{d\tilde{\beta}}{d\chi}\right) > 0 \ .$$

The estimated effect of a one standard deviation increase in average patience (with $\chi_2 = \chi_1 + sd$) on the share of skilled at the aggregate level is then given by

$$
\begin{aligned}
\lambda_2 - \lambda_1 \equiv \Delta\lambda =& \frac{1}{2\varepsilon}\left(\chi_2 + \varepsilon - \tilde{\beta}_2 - (\chi_1 + \varepsilon - \tilde{\beta}_1)\right) \\
=& \frac{1}{2\varepsilon}\left(\chi_1 + sd + \varepsilon - \tilde{\beta}_2 - (\chi_1 + \varepsilon - \tilde{\beta}_1)\right) \\
=& \frac{1}{2\varepsilon}\left(sd - \underbrace{(\tilde{\beta}_2 - \tilde{\beta}_1)}_{\equiv\Delta\beta}\right) \\
\Delta\lambda =& \frac{1}{\sqrt{12}} - \frac{1}{2\varepsilon}\cdot\Delta\beta\ .
\end{aligned}
\tag{19}
$$

Amplification: The aggregate effect is quantitatively larger than the individual effect if

$$
\begin{aligned}
\frac{1}{\sqrt{12}} - \frac{1}{2\varepsilon}\cdot\Delta\beta >& -\frac{1}{\sqrt{12}}\cdot\frac{\left[\tilde{\beta}\left(1+\tilde{\beta}\right)\right]^2}{\frac{\psi^{\frac{-1}{\tilde{\beta}(1+\tilde{\beta})}}}{h}\ln\psi\left(1+2\tilde{\beta}\right)} \\
1 - \frac{\Delta\beta}{sd} >& -\frac{\left[\tilde{\beta}\left(1+\tilde{\beta}\right)\right]^2}{\frac{\psi^{\frac{-1}{\tilde{\beta}(1+\tilde{\beta})}}}{h}\ln\psi\left(1+2\tilde{\beta}\right)} \\
1 - \frac{d\,\tilde{\beta}}{d\,\chi} >& -\frac{\left[\tilde{\beta}\left(1+\tilde{\beta}\right)\right]^2}{\frac{\psi^{\frac{-1}{\tilde{\beta}(1+\tilde{\beta})}}}{h}\ln\psi\left(1+2\tilde{\beta}\right)}\ .
\end{aligned}
$$

Note that the right-hand side goes to zero as $\tilde{\beta} \to 0$, while the left-hand side is strictly positive (due to $\frac{d\,\tilde{\beta}}{d\,\chi} in(0,1)$). On the other hand, as $\tilde{\beta} \to 1$ the left-hand side converges to one. As the right-hand side is increasing $\tilde{\beta}$, the aggregate effect is quantitatively larger than the individual effect if the following parametric restriction holds

$$
1 > -\frac{\left[(\chi+\varepsilon)\left(1+\chi+\varepsilon\right)\right]^2}{\frac{\psi^{\frac{-1}{(\chi+\varepsilon)(1+\chi+\varepsilon)}}}{e^{\rho(1-\psi)}}\ln\psi\left[1+2\left(\chi+\varepsilon\right)\right]}\ .
$$

G.2 The Effect of Patience on Savings

Effect of a one standard deviation increase in patience on the savings rate at the individual level: The savings rate of individual i is given by

$$s(i) = \frac{\beta(i)}{1+\beta(i)} ,$$

as such the marginal effect of an increase in patience is given by

$$\frac{\partial s(i)}{\partial \beta(i)} = \frac{1}{[1+\beta(i)]^2} .$$

Therefore the marginal effect of a one standard deviation increase in patience at the mean on the individual savings rate is given by

$$sd \cdot \left.\frac{\partial s(i)}{\partial \beta(i)}\right|_{\beta(i)=\chi} = sd \cdot \frac{1}{(1+\chi)^2} .$$

Effect of a one standard deviation increase in patience on (average) log savings at the individual level: The average marginal effect on individual savings is given by

$$\frac{\partial \bar{S}}{\partial \beta(i)} = \frac{1}{2\varepsilon}\left[w^L \int_{\chi-\varepsilon}^{\tilde{\beta}} \frac{1}{[1+\beta(i)]^2}\, d\,\beta(i) + w^H h \int_{\tilde{\beta}}^{\chi+\varepsilon} \frac{1}{[1+\beta(i)]^2}\, d\,\beta(i)\right]$$

$$= \frac{1}{1+\tilde{\beta}}\left[\frac{(1-\lambda)\cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda \cdot w^H h}{1+\chi+\varepsilon}\right] .$$

The marginal effect of an increase in patience on (average) log individual savings is given by

$$\frac{\partial \ln \bar{S}}{\partial \beta(i)} = \frac{1}{\bar{S}}\frac{\partial \bar{S}}{\partial \beta(i)} .$$

Therefore the marginal effect of a one standard deviation increase in patience on (average) log individual savings is given by

$$sd \cdot \frac{\partial \ln \bar{S}}{\partial \beta(i)} = sd \cdot \frac{1}{\bar{S}}\left\{\frac{1}{1+\tilde{\beta}}\left[\frac{(1-\lambda)\cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda \cdot w^H h}{1+\chi+\varepsilon}\right]\right\} . \tag{20}$$

Aggregate effect of a one standard deviation increase in patience on log capital per capita: As in the previous cases, consider a one standard deviation shift in patience at the mean. For convenience, denote capital per capita by k.[27] The increase in log capital per capita is then given by

[27] Note that with each generation having size (mass) 1, this is equivalent to $\frac{K}{3}$.

$$\frac{k_2 - k_1}{k_1} \; . \tag{21}$$

Amplification: The aggregate effect is quantitatively larger then the individual effect if

$$\frac{k_2 - k_1}{k_1} > sd \cdot \frac{1}{\bar{S}} \left\{ \frac{1}{1+\tilde{\beta}} \left[\frac{(1-\lambda) \cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda \cdot w^H h}{1+\chi+\varepsilon} \right] \right\}$$

$$\frac{k_2 - k_1}{sd} > \frac{1}{1+\tilde{\beta}} \left[\frac{(1-\lambda) \cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda \cdot w^H h}{1+\chi+\varepsilon} \right] \; .$$

Notice that from the equivalence of aggregate capital with aggregate savings, it holds that

$$\begin{aligned} \frac{dK}{d\chi} =& S^L \cdot \frac{dw^L}{d\chi} + S^H \cdot \frac{dw^H h}{d\chi} + \frac{\tilde{\beta}}{1+\tilde{\beta}} (w^H h - w^L) \cdot \frac{d\lambda}{d\chi} + \frac{1}{1+\tilde{\beta}} \left[\frac{(1-\lambda) \cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda \cdot w^H h}{1+\chi+\varepsilon} \right] \\ =& \underbrace{S^L \cdot \frac{dw^L}{d\chi} + S^H \cdot \frac{dw^H h}{d\chi} + \frac{\tilde{\beta}}{1+\tilde{\beta}} (w^H h - w^L) \cdot \frac{d\lambda}{d\chi}}_{\text{GE effects}} + \underbrace{\frac{\partial \bar{S}}{\partial \beta(i)} \frac{\partial \beta(i)}{\partial \chi}}_{\text{individual effect}} \; , \end{aligned}$$

where $\frac{\partial \beta(i)}{\partial \chi} = 1$ as consequence of a shift of a uniform distribution. Hence, $\frac{dK}{d\chi} > \frac{\partial \bar{S}}{\partial \beta(i)}$ holds if

$$S^L \cdot \frac{dw^L}{d\chi} + S^H \cdot \frac{dw^H h}{d\chi} + \frac{\tilde{\beta}}{1+\tilde{\beta}} (w^H h - w^L) \cdot \frac{d\lambda}{d\chi} > 0 \; ,$$

i.e. the general equilibrium effects of a change in average patience do not counteract the individual effects. Notice that this expression simplifies to

$$\tilde{a} - \left(\tilde{a} - \tilde{\psi} \tilde{b} \right) \cdot \frac{d\tilde{\beta}}{d\chi} > 0 \; ,$$

with

$$\tilde{a} = \frac{(1-\sigma)(1-\theta)}{\sigma - \theta} \frac{\left(S^L + \eta h S^H\right) \tilde{x}}{\lambda \left[(2-\psi)(1-\lambda) + \psi\right]} + \frac{1}{2\varepsilon} \frac{\tilde{\beta}}{1+\tilde{\beta}} (\eta h - 1) > 0$$

$$\tilde{b} = \frac{1-\sigma}{\sigma-\theta} \left(S^L + \eta h S^H\right) \tilde{x} + \eta h S^H > 0$$

For this to hold, it is sufficient that the capital-skill complementarity is strong enough as, for a given σ, $\frac{d\tilde{\beta}}{d\chi}$ falls with decreasing θ (i.e. greater capital-skill complementarity).

G.3 The Effect of Patience on Income

Effect of a one standard deviation increase in patience on (average) log household income: Individual income can be written as

$$\bar{y} = \left[2\left[1 - \mathbb{I}\left(\eta(i) < \eta^*\right)\right] + \psi\right] w^L + \mathbb{I}\left(\eta(i) < \eta^*\right) w^H h + R \cdot \bar{S} \ ,$$

where $\mathbb{I}$ is the indicator function that takes value one if the expression in parentheses is true and zero otherwise.[28]

The average marginal effect on individual income is then given by

$$\begin{aligned}
\frac{\partial \bar{y}}{\partial \beta(i)} =& \frac{\partial \overline{Pr}(\eta(i) < \eta^*)}{\partial \beta(i)} \left[w^H h - (2-\psi) w^L\right] + \frac{1}{2\varepsilon}\left[w^L \int_{\chi-\varepsilon}^{\tilde{\beta}} \frac{1}{\left[1+\beta(i)\right]^2}\, d\,\beta(i) + w^H h \int_{\tilde{\beta}}^{\chi+\varepsilon} \frac{1}{\left[1+\beta(i)\right]^2}\, d\,\beta(i)\right] \cdot R \\
=& \frac{\partial \overline{Pr}(\eta(i) < \eta^*)}{\partial \beta(i)} \left[w^H h - (2-\psi) w^L\right] + \frac{R}{1+\tilde{\beta}}\left[\frac{(1-\lambda)\cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda \cdot w^H h}{1+\chi+\varepsilon}\right] \\
=& \frac{\partial \overline{Pr}(\eta(i) < \eta^*)}{\partial \beta(i)} \left[w^H h - (2-\psi) w^L\right] + R \cdot \frac{\partial \bar{S}}{\partial \beta(i)} \ .
\end{aligned}$$

The marginal effect of an increase in patience on (average) log household income is given by

$$\frac{\partial \ln \bar{y}}{\partial \beta(i)} = \frac{1}{\bar{y}} \frac{\partial \bar{y}}{\partial \beta(i)} \ .$$

Therefore the (average) marginal effect of a one standard deviation increase in patience (with the effect of education evaluated at the threshold) on log household income is given by

$$sd \cdot \left.\frac{\partial \ln \bar{y}}{\partial \beta(i)}\right|_{\beta(i)=\tilde{\beta}} = sd \cdot \frac{1}{\bar{y}} \left\{ \left.\frac{\partial Pr(\eta(i) < \eta^*)}{\partial \beta(i)}\right|_{\beta(i)=\tilde{\beta}} \left[w^H h - (2-\psi) w^L\right] + R \cdot \frac{\partial \bar{S}}{\partial \beta(i)} \right\} \ . \tag{22}$$

Effect of a one standard deviation increase in patience on GDP per capita: Consider a one standard deviation shift in patience at the country mean. Consequently the percentage increase in GDP per capita is given by

$$\frac{y_2 - y_1}{y_1} \ . \tag{23}$$

Amplification: The aggregate effect is quantitatively larger then the individual effect if

[28] For simplicity and without affecting the results, here and below when considering the effects on the aggregate level, we disregard the fact that, with each generation being unit mass, the total population size is 3.

$$\frac{y_2 - y_1}{y_1} > sd \cdot \frac{1}{\bar{y}} \left\{ \left. \frac{\partial Pr(\eta(i) < \eta^*)}{\partial \beta(i)} \right|_{\beta(i)=\tilde{\beta}} \left[w^H h - (2-\psi) w^L \right] + R \cdot \frac{\partial \bar{S}}{\partial \beta(i)} \right\}$$

$$\frac{y_2 - y_1}{sd} > \left. \frac{\partial Pr(\eta(i) < \eta^*)}{\partial \beta(i)} \right|_{\beta(i)=\tilde{\beta}} \left[w^H h - (2-\psi) w^L \right] + R \cdot \frac{\partial \bar{S}}{\partial \beta(i)}$$

$$\frac{dY}{d\chi} > \left. \frac{\partial \bar{y}}{\partial \beta(i)} \right|_{\beta(i)=\tilde{\beta}} ,$$

where

$$\frac{dY}{d\chi} = \left[2(1-\lambda)+\psi\right] \frac{dw^L}{d\chi} + \lambda \frac{dw^H h}{d\chi} + K \cdot \frac{dR}{d\chi} + \frac{d\lambda}{d\chi} \left[w^H h - (2-\psi) w^L \right] + R \cdot \frac{dK}{d\chi}$$

$$= \underbrace{\left[2(1-\lambda)+\psi\right] \frac{dw^L}{d\chi} + \lambda \frac{dw^H h}{d\chi} + K \cdot \frac{dR}{d\chi}}_{\text{GE effects}} + \underbrace{\frac{d\lambda}{d\chi} \left[w^H h - (2-\psi) w^L \right] + R \cdot \frac{dK}{d\chi}}_{\text{direct effects}} .$$

Hence, the effect of patience on income is amplified on the aggregate level if

$$\left[2(1-\lambda)+\psi\right] \frac{dw^L}{d\chi} + \lambda \frac{dw^H h}{d\chi} + K \cdot \frac{dR}{d\chi} > 0 ,$$

i.e., if the general equilibrium effects of a change in average patience are positive. Notice that this expression simplifies to

$$\underbrace{\tilde{c} - \eta(1-\theta)\tilde{\psi}\bar{K}_{w^H} \left(\frac{H}{K}\right)^{1-\theta}}_{>0} - \left(\tilde{c} - \tilde{\psi}\tilde{d}\right) \cdot \frac{d\tilde{\beta}}{d\chi} > 0 ,$$

with

$$\tilde{c} = \frac{(1-\sigma)(1-\theta)}{\sigma-\theta} \frac{\left[2(1-\lambda)+\psi+\lambda\eta h\right]\tilde{x}}{\lambda\left[(2-\psi)(1-\lambda)+\psi\right]} + \eta\Delta K \left(\frac{H}{K}\right)^{1-\theta} > 0$$

$$\tilde{d} = \frac{1-\sigma}{\sigma-\theta} \left[2(1-\lambda)+\psi+\lambda\eta h\right]\tilde{x} + \eta\left(\lambda h + \frac{1-\sigma(2-\theta)}{\sigma-\theta} K^\theta H^{1-\theta}\right) > 0$$

For this to hold, it is sufficient that the capital-skill complementarity is strong enough as, for a given σ, $\frac{d\tilde{\beta}}{d\chi}$ falls with decreasing θ (i.e., greater capital-skill complementarity).

H Structural Estimation

H.1 Estimation Approach

The estimation of the model follows an indirect inference approach, where structural parameters that are relevant for the amplification are estimated by minimizing the distance between the empirical patience elasticities obtained with reduced form estimation and their theoretical counterparts. Parameters that are of no immediate relevance or that have direct empirical counterparts are calibrated. The vector of remaining parameters to be estimated is given by $\Theta = (\chi, \epsilon, A)$. These parameters are estimated by matching the patience elasticities of education and income at the individual and aggregate level. Denote by Z the vector of elasticities obtained from reduced form regressions, and by $\tilde{Z}$ the corresponding vector of elasticities from the quantified model represented by (18), (19), (22) and (23). The vector of parameter estimates $\hat{\Theta}$ is the solution to the minimization of the squared residuals

$$\hat{\Theta} = \arg\min_{\Theta} \quad \psi(\Theta)' \, \psi(\Theta) \, , \tag{24}$$

where

$$\psi(\Theta) = \frac{\tilde{Z} - Z}{Z} \, ,$$

denotes the vector of residuals that corresponds to the relative mismatch between the model elasticities and the empirical targets.

Table H.5: Raw Data Moments

	(1) Baseline "Country 1" (mean)	(2) Alternative "Country 2" (mean)
Percentage with at least secondary education	58.45192	82.25101
Percentage with Secondary Schooling, Female and Male (25+)	42.2795	56.71792
Percentage with Primary Schooling, Female and Male (25+)	27.3666	14.03117
Percentage with Tertiary Schooling, Female and Male (25+)	16.17243	25.53309
Avg. years of education	8.228714	11.00282
Capital Stock per capita	34115.12	110302.4
GDP p.c. (2010, WDI)	13327.11	32866.38
GDP p.c. (in const. 2010 USD, Maddison)	9434.517	34778.07
GDP p.c. (in const. 2010 USD, PWT)	9963.385	29704.52
Observations	36	16

The table provides average values for various moments for two groups of countries in the data. Baseline countries correspond to countries with an average patience level within the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentile of the cross-country distribution of patience. The alternative group of countries corresponds to countries with an average patience level within the 75$^{\text{th}}$ and 90$^{\text{th}}$ percentile of the cross-country distribution.