Patience and Comparative Development

Uwe Sunde (LMU Munich)
Thomas Dohmen (University of Bonn)
Benjamin Enke (Harvard University)
Armin Falk (briq and University of Bonn)
David Huffmann (University of Pittsburgh)
Gerrit Meyerheim (LMU Munich)

Discussion Paper No. 292

November 11, 2021

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Patience and Comparative Development*

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

March 11, 2021

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 per capita income and the accumulation of physical capital, human capital and productivity. These correlations hold across countries, subnational regions, and individuals. Second, the 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, aggregate production is characterized by capital-skill complementarities, and productivity implicitly depends on patience through a human capital externality. In our model estimations, general equilibrium effects alone account for a non-trivial share of the observed amplification effects, and an extension to human capital externalities can quantitatively match the empirical evidence.

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. 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) through 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 production factors and productivity play an important role in explaining international income differences (Hall and Jones, 1999; Caselli, 2005; Hsieh and Klenow, 2010). 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 ultimately arise 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). Perhaps due to a previous lack of reliable and comparable data on time preference on a global scale, however, the relationship between patience and comparative development is not well-explored.

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 level and macro level.

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 is strongly positively correlated with income and 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, different measures of 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, educational attainment, and standardized achievement test scores. 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 or altruism) 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.

Finally, we present conceptually analogous analyses across individuals, holding fixed the country or subnational region of residence. Here, again, patience is robustly correlated with higher household income, a greater propensity to save, and higher educational attainment. Taken together, our analyses show 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 are drivers of the correlation between patience and income.

A salient finding that emerges from the analysis at different levels of aggregation 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.

Most likely, *some* fraction of the differences in coefficient estimates across levels of aggregation are driven by measurement error and resulting attenuation bias. After all, across-individual and across-region variation in patience is likely measured with more error than cross-country patience. At the same time, our data also strongly suggest that attenuation alone is very unlikely to generate the observed aggregation 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 effect but cannot be explained by greater measurement error since all individual-level regressions (with or without country fixed effects) rely on the same individual-level data. This suggests that the amplification effects reflect an economic mechanism rather than a statistical artifact.

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, 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 differences in coefficient estimates across levels of aggregation can plausibly be generated by the model. 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 standard parameters based on estimates from the literature. We then estimate the remaining structural parameters (for which no agreed-upon estimates exist) using an indirect inference approach. We implement these estimations by targeting as estimation moments the empirical patience elasticities that we observe in our regressions at different levels of aggregation.

In the baseline version of the model, 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. Thus, potential amplification effects only arise as a result of price effects in general equilibrium. A helpful way to think about this model variant is that it corresponds to the empirical estimates across subnational regions, where human and physical capital may vary but the broader productivity environment (institutions, national policies etc.) is largely kept fixed.

Estimation of this baseline model delivers sensible parameter values. For example, we estimate average annual discount factors of 0.93 – 0.95. When we simulate the baseline model using the estimated parameter values, the implied patience elasticity is about twice as large at the aggregate relative to the individual level. This shows that in the model general equilibrium effects alone can lead to substantial amplification. The magnitude of these simulated amplification effects resonates with the empirical amplification effects observed going from individual- to regional-level estimates. While this amplification effect is substantial, however, it is not large enough to entirely account for the empirically-observed amplification in cross-country regressions.

Thus, in a second step, we estimate model variants in which productivity is allowed to vary, and implicitly depends on patience through a human capital externality. We think of these specifications as mirroring our cross-country regressions, in which the broader productivity environment also varies. In these analyses, we find that the amplification of the elasticity of income and skill shares with respect to patience increases substantially, and comes close to matching the empirically-observed patterns. Through a series of sensitivity checks, we document that the magnitude of amplification effects is largely governed by (i) the magnitude of capital-skill complementarities and (ii) the size of human capital externalities. Taken together, the model offers an internally consistent way to think about the empirical results, tying together the correlations between patience and economic outcomes across levels of aggregation, while simultaneously shedding light on the substantial amplification effects. Moreover, the estimation results clarify that – in the context of our model – the empirically-observed variation in patience can rationalize the observed development differences and amplification effects.

This paper contributes to two lines of research in the literature on comparative development. The first, using development accounting, decomposes national income into production factors and productivity (the proximate determinants of development). The second involves research on the deep determinants of development and 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 related to a cultural factor that can generate variation in the proximate determinants of development (e.g., Doepke and Zilibotti, 2008, 2014). 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 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 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 results also complement recent work that studies the intergenerational transmission and evolution of patience in response to economic incentives, and the overall economic environment, in a setting where patience determines human capital investment (Doepke and Zilibotti, 2018).

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. The GPS contains a set of survey items that were explicitly designed to measure a respondent's time preferences, risk preferences, social preferences, and trust, that were part of the regular 2012 questionnaire of the Gallup World Poll (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 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

[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.

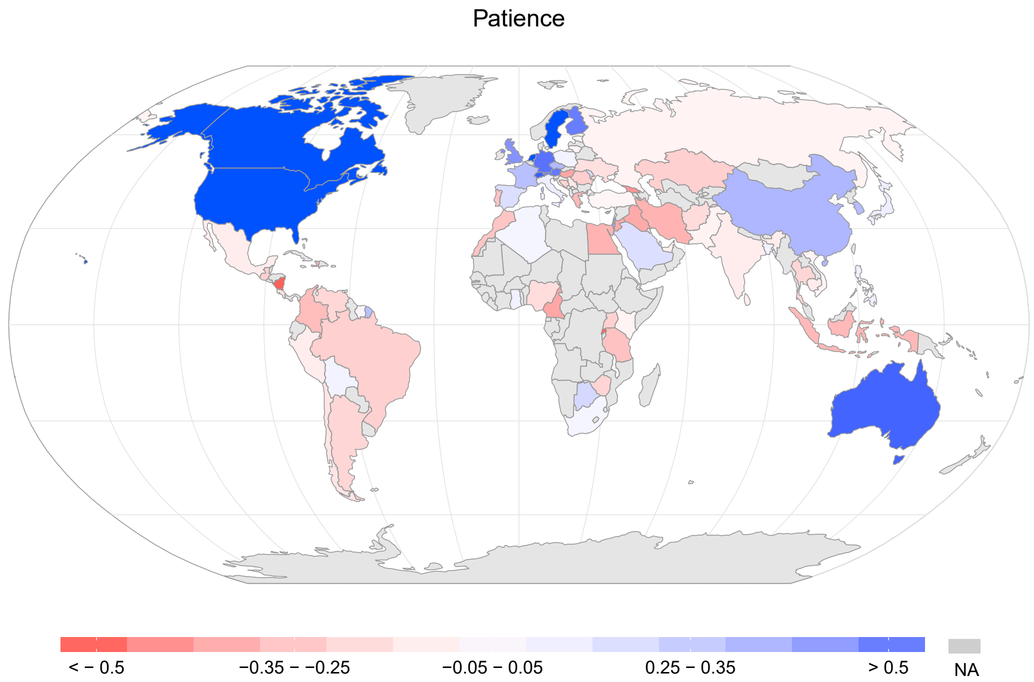


Figure 1: Distribution of patience across countries

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

[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 $\rho = 0.23$, respectively. Moreover, as shown below, the well-known correlation between trust and per capita income vanishes once patience is controlled for.

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. 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, and patience alone statistically accounts for 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, including controls for world regions, absolute latitude, longitude, the fraction of arable land, land suitability for agriculture, average precipitation and temperature as well as the fractions of the population that live in the (sub-) tropics or in areas where there exists the risk of contracting malaria.[4] Finally, column (5) additionally controls for genetic diversity and its square, and trust. 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), has little explanatory power once patience is included in the analysis. Figure 2 illustrates the conditional relationship for the estimates in column (5).

Robustness checks. Appendix D presents two sets of robustness checks. First, the results are robust to additionally controlling for average risk aversion, other geographical variables, linguistic, religious, and ethnic fractionalization, legal origin dummies, major religion shares, the fraction of European descent, and the genetic distance to the U.S. (Table D.1). Second, the relationship between patience and per capita income robustly appears in various sub-samples, e.g., within each world region, within OECD or non-OECD

[3]The coefficient estimate in column (1) differs slightly 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. * $(p < 0.10)$, ** $(p < 0.05)$, *** $(p < 0.01)$

countries, or within former colonies and countries that have never been colonized (Table D.2).

Growth extension. Appendix D also 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 when we control for per capita income in the base year and additional covariates (Table D.3).

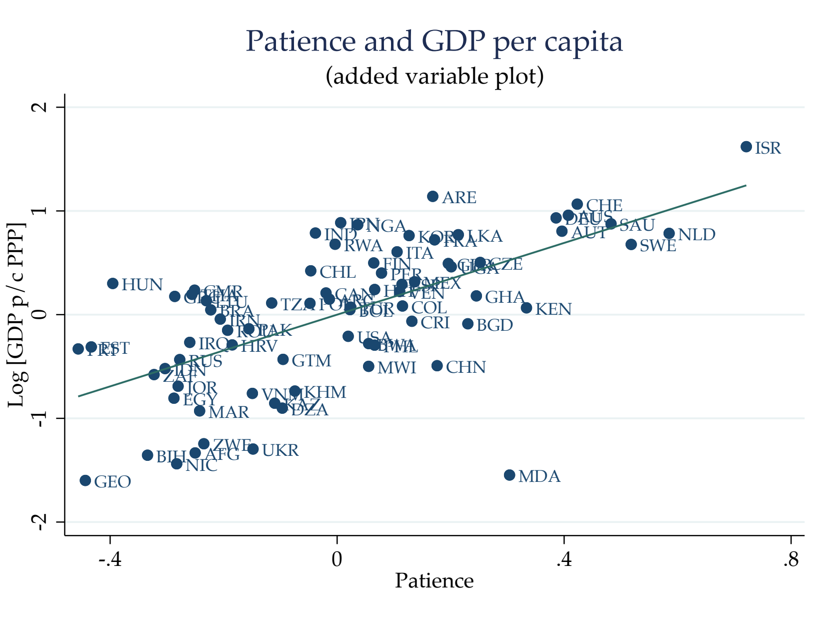


Figure 2: Patience and national income (added variable plot conditional on the full set of covariates in column (5) of Table 1).

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 corresponding 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. The data on household savings rates are based on surveys and are only available for OECD countries. Throughout, the results reveal a significant positive relationship between patience and

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. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

savings. The finding that variation in patience is related to cross-country variation in household savings rates even within OECD countries is noteworthy, given the similarity of this subset of countries in terms of economic development and other characteristics.

Human Capital. As baseline measures of human capital, we consider conventional quantitative measures 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 these variables.[5]

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. Columns (5) – (8) in Table 3 document that patience is strongly correlated with both the TFP measure from the PWT and the number of researchers in research and development. 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

[5]Comparable results are obtained with alternative measures of human capital, such as the fraction of the population aged over 25 that has obtained tertiary education, or a measure of the quality of human capital as reflected by a measure of standardized math and science test scores (Hanushek and Woessmann, 2012), see Table D.4 in Appendix D.

Table 3: Patience, human capital and productivity

	Human capital				Productivity			
	Dependent variable:							
	% Skilled		Yrs. of schooling		TFP		Log [# Researchers in R&D]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience	38.5***	20.1***	4.34***	2.47***	0.29***	0.17**	2.70***	1.49***
	(5.45)	(7.20)	(0.58)	(0.86)	(0.05)	(0.07)	(0.35)	(0.50)
Continent FE	No	Yes	No	Yes	No	Yes	No	Yes
Additional controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	72	71	72	71	59	58	69	68
R^2	0.30	0.73	0.34	0.76	0.29	0.70	0.35	0.83

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). Number of researchers in R & D are per $1,000$ population. Columns $(5)-(6)$ 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. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

the external environment such as institutions, inflation, or interest rates; and (ii) the OLS correlations could 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 analysis as a first contribution that studies the systematic relationship between patience, accumulation and income, and that 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 expect higher incomes in the future 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 4, 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 4.

Subjective Uncertainty. In the quantitative decision tasks between payment today and in twelve months, respondents may face subjective uncertainty about whether they would actually receive the (hypothetical) payment 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 4 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 4 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, 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 4: 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***	0.77***	1.52***	1.04***	1.17***	1.37***	
		(0.33)	(0.27)	(0.41)	(0.24)	(0.24)	(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. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

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 (see right panel of Figure D.1 in Appendix D). While these results do not necessarily rule out reverse causality from 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 5 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

Table 5: Other preference measures

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

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

preferences and trust. Other measures are only inconsistently related with outcomes (see 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.[6]

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.

[6]See Appendix C.1 for an overview of the number of regions per country.

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.[7]

To provide some perspective on the variation in average regional patience, we discuss a few summary statistics. Recall that individual patience is standardized to have mean zero and standard deviation one. Average regional patience has a standard deviation of $\sigma = 0.45$ (average country patience has standard deviation $\sigma = 0.37$). Moreover, only 72% of the variation in regional patience is explained by country fixed effects. This suggests that our data exhibit sufficient within-country variation to meaningfully explore the link between regional patience and regional development.

Table 6 reports regression results for 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 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, a noteworthy observation is the change in 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)). Moreover, the across-region coefficient estimates are substantially smaller than the corresponding across-country estimates reported in Table 1 and Table 3. 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 7 presents the results of OLS regressions with three dependent variables: log household income per capita, a binary indicator for

[7]For details see Appendix C.2.

Table 6: 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 the country level) in parentheses. Patience is shrunk patience, see equation (C.1) in Appendix C.2. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

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.[8] This pattern holds conditional on a comprehensive vector of individual-level covariates including religion fixed effects, age, age squared, gender, 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 B).

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.6 in Appendix D additionally controls for this binary indicator, with very similar results as in Table 7.

Moving beyond the qualitative patterns, we again see that the coefficient estimate of patience drops by a factor of six in the income regressions 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.

[8]Comparable results are obtained with a more restrictive definition of being skilled, or for subjective measure related to the quality of human capital in terms of math skills (see Table D.5 in Appendix D).

Table 7: 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				0.58***				-0.059				0.20
				(0.20)				(0.32)				(0.24)
Age squared				-0.38				-0.056				-0.94***
				(0.23)				(0.30)				(0.22)
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
Subnational region 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 the 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 B for a detailed description of all dependent variables. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

3.4 Potential Statistical Reasons for Amplification

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 substantially (roughly by a factor of seven) once country fixed effects are included. This result is not due to the use of different specifications or data sources at different levels of aggregation. In fact, very similar aggregation effects emerge when we use the GPS data on patience and income and aggregate them up to the regional or country level.[9]

From an empirical perspective, the two obvious candidate explanations for the differences in the estimates across different levels of aggregation are measurement error and omitted variables at the aggregate level that correlate with average patience. In the following, we provide a brief discussion of both.

Measurement error constitutes a potential explanation for the large variation in coefficient estimates across levels of aggregation due to attenuation bias. In particular, the relationship between individual income and patience should be more attenuated than the country-level relationship if individual patience is measured with more noise than country-level patience. This is likely the case as measurement error washes out when aggregating patience at the country level. 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, part of the difference in patience coefficients between country-, regions- and individual-level analysis is likely to be due to measurement error.

At the same time, two pieces of evidence strongly indicate that measurement error alone is unlikely to generate the observed aggregation effects. First, an argument that is based on measurement error cannot explain why – within individual-level or region-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 level of aggregation (either individual or region). Instead, it appears that moving from a cross-country to a purely subnational comparison per se reduces the magnitude of the patience coefficient.[10]

[9] See Table D.7 in Appendix D.

[10] 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 use of different patience measures and income variables, the few available benchmarks reveal encouraging similarities. In the nationally representative German sample of Dohmen et al. (2010), the corresponding coefficient of individual patience in a regression with log income per capita 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.

A second piece of evidence against a pure measurement error explanation is the required magnitude of noise. 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 $p_o = p_t + \alpha \cdot \epsilon$, 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 E, 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 $\rho = 0.02$.[11] 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 amplification of effects.[12]

The other candidate explanation for the differences in the estimates across different levels of aggregation is omitted variables in the form of correlated aggregate effects as the result of equilibrium interactions or other externalities. In particular, abstracting from measurement error, it is known that, in the presence of omitted variables that correlate with patience at the country level, the country-level estimates and the (within-country) individual-level estimates of the patience elasticity estimate different parameters, since the country-level estimate also contains the correlated group effects (which could reflect, e.g., equilibrium effects or externalities). To see this, suppose that the structural model underlying the data is the same for all individuals and consider a model for the relationship between an outcome variable, e.g., income, y, and patience p for individuals i in countries j, with latent country-specific effects κ_j and a homogeneous slope, $y_{ij} = a_0 + a_1 p_{ij} + \kappa_j + u_{ij}$. Following Mundlak (1978) and Pakes (1983), let the latent effect be a function of average patience, with $\mathbf{E}[\kappa_j | \bar{p}_{.j}] = \bar{p}'_{.j}\gamma$. Then $y_{ij} = p'_{ij}a_1 + \bar{p}'_{.j}\gamma + e_{ij}$, and taking country means gives $\bar{Y}_{.j} = \bar{P}'_{.j}(a_1 + \gamma)$. Thus, estimates at the country level also reflect equilibrium effects or other externalities in addition to the individual-level relationship. Considering the possibility that patience is measured with error implies an observationally similar difference in the individual and aggregate estimates and it is not possible to disentangle both mechanisms in general. However, the availability of an intermediate level of aggregation,

[11] 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.

[12] An additional measurement-related issue that could 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 E.2 discusses this issue in detail.

as in our case sub-national regions, provides further insights as to whether the observed amplification effect is due to omitted variables at the aggregate level or due to measurement error (for details, see Pakes, 1983). A comparison of the estimates at the levels of individuals, regions, and countries reveals that correlated country-level effects, and to a lesser extent correlated region-effects, are likely candidates for explaining the amplification effect, in contrast to measurement error.[13]

In the next section, we consider an model that features heterogeneity in patience across individuals and across countries. This model illustrates a potential economic (rather than statistical) mechanism behind aggregation effects by rationalizing the existence of correlated group effects as consequence of general equilibrium mechanisms and externalities.

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 potential human capital externalities on productivity (Lucas, 1988).[14] 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 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

[13] Suppose that the correlated group effects and measurement error are mutually uncorrelated, then it can be shown that the difference between country-level estimates and estimates between regions within countries (region-level estimates conditional on country fixed effects) identifies the correlated effect of variation in country-level patience (for details, see Pakes, 1983). For instance, considering the results in Table D.7, the estimate for this effect corresponds to $2.14 - 0.15 = 1.99$. This is similar to the effect obtained under the assumption of no measurement error, which corresponds to the differences between country-level estimates and estimates between individuals within countries (individual-level estimates conditional on country fixed effects) of $2.14 - 0.056 = 2.084$.

[14] 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.

of individual i are represented by

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

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 devoting a fraction $(1-\psi)$ of the first period of life to skill acquisition. 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 individual human capital corresponding to $h = e^{\rho(1-\psi)}$, where $\rho > 0$ is the 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 $w^H h$, 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.[15] The respective budget constraints are then

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

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

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 (1) subject to (2). Analogously, the optimal savings decision for individual i conditional on becoming a skilled worker is determined by maximizing (1) subject to (3). Solving the individual decision problems delivers the optimal savings rate as

$$s_{t+1}^L = s_{t+1}^H = \frac{\beta(i)}{1 + \beta(i)} \ , \tag{4}$$

[15] 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.

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 $\eta_{t+1} = \frac{w^H_{t+1}}{w^L_{t+1}}$, 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 is given by

$$\eta_{t+1} > \eta(i) = \frac{\psi^{\frac{-1}{\beta(i)(1+\beta(i))}}}{h} , \tag{5}$$

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 $\eta_{t+1} = \frac{w^H_{t+1}}{w^L_{t+1}}$, condition (5) therefore implicitly determines a threshold level for patience, $\tilde{\beta}_t$, that determines the population share of skilled individuals.[16]

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 (4). Likewise, more patient individuals have a higher propensity to become skilled due to (5). 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

[16] The simple representation of the individual education decision as a binary choice problem illustrates the main mechanism while keeping the model tractable. In order to account for other unobserved and idiosyncratic factors that might influence education choices in reality and thus the patience elasticity, the decision to become skilled as determined by (5) could be extended by incorporating idiosyncratic heterogeneity that is orthogonal to the mechanism related to patience. In the empirical implementation below this is done by including a symmetrically distributed additive noise term with mean zero. The main empirical predictions regarding the role of patience for individual decisions remain unaffected by this since the effects of idiosyncratic heterogeneity wash out on average.

evidence regarding capital-skill complementarities (Duffy et al., 2004), we assume that the production function takes the form

$$Y_t = A_t \left[\left(K_t^{\theta} + H_t^{\theta} \right)^{\frac{\sigma}{\theta}} + L_t^{\sigma} \right]^{\frac{1}{\sigma}} \ , \tag{6}$$

with the aggregate capital stock in period t 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).[17] 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 competitive 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

$$L_t = \frac{1}{2\varepsilon} \left[\int_{\chi-\varepsilon}^{\tilde{\beta}_{t-1}} 1 \, d\beta + \int_{\chi-\varepsilon}^{\tilde{\beta}_t} 1 \, d\beta + \int_{\tilde{\beta}_t}^{\chi+\varepsilon} \psi \, d\beta \right] \ , \tag{7}$$

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

$$H_t = \frac{1}{2\varepsilon} \int_{\tilde{\beta}_{t-1}}^{\chi+\varepsilon} e^{\rho(1-\psi)} \, d\beta \ . \tag{8}$$

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

$$K_{t+1} = \frac{1}{2\varepsilon} \left[\int_{\chi-\varepsilon}^{\tilde{\beta}_{t-1}} \frac{\beta(i)}{1+\beta(i)} (w_t^L \cdot 1) \, d\beta(i) + \int_{\tilde{\beta}_{t-1}}^{\chi+\varepsilon} \frac{\beta(i)}{1+\beta(i)} (w_t^H \cdot h) \, d\beta(i) \right] \ . \tag{9}$$

[17] We abstract from the consideration of different types of capital in terms of equipment and structures, as in Krusell et al. (2000).

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

$$A_t = \bar{A} \cdot H_t^{\gamma} \; , \tag{10}$$

where $\bar{A}$ captures heterogeneity in productivity that is orthogonal to accumulated factors (in the sense of a Solow residual) and $\gamma \geq 0$.

Equilibrium. The remaining analysis focuses on the steady-state equilibrium. 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.[18] The key condition for equilibrium is the consistency of the indifference condition for education (5) with the earnings premium that emerges from the relative supply of skilled labor, and the corresponding capital supply and demand.[19]

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 (5) 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 $\lambda = \frac{\chi+\varepsilon-\tilde{\beta}}{2\varepsilon} \Leftrightarrow \tilde{\beta} = \chi + \varepsilon - 2\varepsilon\lambda$.[20]

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 $\frac{d\lambda}{d\chi} = \frac{1}{2\varepsilon}\left(1 - \frac{d\tilde{\beta}}{d\chi}\right) > 0$.[21] These conditions also imply that the threshold in terms of individual patience for becoming skilled is higher in a country with a higher average level of patience.

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

[19] The consideration of orthogonal idiosyncratic heterogeneity implies a negative second-order effect on aggregate savings but does not affect the existence and uniqueness of equilibrium and thus only has a minor quantitative effect on aggregation.

[20] 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 $\bar{\beta} = \chi + \varepsilon(1 - \lambda)$.

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

5 Bringing the Model to the Data

5.1 Testable Implications

Approach. We are interested in evaluating the effect of an increase in individual or average patience on education and income, in particular how this effect varies with the level of aggregation. To construct model analogues for the regression results, the model analyzes two thought experiments. At the individual level, the thought experiment is to determine the cross-sectional average elasticity of income and education with respect to a change in individual patience $\beta(i)$, 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 (or subnational region) fixed effects, where the 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 an increase in patience 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 a baseline country (country 1) to that in a second country with $\chi_2 = \chi_1 + sd > \chi_1$, all else equal. Note how this thought experiment corresponds to the empirical analyses above, in which the OLS coefficients are estimated using a patience variable that has standard deviation one.

Education. First, consider the effect of patience on an individual's decision to become skilled. It is clear from (5) that the propensity to become skilled can be expressed as a binary choice problem with the compensation that an individual requires for becoming skilled, $\eta(i)$, as latent variable. If the market compensation, η^*, is greater than this minimum compensation, the individual becomes skilled. In reality, other unobserved and idiosyncratic factors beyond patience influence the education choice. Therefore, we represent the empirical analogue of the decision to become skilled as a linear probability model in which the decision to become skilled is determined by (5) with an additive noise term u that is symmetrically distributed around zero, $\mathcal{I}_{\text{skilled}}(i) = \mathbb{1}\{\eta^* - \eta(\beta(i)) + u(i) > 0\}$.[22] Consequently, the marginal effect of an increase in patience on the propensity to become skilled is given by

[22] See Appendix G for details.

$$\frac{\partial \mathcal{I}_{\text{skilled}}(i)}{\partial \beta(i)} = \frac{1}{2\varepsilon} \cdot \left| \frac{d\,\beta(i)}{d\,\eta(i)} \right| . \tag{11}$$

Notice that the effect of patience on the individual propensity to become skilled depends on the level of patience. The empirical estimate of the elasticity of individual education with respect to patience (Table 7) corresponds to the population average of a linearized version of this marginal effect. In our quantitative model analysis, we evaluate this expression 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

$$\frac{d\,\lambda}{d\,\chi} = \frac{1}{2\varepsilon} \left(1 - \frac{d\,\tilde{\beta}}{d\,\chi} \right) > 0 . \tag{12}$$

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

We are interested in whether this aggregate elasticity is larger than the corresponding individual-level elasticity. 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. 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 a 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 certain conditions (see Appendix G for details).

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 (4), the average marginal effect of an increase in patience on individual savings is given by

$$\frac{\partial \bar{S}}{\partial \beta(i)} = \frac{1}{1+\tilde{\beta}} \left[\frac{(1-\lambda) \cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda \cdot w^H h}{1+\chi+\varepsilon} \right] , \tag{13}$$

where $\bar{S}$ denotes average individual savings.[23] This implies that the average effect of

[23]See Appendix G for details.

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 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 aggregate capital is given by

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

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 $\sigma > \theta > 0$), 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 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

$$\frac{\partial \bar{y}}{\partial \beta(i)} = \frac{\partial \bar{\mathcal{I}}_{\text{skilled}}}{\partial \beta(i)} \left[w^H h - (2-\psi) w^L\right] + R \cdot \frac{\partial \bar{S}}{\partial \beta(i)} , \tag{14}$$

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 and the corresponding prices.

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

$$\frac{dY}{d\chi} = \underbrace{\frac{d\lambda}{d\chi}\left[w^H h - (2-\psi)w^L\right] + R\cdot\frac{dK}{d\chi}}_{\text{direct effects}} + \underbrace{\left[2(1-\lambda)+\psi\right]\frac{dw^L}{d\chi} + \lambda\frac{dw^H h}{d\chi} + K\cdot\frac{dR}{d\chi}}_{\text{general equilibrium effects}} . \quad (15)$$

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 at 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.[24]

5.2 Parameter Calibration and Estimation Approach

We use a combination of model calibration and estimation to quantify the model. The baseline model contains eight parameters. The extension with a human capital externality involves an additional parameter. We calibrate parameters that are standard in macro-models using conventional estimates from the literature and estimate the remaining parameters as described below. In particular, we calibrate the CES elasticities σ and θ based on empirical estimates by Duffy et al. (2004).[25] We set the time requirement for becoming a skilled worker in terms of the fraction of the first period of life to $(1-\psi) = 0.2$, which is equivalent to five years with the length of a generation being 25 years. Finally, we assume a Mincerian return of 7%, which is in line with empirical estimates (e.g., Acemoglu and Angrist, 2000; Card, 2001; Belzil and Hansen, 2002; Psacharopoulos and Patrinos, 2018).[26] More precisely, given an average return of 7% over five additional years, this implies for the model $e^{0.07\cdot 5} = e^{\rho(1-\psi)}$. Inserting the calibrated value of ψ and solving for ρ yields $\rho = 1.75$. Table 8 summarizes the calibration of these parameters.

The remaining parameters include the distributional parameters of the patience distribution, χ and ε, and the level of the Solow residual, $\bar{A}$. The extension of the model to a human capital externality invokes an additional parameter, γ. These parameters are either model-specific or no commonly agreed estimates exist that can be used for calibration. For instance, the literature has not settled on how large human capital externalities are in the

[24] See Appendix G for details.

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

[26] Regressing household income on years of schooling in our global individual-level data delivers an average Mincerian return of approximately 6.5%. However, this estimate has to be interpreted with caution because of the income measure and potential measurement error in the Gallup data.

Table 8: Calibrated parameters

Parameter	Value	Calibration Details
$1-\psi$	0.2	Fraction of young age required to become skilled (five additional years) (Caselli, 2017)
ρ	1.75^a	Corresponds to a (private) Mincerian return of 7% (Card, 2001; Psacharopoulos and Patrinos, 2018)
σ	0.62	CES (inverse): labor/capital compound (Duffy et al., 2004)
θ	0.38	CES (inverse): physical/human capital (Duffy et al., 2004)

Calibrated parameters. [a] With the Mincerian human capital production function, a return of $x = 0.07$ for five years of schooling during a 25-year period of youth corresponds to $\rho = \frac{\ln\left(e^{0.07 \cdot 5}\right)}{0.2} = 1.75$.

data.[27] Consequently, we estimate these parameters using an indirect inference approach that allows us to estimate the remaining free model parameters by matching the patience elasticities of the variables of interest in the model to those obtained from the regressions in Section 3 above. In particular, the parameters are estimated by matching as empirical moments the patience elasticities of education at the individual level and aggregate level, as well as the patience elasticities of income at the individual level and aggregate level.[28]

To keep this analysis directly comparable to the reduced-form patterns, the simulated individual 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 at the country level. As a result, the remaining parameters to be estimated in the baseline estimation are given by the vector $\Theta = (\chi_1, \varepsilon, \bar{A})$, as χ_2 is implicitly determined by

$$\Delta\chi = \chi_2 - \chi_1 = sd = \frac{2\varepsilon}{\sqrt{12}} \; .$$

Unless noted otherwise, $\bar{A}$ is restricted to be the same across countries.

Estimation is based on a Wald-type minimization of the vector of quadratic differences of the standardized elasticities. Denote by Z the vector of elasticities obtained from reduced form regressions, and by $\tilde{Z}(\Theta)$ the corresponding vector of elasticities from the quantified model. The vector of parameter estimates $\hat{\Theta}$ is the solution to the minimization of the squared residuals

$$\hat{\Theta} = \arg\min_{\Theta} \quad \vartheta(\Theta)' \, \vartheta(\Theta) \; , \tag{16}$$

[27] See, e.g., Acemoglu and Angrist (2000); Moretti (2004); Ciccone and Peri (2006); Acemoglu and Autor (2012); Thönnessen and Gundlach (2013); Psacharopoulos and Patrinos (2018).

[28] 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.

Table 9: Matched elasticities and targets

Effect of patience	Model Moment $\left[\tilde{Z}(\Theta)\right]$		Empirical Moment $[Z]$	Target Value
Individual level				
Income	$\frac{\partial \ln \bar{y}}{\partial \beta(i)}$	as in (14)	Table 7, col. (3)	0.05
Education	$\frac{\partial \mathcal{I}_{\text{skilled}}(i)}{\partial \beta(i)}$	as in (11)	Table 7, col. (11)	0.03
Country level				
Income	$\frac{Y_2 - Y_1}{Y_1}$	as in (15)	Table 1, col. (5)	1.73
Fraction skilled	$\lambda_2 - \lambda_1$	as in (12)	Table 3, col. (2)	0.20

The model moments for the elasticities at the country level are discretized versions of equations (15), and (12). For the fully parametric versions as implemented in the estimation, see (G.5), (G.1), (G.6), and (G.2) in Appendix G.

where

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

denotes the vector of residuals that corresponds to the relative mismatch between the model elasticities and the empirical targets. Table 9 provides an overview of the matched model quantities (elasticities) and corresponding empirical moments.

5.3 Model Specifications

To be able to shed light on the mechanisms behind the observed amplification effects in the data, we consider four variants of the model.[29]

Baseline. In the baseline version, we consider two model economies that only differ in their patience distribution, but without a human capital externality on TFP. Thus, patience only affects economic performance through the accumulation of physical capital and human capital. We think of this specification as conceptual analogue to the within-country-across-region regressions, reported in Section 3.2. Here, patience might affect the formation of physical and human capital, but the broader productivity environment is effectively held constant in these regressions. For example, national institutions, policies or the supply of educational resources plausibly affect the productivity environment, but are largely fixed when comparing subnational regions within the same country.

[29] We discuss additional robustness checks below in Section 5.5.

HC externality (estimated). To account for systematic differences in productivity across countries, which might also influence the accumulation of factors (Hsieh and Klenow, 2010), we then estimate an extended version of the model that accounts for the observed differences in TFP across countries. Specifically, we estimate the parameter that governs a potential human capital externality, γ while keeping $\bar{A}$ fixed across both economies. We think of this model variant as analogue of the cross-country regressions, where the broader productivity environment also varies, and where patience could implicitly affect the supply of national policies or other productivity shifters through the aggregate stock of human capital.

HC externality (calibrated). As we will see below, allowing for a human capital externality increases the amplification effects between individual-level and country-level analyses substantially. This raises the natural question how sensitive our results are to the magnitude of the human capital externality. Given that no widely agreed-upon magnitude for this externality exists, we check sensitivity by calibrating $\gamma = \frac{\rho}{2}$. We view this calibration exercise as a conservative approach that complements the less conservative approach of directly estimating γ in light of evidence that social returns to human capital are usually larger than private returns (Psacharopoulos and Patrinos, 2018).

Development accounting: TFP variable, patience fixed. A natural concern with our empirical analyses is the presence of omitted variables. In particular, it is conceivable that patience is strongly correlated with income (especially across countries) because variables that are typically summarized as contributing to TFP might covary with patience, such as institutions or the quality of national policies. If this was the case, the amplification effects documented above would partly reflect omitted variable bias. To assess the plausibility of such an account, we estimate a model variant in which average patience χ is held fixed across the two economies under consideration. Instead, in this model variant we estimate two separate levels for $\bar{A}_1$ and $\bar{A}_2$, as is commonly done in the development accounting literature. That is, in these estimations, any differences in aggregate outcomes are exclusively driven by exogenous differences in TFP. We will then conduct the following thought experiment: Suppose that both economies are equally patient (fixed χ), yet the high-TFP one *appears* more patient in the GPS survey data. Then, can the observed amplification effects (and outcome differences between seemingly patient and impatient countries) be rationalized as a result of TFP differences? By addressing this thought experiment, the analysis will shed light on two aspects: whether exogenous variation in TFP alone is sufficient to rationalize the patterns in the data, and the potential value added of a patience-related amplification mechanism.

Table 10: Estimated parameters

Model	Baseline	Extensions		Dev. Acc.
		HC-Externality (estimated)	HC-Externality (calibrated)	
χ_1, χ_2	0.16, $[0.25]^a$	0.16, $[0.25]^a$	0.12, $[0.19]^a$	0.16, $[0.16]^c$
ϵ	0.16	0.16	0.12	0.16^c
$\frac{\bar{A}_1}{\bar{A}_2}$	1	1	1	2.55
γ	0	1.98	0.88^b	0^c

Parameters in brackets [ ] are derived from estimated parameters. [a] Level of χ_2 as implied by the estimated values of χ_1 and ε. [b] Calibrated to $0.88 = \frac{\rho}{2}$. [c] Values fixed as in baseline model.

5.4 Estimation Results

Amplification effects. Table 10 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 (a discount rate of 5 – 7%).

Table 11 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. We begin 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 11. The individual-level elasticities of income and education with respect to variation in patience obtained with the model closely resemble the empirical estimates, as shown in the upper panel. The bottom panel of the table shows the corresponding elasticities for variation in patience across economies. The baseline version of the model delivers a moderate amplification of the elasticity in income by a factor of about two. Interestingly, this magnitude of amplification corresponds to the patterns observed in the reduced-form regressions across subnational regions in column (3) of Table 6. The fact that the observed amplification is much larger at the country level – and that this cannot be reproduced by our baseline specification – suggests a potential role for TFP differences. 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.

Table 11: Quantified model vs. data

	Effect of one SD increase of patience				Fixed χ
	Data	Model			Model
	(Baseline Controls)	Baseline	Extensions		Dev. Acc.
			HC Externality (estimated)	HC Externality (calibrated)	
	(1)	(2)	(3)	(4)	(5)
			Individual level		
Income	0.05	0.05	0.05	0.05	0.05^a
Education	0.03	0.03	0.03	0.02	0.03^a
			Country level		
Income	1.73	0.11	1.74	0.80	1.79^b
Fraction skilled	0.20	0.16	0.18	0.16	0.02^b

The effect sizes in the simulated model are obtained after estimating the parameters through indirect inference as reported in Table 10. In the baseline, estimated parameters are χ_1, ε, and $\bar{A}$ by matching as moments the effects of patience on individual income, individual propensity to become skilled, aggregate income and aggregate skill share. For details on the target moments from the data see Table 9. ^{a} Effect of one standard deviation increase in individual patience β. ^{b} Effects of comparing across two countries with identical patience distributions, but with different levels of $\bar{A}$.

To account for these national productivity factors in a parsimonious way, we then estimate the model variant in which γ (the externality) is a free parameter to be estimated. In Table 10, the estimation yields $\gamma = 1.98$, which is slightly larger than the value of the private return to human capital ($\rho = 1.75$) that is implied by a Mincerian return of 7%. Column (3) of Table 11 presents the results on amplification. In this version of the model, the individual-level patience elasticities are matched, and the model also delivers a large amplification of elasticities at the aggregate level that closely matches the data.

Given the strong increase in the observed amplification effect as a result of allowing for a human capital externality, we investigate the sensitivity of our results by calibrating $\gamma = \frac{\rho}{2}$. In column (4) of Table 11, we see that this version of the model again matches the empirical individual-level elasticities well. In addition, the elasticity of the fraction skilled with respect to patience is also matched closely. Regarding income, the observed amplification is now by a factor of 16. This is about half as much as with an estimated human capital externality, but nevertheless substantial in magnitude.

Finally, we present the results for the development accounting scenario in which average patience is fixed across economies, yet TFP levels are estimated for both countries. Column (5) of Table 11 presents the corresponding results. For the patience elasticity of income at the individual level, the patterns are similar to the other versions of the model. At the country level, we see that this version of the model does a very good job at matching the income difference between the high and low TFP countries (which we here interpret as "seemingly patient and impatient countries in the GPS"). At the same time, the difference in share skilled between the high and low TFP country are much smaller than the ones observed in the data, and also much smaller than the ones generated by our model variants that feature variable country-level patience χ.

Non-Targeted Moments. To further assess model performance and the plausibility of the estimated parameters, we also compared other, non-targeted data moments obtained from reduced form estimates or raw data to the corresponding moments obtained from the model estimation. These moments include elasticities of physical and human capital accumulation with respect to patience that 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, we consider other moments that are relevant from the perspective of comparative development. The results reveal that, by and large, the moments implied by the estimates of the different specifications of the model resemble the empirical moments, where the model extension with a human capital externality on TFP again provides the best overall fit; in comparison, the model variant with fixed patience but variable (exogenous) TFP fits the data rather poorly (see Table H.8 in Appendix H for details).

Sensitivity analyses: capital-skill complementarity. To assess the sensitivity of the results with respect to the magnitude of the capital-skill complementarity, we present a modification of our baseline model (without human capital externality) in which we do not calibrate the CES parameters σ and θ, but instead estimate them. The results of these estimations are similar to the baseline results (see Tables H.9, H.10 and H.11 in Appendix H for details). If anything, this exercise delivers an even stronger capital-skill complementarity while improving the fit moderately. In a second sensitivity check, we estimate a model version in which the complementarity is calibrated to be considerably smaller than in our baseline specification ($\sigma = 0.51$ and $\theta = 0.49$). This reduces the observed income amplification relative to the baseline (from a factor 2.2 to 1.8 for income).

Sensitivity analyses: human capital. Appendix H shows that our results are robust to allowing for both a human capital externality and TFP differences that are unrelated to human capital. The estimation of the individual return to education, ρ, delivers a slightly more pronounced amplification of the aggregate patience elasticities, but an otherwise fairly similar performance as the baseline model. Finally, we also estimated a version of the model that focuses on the role of upper-tail human capital. This version is motivated by arguments that the social returns to education are plausibly larger than is commonly estimated in cross-sectional data because the latter ignore the “level” effect that results from having highly skilled workers or entrepreneurs who run more productive firms and thereby increase the productivity of the entire workforce (Gennaioli et al., 2013). When incorporating a reduced-form version of this perspective, the model delivers a similar amplification and performance. See Appendix H for details.

5.5 Discussion

Overall, the estimations yield three findings. First, even without variation in TFP, the model generates a non-trivial amplification effect in coefficient estimates going from the individual level to the aggregate level. This amplification is comparable to the amplification observed in the data when comparing individual-level and regional-level results.

Second, once variation in TFP is incorporated, the model predictions get closer to the coefficients obtained in cross-country analyses. This is consistent with productivity differences (such as national policies or supply of schooling) that are endogenous to average patience contributing to the observed amplification patterns.[30] For example, if patient populations opted for institutions designed to foster long-term growth as opposed to short-term rent extraction, these institutions may entail additional positive effects on factor accumulation and income.

Third, an alternative model in which the true underlying variation is not in patience but instead in TFP provides a less convincing model fit. In particular, in the model simulations, exogenous differences in TFP are unable to quantitatively match the observed variation in skill shares. In contrast, the assumed exogenous variation in patience induces large variation in both income and skill shares. These results speak to the literature on development accounting. 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 neo-classical growth model typically requires large TFP differences between countries to account for observed 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.[31] Our approach complements these contributions by highlighting the potential role of patience rather than education (which is an endogenous object) itself.

[30] 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).

[31] 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 differences in human capital substantially amplify TFP differences across countries. Schoellman (2012) makes a related point based on a novel methodology designed to measure differences in human capital quality.

6 Concluding Remarks

In this paper, we have documented two sets of stylized facts. First, across levels of aggregation, differences in income as well as the accumulation of human capital, physical capital, and the stock of knowledge are systematically linked to variation in patience. Second, the data reveal strong aggregation effects with respect to patience. The analysis of a stylized general equilibrium model that allows for heterogeneity in patience within and across countries has shown that both patterns are consistent with economic theories of intertemporal choice. The results from a quantitative analysis of our model are consistent with the idea that the difference in magnitude of coefficients across levels of aggregation is partly driven by general equilibrium effects and human capital externalities.

We highlight three broad avenues for future research. First, 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) rule out categorically that heterogeneity in patience reflects general circumstances 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, 2016), 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).

A second open question concerns the scope of the amplification mechanism at the regional level. Our main argument in the model estimation section was that aggregate productivity is largely held constant in across-region comparisons, so that – from the perspective of the model – potential amplification effects in cross-regional regressions reflect price effects. At the same time, the magnitude of cross-region differences in TFP and its link to patience is still an open question.

Third, while prior micro studies have focused on linking patience to human capital and physical capital accumulation, less is known about the effects that variation in patience might exert on productivity differences. This question seems particularly relevant from the perspective of our model estimations, in which the empirically-observed amplification patterns can only be explained in the presence of human capital externalities on productivity.

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, 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, "Ancient Origins of the Global Variation in 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.

Card, David, "Estimating the return to schooling: Progress on some persistent econometric problems," *Econometrica*, 2001, *69* (5), 1127–1160.

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.

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," *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.

_ **and** _ , "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission," *Econometrica*, 2018, *85* (5), 1331–1371.

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, Chris 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.

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," *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.*

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

Galor, Oded and Ömer Özak, "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 I. 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.

— , **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.

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

Krusell, Per, Lee. E. Ohanian, José-Victor Rios-Rull, and Giovanni L. Violante, "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

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.

Moretti, Enrico, "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," *American Economic Review*, 2004, *94* (3), 565–690.

Mundlak, Yair, "On the Pooling of Time Series and Cross-Section Data," *Econometrica*, 1978, *46* (1), 69–86.

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

Pakes, Ariel, "On Group Effects and Errors in Variables in Aggregation," *Review of Economics and Statistics*, 1983, *65* (1), 168–173.

Psacharopoulos, George and Harry Anthony Patrinos, "Returns to investment in education: a decennial review of the global literature," *Education Economics*, 2018, *26* (5), 445–458.

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

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 reciprocity, 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 reciprocity, 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 ten countries in central Asia (Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Russia, Tajikistan, Turkmenistan, and Uzbekistan), two countries in South-East Asia (Bangladesh and Cambodia), five countries in Southern and Eastern Europe (Croatia, Hungary, Poland, Romania, Turkey), four countries in the Middle East and North Africa (Algeria, Lebanon, Jordan, and Saudi-Arabia), and one 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.[32]

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.

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

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 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.[33] 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.

[33] 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.

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 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 twelve 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 twelve months. The values in the tree denote the amounts of money to be received in twelve 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.

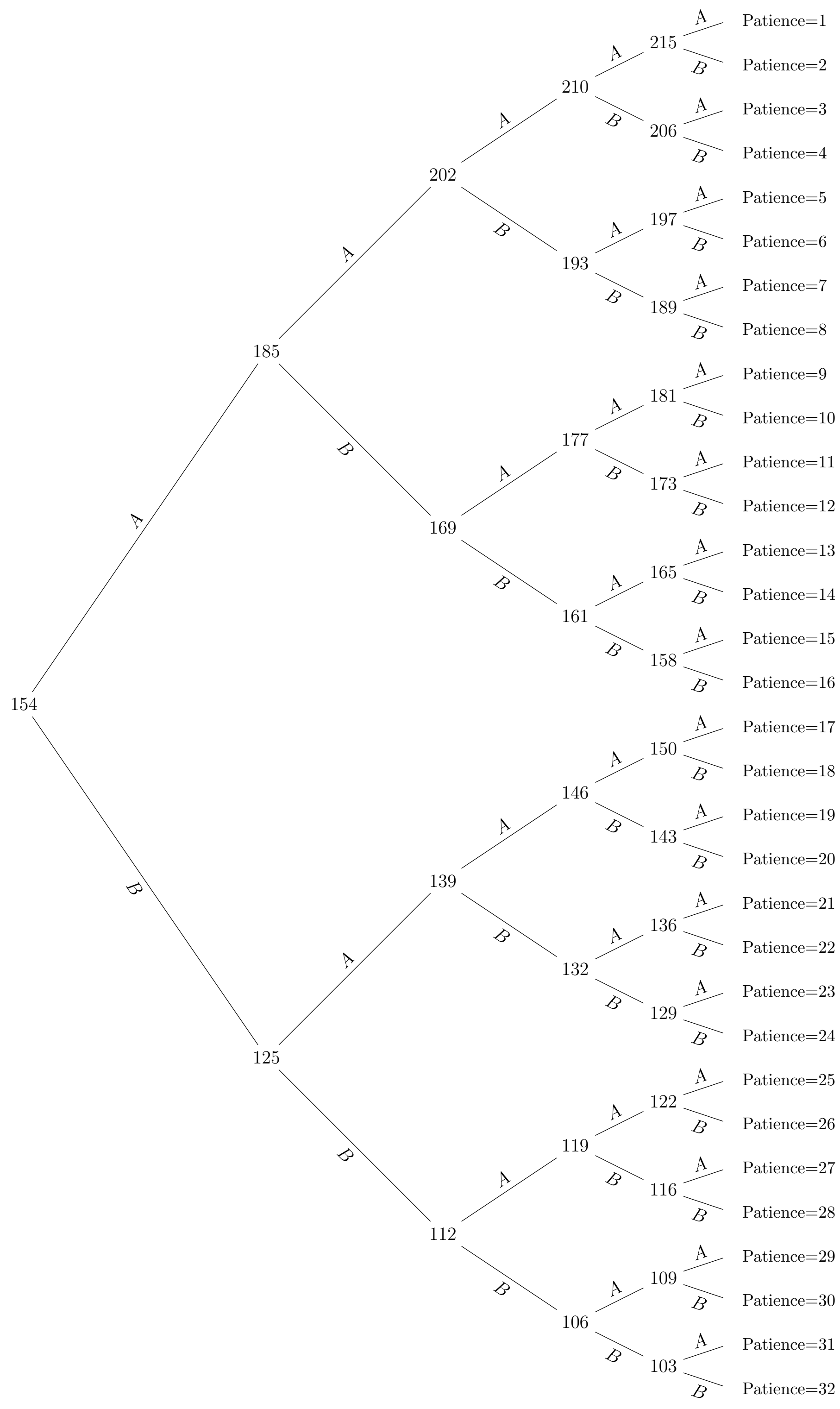


Figure 3: Tree for the staircase time task as implemented in Germany (numbers = payment in twelve months, A = choice of "100 Euros today", B = choice of "x Euros in twelve months"). First, each respondent was asked whether they would prefer to receive 100 Euros today or 154 Euros in twelve months from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

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:

- 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 the 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):

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

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 Description and Sources of Main Variables

B.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.

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.

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 quantifies 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 QOG dataset.

B.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.

B.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 $\ln(1 + \text{household income})$.

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?*.

C Details for Regional-Level Analysis

C.1 Data

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.2 Adjustment of region-level patience

We shrink observed average patience in a region towards the average of all region averages in our global sample. Formally, suppose that true average patience in region j is given by β_j^t and observed average region-level patience by $\beta_j = \beta_j^t + \epsilon_j$. Then, shrunk patience of region j, β_j^s, is computed as a convex combination of observed average patience in region j and the mean of all region averages $\bar{\beta}$:

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

Here, w_j represent region-specific weights that capture the signal-to-noise ratio. Recall that the traditional James-Stein estimator is derived by constructing the ratio of the total variance of β_j and its signal variance. Computing the variance of β_j and rearranging, the signal variance is given by $Var(\beta_j^t) = Var(\beta_j) - Var(\epsilon_j) = Var(\beta_j) - \mathbb{E}[(se_j)^2]$, where $Var(\beta_j)$ is the variance of the region means and se_j the standard error of β_j in region j. Then, the traditional James-Stein ratio of signal variance to total variance is given by $w = \frac{Var(\beta_j) - \mathbb{E}[(se_j)^2]}{Var(\beta_j)}$. Note that this weight would be constant across regions. Following Chetty and Hendren (2016), we adjust this weight for heteroscedasticity (the regional means are estimated with different levels of precision across regions) as follows:

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

The weights are decreasing in the standard error of the sample mean in a given region. These weights can be computed in the data because the variance of the regional averages, the within-region squared standard errors, and their average across regions are all observable.

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 (5) 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

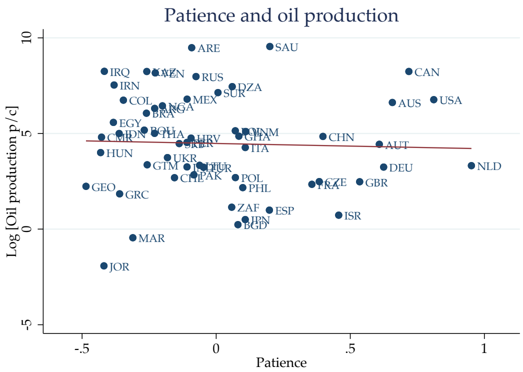
	Dependent variable: Annual growth rate in GDP p/c (in %) since...					
	1950			1975		
	(1)	(2)	(3)	(4)	(5)	(6)
Patience	0.83**	1.00**	1.25***	0.75*	1.42***	1.96***
	(0.32)	(0.41)	(0.30)	(0.41)	(0.40)	(0.43)
Log [GDP p/c base year]		-0.81***	-1.18***		-1.02***	-1.65***
		(0.22)	(0.19)		(0.19)	(0.27)
Continent 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. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

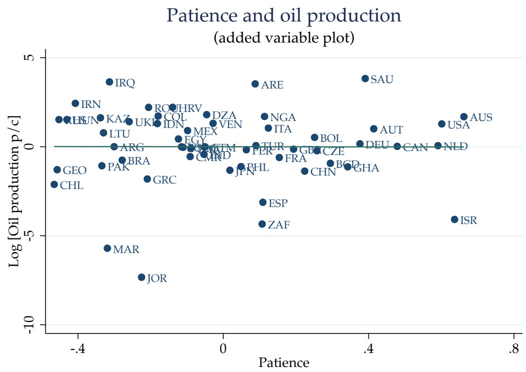
Table D.4: Patience and human capital: Alternative measures

	Dependent variable:					
	% Tertiary Education			Test Scores		
	(1)	(2)	(3)	(4)	(5)	(6)
Patience	18.7***	9.45**	7.49	0.81***	0.41***	0.34*
	(3.51)	(3.71)	(5.24)	(0.13)	(0.13)	(0.19)
Continent FE	No	Yes	Yes	No	Yes	Yes
Additional controls	No	No	Yes	No	No	Yes
Observations	72	72	71	49	49	48
R^2	0.27	0.55	0.68	0.28	0.65	0.75

OLS estimates, robust standard errors in parentheses. Dependent variables are the percentage aged 25+ that has tertiary education (Barro and Lee, 2012) and harmonized test scores (Hanushek and Woessmann, 2012). See column (5) of Table 1 for a complete list of the additional controls. * $(p < 0.10)$, ** $(p < 0.05)$, *** $(p < 0.01)$



(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)

Table D.5: Individual patience, and human capital: Alternative measures

	Dependent variable:							
	1 if tertiary educ.				Subj. cognitive Skills (Math)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience	0.054***	0.034***	0.034***	0.027***	0.37***	0.29***	0.27***	0.25***
	(0.01)	(0.00)	(0.00)	(0.00)	(0.03)	(0.02)	(0.02)	(0.02)
Age				0.89***				-1.19
				(0.09)				(0.87)
Age squared				-1.00***				-0.70
				(0.09)				(0.81)
1 if female				0.0096				-0.63***
				(0.01)				(0.05)
Subj. math skills				0.019***				
				(0.00)				
Subjective institutional quality				0.0036				0.023
				(0.01)				(0.07)
Confidence in financial institutions				0.62				8.11*
				(0.49)				(4.47)
Subjective law and order index				-0.0049				0.088
				(0.01)				(0.06)
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	79357	79357	78403	46550	78755	78755	77837	46688
R^2	0.02	0.12	0.15	0.18	0.02	0.07	0.12	0.13

Individual-level OLS estimates, standard errors (clustered at country level) in parentheses. The dependent variable in (1)–(4) is 1 if the individual has tertiary education. The dependent variable in (5)–(8) subjective assessment of math skills. Age is divided by 100. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

Table D.6: 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				0.28				-0.62*
				(0.28)				(0.31)
Age squared				0.20				-0.37
				(0.38)				(0.30)
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
Subnational region 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 the 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. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

Table D.7: Individual patience, and income: Aggregation

	Dependent variable: Log [Income p/c]					
Level of aggregation:	Individual			Region		Country
	(1)	(2)	(3)	(4)	(5)	(6)
Patience	0.34***	0.056***	0.049***	1.51***	0.15***	2.14***
	(0.05)	(0.01)	(0.01)	(0.21)	(0.05)	(0.18)
Country FE	No	Yes	No	No	Yes	No
Subnational region FE	No	No	Yes	No	No	No
Observations	79245	79245	78271	1046	1046	76
R^2	0.05	0.61	0.64	0.29	0.90	0.47

OLS estimates, based on GPS/Gallup data at individual and aggregate levels. In each column, we aggregate the GPS data up to the relevant level. Regional patience is shrunk using the same procedure as described in Section 3.2, where we again exlude regions with fewer than 15 respondents. Robust standard errors (clustered at country level) in parentheses. * ($p < 0.10$), ** ($p < 0.05$), ***($p < 0.01$)

E 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.

E.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 $p_o = p_t + \alpha \cdot \epsilon$, 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.

E.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 regressing 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.

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 (F.5)*, respectively.*

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

Proof. Consider the indifference condition for education (5). 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,

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

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,

$$\eta^d = \frac{w^H}{w^L} = \left[K^\theta + H^\theta\right]^{\frac{\sigma-\theta}{\theta}} \frac{L^{1-\sigma}}{H^{1-\theta}} = \left[\left(\frac{K}{H}\right)^\theta + 1\right]^{\frac{\sigma-\theta}{\theta}} \left(\frac{L}{H}\right)^{1-\sigma} \; , \tag{F.2}$$

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

$$K^d = \left[\eta^{\frac{\theta}{\sigma-\theta}} \cdot \left(\frac{H}{L}\right)^{\frac{\theta(1-\sigma)}{\sigma-\theta}} - 1\right]^{\frac{1}{\theta}} \cdot H \; , \tag{F.3}$$

while in steady state the supply of capital given by (9) simplifies to

$$K^s = \frac{1}{2\varepsilon} w^L \left\{ \left[2\varepsilon(1-\lambda) - \ln\left(1 + \frac{2\varepsilon(1-\lambda)}{1+\chi-\varepsilon}\right)\right] + \eta h \left[2\varepsilon\lambda + \ln\left(1 - \frac{2\varepsilon\lambda}{1+\chi+\varepsilon}\right)\right] \right\} \; , \tag{F.4}$$

which is an increasing function in η.

Now note that *(a)* is satisfied via (F.2) 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 (F.4) with (F.2) and using the expressions for L and H in *(c)* gives

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

$$\cdot \frac{A\left[\left(\frac{H}{L}\right)^{\frac{\sigma(1-\theta)}{\sigma-\theta}} \cdot \left(\eta^d\right)^{\frac{\sigma}{\sigma-\theta}} + 1\right]^{\frac{1-\sigma}{\sigma}}}{2\varepsilon}, \tag{F.5}$$

so that any solution to *(b)* will also fulfill *(c)*. Existence and uniqueness then follows since (F.3) and (F.5) determine a unique tuple of η and K, which in turn via (F.2) and (F.1) determines a unique tuple of η and $\tilde{\beta}$ for a given (unique) K. To see this, note that from (F.3) K^d is smaller than zero for $\eta = 0$ and increasing in η while from (F.5) 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 $\eta > 0$ 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 (F.1) is a decreasing, convex function in the patience threshold $\tilde{\beta}$ that approaches infinity from below as $\tilde{\beta} \to 0$, while (F.2) 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 as well as 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 (F.1) and (F.2), inserting (F.4), $H = \lambda e^{\rho(1-\psi)}$, $L = 2(1-\lambda)+\psi\lambda$, 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

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

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

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

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

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

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

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

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

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

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

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

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 $0 < \hat{\theta} < \theta < \sigma < 1$

$$\frac{d\,\tilde{\beta}}{d\,\chi} > 0 \ , \quad \frac{d\,\tilde{\beta}}{d\,\chi} < 1 \ , \quad \frac{d\,\lambda}{d\,\chi} > 0 \ , \quad \frac{d\,L}{d\,\chi} < 0 \ , \quad \frac{d\,H}{d\,\chi} > 0 \ , \quad \frac{d\,K}{d\,\chi} > 0 \ .$$

G Testable Predictions: Derivation of Expressions

Set-up: Consider a country with a uniform distribution of patience between $\chi - \varepsilon$ and $\chi + \varepsilon$. 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 (5),

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

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 $\mathcal{I}_{\text{skilled}}(i) = \mathbb{1}\{\eta^* - \eta(\beta(i)) + u(i) > 0\}$, where the expression in brackets corresponds to the economic considerations related to the decision to become skilled, and $u(i)$ is an additive noise term that is symmetrically distributed around zero and reflects idiosyncratic factors influencing this decision.[34] From a population perspective, it then holds that the predicted share of skilled individuals is given by

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

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.[35] The marginal effect of an increase in patience on the propensity of becoming skilled can then be expressed as

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

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

[34] To keep a meaningful role for patience, the distribution of the noise term is also assumed to be bounded on the admissible support for η.

[35] 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

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

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

Notice that the OLS estimates in Table 7 correspond to the population average of a linearized version of this effect under the assumption that $\eta(i)$ is affected by an additive 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

$$\lambda_2 - \lambda_1 \equiv \Delta\lambda = \frac{1}{2\varepsilon}\left(\chi_2 + \varepsilon - \tilde{\beta}_2 - (\chi_1 + \varepsilon - \tilde{\beta}_1)\right)$$

$$= \frac{1}{2\varepsilon}\left(\chi_1 + sd + \varepsilon - \tilde{\beta}_2 - (\chi_1 + \varepsilon - \tilde{\beta}_1)\right)$$

$$= \frac{1}{2\varepsilon}\left(sd - \underbrace{(\tilde{\beta}_2 - \tilde{\beta}_1)}_{\equiv \Delta\beta}\right)$$

$$\Delta\lambda = \frac{1}{\sqrt{12}} - \frac{1}{2\varepsilon} \cdot \Delta\beta \ . \tag{G.2}$$

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

$$\frac{1}{\sqrt{12}} - \frac{1}{2\varepsilon} \cdot \Delta\beta > -\frac{1}{\sqrt{12}} \cdot \frac{\left[\tilde{\beta}\left(1+\tilde{\beta}\right)\right]^2}{\frac{\psi^{\frac{-1}{\tilde{\beta}(1+\tilde{\beta})}}}{h} \ln\psi\left(1+2\tilde{\beta}\right)}$$

$$1 - \frac{\Delta\beta}{sd} > -\frac{\left[\tilde{\beta}\left(1+\tilde{\beta}\right)\right]^2}{\frac{\psi^{\frac{-1}{\tilde{\beta}(1+\tilde{\beta})}}}{h} \ln\psi\left(1+2\tilde{\beta}\right)}$$

$$1 - \frac{d\,\tilde{\beta}}{d\,\chi} > -\frac{\left[\tilde{\beta}\left(1+\tilde{\beta}\right)\right]^2}{\frac{\psi^{\frac{-1}{\tilde{\beta}(1+\tilde{\beta})}}}{h} \ln\psi\left(1+2\tilde{\beta}\right)} \ .$$

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 in $\tilde{\beta}$, 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 (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

$$sd \cdot \frac{\partial \ln \bar{S}}{\partial \beta(i)} = sd \cdot \frac{1}{\bar{S}} \left\{ \frac{1}{1+\tilde{\beta}}\left[\frac{(1-\lambda)\cdot w^L}{1+\chi-\varepsilon} + \frac{\lambda \cdot w^H h}{1+\chi+\varepsilon}\right]\right\} . \tag{G.3}$$

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.[36] The increase in log capital per capita is then given by

$$\frac{k_2 - k_1}{k_1} . \tag{G.4}$$

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

[36] 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.

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

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

$$\frac{d\,K}{d\,\chi} > \frac{\partial \bar{S}}{\partial \beta(i)} \ .$$

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

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

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

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

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

$$\bar{y} = \left[2\left(1 - \overline{\mathcal{I}}_{\text{skilled}}\right) + \psi\right] w^L + \overline{\mathcal{I}}_{\text{skilled}} w^H h + R \cdot \bar{S} \ .$$

The average marginal effect on individual income is then given by

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

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

$$sd \cdot \left.\frac{\partial \ln \bar{y}}{\partial \beta(i)}\right|_{\beta(i)=\tilde{\beta}} = sd \cdot \frac{1}{\bar{y}} \left\{ \left.\frac{\partial \mathcal{I}_{\text{skilled}}(i)}{\partial \beta(i)}\right|_{\beta(i)=\tilde{\beta}} \left[w^H h - (2-\psi) w^L\right] + R \cdot \frac{\partial \bar{S}}{\partial \beta(i)} \right\} \ . \tag{G.5}$$

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

$$\frac{y_2 - y_1}{y_1} \ . \tag{G.6}$$

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

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

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

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

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 – Additional Results

H.1 Results for Non-Targeted Moments

Table H.8: Simulated moments

	Data	Model			
	(Baseline Controls)	Baseline	Extensions		Dev. Acc.
			HC-Ext. (est.)	HC-Ext. (cal.)	
	(A) Elasticities: Effect of 1 SD increase of patience on				
Aggregate capital	1.17	0.50	2.59	1.43	1.75^a
HH savings (% Income)	9.80	6.25	6.26	5.06	0^a
Years of Schooling	2.47	1.94	2.20	1.91	0.19^a
	(B) Moments and Comparative Differences				
Skill shares ($\lambda_1,\ \lambda_2$)	0.59, 0.82	0.36, 0.52	0.36, 0.54	0.27, 0.42	0.35, 0.37
Ratio skill shares ($\frac{\lambda_2}{\lambda_1}$)	1.39	1.43	1.49	1.58	1.04
Capital ratio ($\frac{K_2}{K_1}$)	3.23	1.50	3.59	2.43	2.75
Output ratio ($\frac{Y_2}{Y_1}$)	2.98	1.11	2.74	1.80	2.79

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 10. [a] Differences between the high and low TFP economy instead of patience elasticities.

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

H.2 Additional Results – Model Extensions

This appendix contains the results of alternative model extensions.

- (A): TFP Calibrated: calibration of $\bar{A}_2$ to correspond to the patience elasticity of TFP as Table 3, column (6).
- (B): Combination of estimated human capital externality, γ, and TFP calibration as in A.
- (C): Human capital externality from top tail. Concretely, to account for the role of upper tail human capital, we assume that firms run by highly skilled people exhibit higher productivity. We capture this by including a level effect of average upper tail patience (the average patience of skilled workers, $\bar{\beta}$) on TFP and replacing the externality of aggregate human capital on productivity as in (10) by

  $$A_t = \bar{A} \cdot \left[h \cdot \bar{\beta}\right]^{\varphi}$$

  and estimating φ.
- (D): Re-estimation of ρ while keeping the remaining parameters as calibrated or estimated in the baseline version.
- (E): Re-estimation of $\{\sigma, \theta\}$ while keeping the remaining parameters as calibrated or estimated in the baseline version.
- (F): Alternative calibration of $\{\sigma, \theta\}$ to assess sensitivity of results. All other parameters are re-estimated analogous to the estimation of the baseline model.

Table H.9: Calibrated and estimated parameters: Alternative Specifications and Extensions

	Baseline	Alternative Extensions					
		TFP Cal.	HC-Ext. & TFP-Cal.	HC-Ext. (top-tail)	$\hat{\rho}$	$\left\{\hat{\sigma}, \hat{\theta}\right\}$	$\{\sigma, \theta\}$ Alt. Cal.
	(1)	(A)	(B)	(C)	(D)	(E)	(F)
χ_1, χ_2	0.16, $[0.25]^a$	0.23, $[0.25]^a$	0.22, $[0.25]^a$	0.16, $[0.25]^a$	0.16, $[0.25]^c$	0.16, $[0.25]^c$	0.16, [0.24]
ϵ	0.16	0.05	0.05	0.16	0.16^c	0.16^c	0.14
$\frac{\bar{A}_1}{\bar{A}_2}$	1	1.17^b	1.17^b	1	1	1	1
γ			1.73				
φ				3.65			
ρ					4.72		
σ						0.83	0.51
θ						0.34	0.49

Estimated or calibrated parameters for alternative specifications. Unless noted otherwise, calibrated parameters take the same values as in the baseline calibration, see Table 8. Parameters in brackets [ ] are derived from estimated parameters. See text for details. [a] Level of χ_2 as implied by the estimated values of χ_1 and ε. [b] $\frac{\bar{A}_1}{\bar{A}_2}$ calibrated to match elasticity in TFP as in Table 3(6). [c] Values kept at same level as in baseline model.

Table H.10: Quantified model vs. data: Alternative Specifications and Extensions

	Effect of one SD increase of patience							
	Data	Model						
	(Baseline Controls)	Baseline	Alternative Extensions					
			TFP-Cal.	HC-Ext.& TFP-Cal.	HC-Ext. (top-tail)	$\hat{\rho}$	$\left\{\hat{\sigma}, \hat{\theta}\right\}$	$\{\sigma, \theta\}$ Alt. Cal.
	(1)	(2)	(A)	(B)	(C)	(D)	(E)	(F)
	Individual level							
Income	0.05	0.05	0.06	0.05	0.05	0.06	0.06	0.05
Education	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
	Country level							
Income	1.73	0.11	0.54	1.79	1.75	0.16	0.15	0.09
Fraction skilled	0.20	0.16	0.13	0.17	0.18	0.19	0.21	0.13

The effect sizes in the simulated model are obtained after estimating the parameters through indirect inference as reported in Table H.9. In the baseline, estimated parameters are χ_1, ε, and $\bar{A}$ by matching as moments the effects of patience on individual income, individual propensity to become skilled, aggregate income and aggregate skill share. For details on the target moments from the data see Table 9.

Table H.11: Simulated Moments: Alternative Specifications and Extensions

	Data	Model						
	(Baseline Controls)	Baseline	Alternative Extensions					
			TFP-Cal.	HC-Ext. & TFP-Cal.	HC-Ext.(top-tail)	$\hat{\rho}$	$\left\{\hat{\sigma}, \hat{\theta}\right\}$	$\{\sigma, \theta\}$Alt. Cal.
	(1)	(2)	(A)	(B)	(C)	(D)	(E)	(F)
	(A) Elasticities: Effect of 1 SD increase of patience on							
Aggregate capital	1.17	0.50	0.51	1.62	2.59	0.52	0.59	0.45
HH savings (% Income)	9.80	6.25	1.71	1.98	6.25	6.25	6.25	5.81
Years of Schooling	2.47	1.95	1.63	2.12	2.20	2.40	2.65	1.58
	(B) Moments and Comparative Differences							
Skill shares ($\lambda_1, \ \lambda_2$)	0.59, 0.82	0.36, 0.52	0.87, 1	0.65, 0.82	0.36, 0.54	0.47, 0.67	0.36, 0.57	0.35, 0.48
Ratio skill shares ($\frac{\lambda_2}{\lambda_1}$)	1.39	1.43	1.15	1.26	1.49	1.41	1.60	1.37
Capital ratio ($\frac{K_2}{K_1}$)	3.23	1.50	1.51	2.62	3.59	1.52	1.59	1.45
Output ratio ($\frac{Y_2}{Y_1}$)	2.98	1.11	1.54	2.79	2.75	1.16	1.15	1.09

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 H.9.

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