CAGE

Patience and Subnational Differences in Human Capital: Regional Analysis with Facebook Interests

CAGE working paper no. 731

November 2024

Eric A. Hanushek,
Lavinia Kinne,
Pietro Sancassani,
Ludger Woessmann

WARWICK
THE UNIVERSITY OF WARWICK

UKRI
Economic and Social Research Council

Patience and Subnational Differences in Human Capital: Regional Analysis with Facebook Interests*

Eric A. Hanushek, Lavinia Kinne, Pietro Sancassani, and Ludger Woessmann†

Abstract

Decisions to invest in human capital depend on people's time preferences. This paper shows that differences in patience are closely related to substantial subnational differences in educational achievement, leading to new perspectives on longstanding within-country disparities. We use social-media data – Facebook interests – to construct novel regional measures of patience within Italy and the United States. The approach is first validated with a cross-country analysis of patience and Facebook interests. We then show that patience is strongly positively associated with student achievement across regions in both countries, accounting for three-quarters of the achievement variation across Italian regions and one-third across U.S. states. The finding is confirmed in an identification strategy employing variation in ancestry countries of the current population of U.S. states. Results hold for six other countries with more limited regional achievement data.

Keywords: patience, human capital, student achievement, regions, social media, Facebook

JEL classification: I21, Z10

September 27, 2024

* We gratefully acknowledge comments from the editor Steffen Huck, two anonymous referees, Elliott Ash, Davide Cantoni, David Figlio, Philipp Lergetporer, Ömer Özak, Solomon Polachek, Paola Sapienza, Uwe Sunde, and seminar participants at Cornell University, the University of Rochester, the annual meetings of the American Economic Association and the German Economic Association, the CESifo big data workshop, the IZA education workshop, the Berlin-Munich CRC retreat in Ohlstadt, the IWAEE workshop in Catanzaro, the ifo Center for the Economics of Education, and the CESifo Area Conference on the Economics of Education. This work was supported by the Smith Richardson Foundation. The contribution of Woessmann is also part of German Science Foundation project CRC TRR 190.

† Hanushek: Hoover Institution, Stanford University; CESifo, IZA, and NBER, hanushek@stanford.edu. Kinne: DIW Berlin, lkinne@diw.de. Sancassani: ifo Institute, University of Munich, sancassani.pietro@gmail.com. Woessmann: University of Munich, ifo Institute; Hoover Institution, Stanford University; CESifo, and IZA, woessmann@ifo.de.

1. Introduction

Differences in time preferences have been recognized as an important determinant of individual investments in skills since the earliest human capital theory of Becker (1964). More recent analyses show that differences in patience are closely related to cross-country differences in educational outcomes and in the resulting differences in income and growth. But investigation of the role of such fundamental preferences in explaining historically significant *subnational* variations in education and incomes has been stymied by a lack of representative region-specific measures of time preferences. We combine the massive data available from social media – specifically Facebook interests – with machine-learning algorithms to derive new regional measures of patience. We find that patience has a significant role in accounting for within-country differences in student achievement and, by implication, geographical variations in incomes and economic growth.

Many countries have large, longstanding regional differences in student achievement that in turn affect regional income and growth (Hanushek and Woessmann (2015); Hanushek, Ruhose, and Woessmann (2017)). Our analysis here focuses on Italy and the United States – two countries with achievement data for a large number of subnational regions showing persistent historical education and income variations.[1] The differences in math achievement between the top- and bottom-performing U.S. states equal roughly two-thirds of the achievement differences between top- and bottom-performing countries in the OECD and are equivalent to over two years of learning. Similar differences are found between the top- and bottom-performing regions in Italy.

[1] The large North-South variation in Italy has raised substantial interest in policy and research (e.g., Putnam (1993); Ichino and Maggi (2000); Guiso, Sapienza, and Zingales (2004) and has been shown to be related to concepts of social preferences such as trust and social capital (Bigoni et al. (2016)), beliefs about cooperativeness (Bigoni et al. (2019)), and civicness (Michaeli et al. (2023)).

Patience, the relative valuation of present versus future payoffs, appears not only in individual-level but also in group-level decisions. Students weigh current gratification such as play time with friends against study time that may lead to deferred rewards. Communities trade off present costs against future benefits when deciding how much to invest in school quality, how strongly to motivate children to learn, and whether to design institutions to incentivize learning. Testing any hypothesized contribution of patience in affecting regional differences in educational achievement, however, requires representative regional measures of preferences.

A major innovation of this paper is demonstrating how social-media data can be used to derive subnational measures of preferences. The fundamental idea, building on recent international analysis of culture by Obradovich et al. (2022), is that social-media data contain information that characterizes people's underlying preferences such as patience at geographically granular levels. This in turn permits investigation of previously unexplained place-based heterogeneity of skills.

For marketing purposes, Facebook has developed an algorithm to classify the "interests" of over two billion users. We identify the 1,000 Facebook interests with the largest audiences worldwide and then use Facebook's marketing application programming interface (API) to extract data on the prevalence of these interests by country and subsequently by region. After reducing their dimensionality with a principal component analysis (PCA), we employ machine-learning techniques to train an international model that predicts the experimentally validated patience measure of the Global Preference Survey (GPS) of Falk et al. (2018).

We first validate the use of our Facebook-derived measure of patience for characterizing educational outcomes through an international analysis that mimics existing investigations of preferences and cross-country achievement (Figlio et al. (2019); Hanushek et al. (2022)).

We then use the parameters from the international Facebook analysis along with observed regional Facebook interests to construct subnational patience measures across 20 Italian regions and 50 U.S. states. In both countries, the geographic pattern of the Facebook-derived patience measure coincides with longstanding North-South economic disparities. Regional differences in patience account for over three-quarters of the variation in student achievement across Italian regions and for over one-third across U.S. states. In Italy, a one standard deviation (SD) higher regional patience relates to 1.2-1.6 SD higher math achievement, only slightly lower than the cross-country relationship. Possibly related to the substantial internal mobility, the equivalent U.S. estimate is about one quarter that for Italy albeit still statistically significant.

Two analyses, while not conclusive, are in line with a causal interpretation of the descriptive baseline association. First, in a cross-country analysis we consider differential outcomes for migrants within each country based on the Facebook preferences of their origin country. By conditioning on fixed effects for the migrants' country of residence, we shield against unobserved features of students' residence countries and against reverse causation from education or ability to preferences (as suggested, e.g., by Dohmen et al. (2010) or Benjamin, Brown, and Shapiro (2013)). Results are qualitatively similar to the aggregate cross-country results. Compared to the international analysis, the within-country estimation for Italy and the United States is less prone to confounding from unobserved national traits such as languages, constitutions, and institutional factors. Second, a complementary instrumental-variable approach goes further in the subnational analysis to address potential confounders. Ancestors of the current populations of the different U.S. states migrated from different countries, giving rise to variation in patience that is not jointly determined with other current state characteristics. Using the weighted patience level of the ancestry countries as an instrument for states' patience confirms a significant effect of patience on student achievement of the same order of magnitude as the baseline model.

Results are stable in robustness analyses that include using reading achievement, differentiating by gender or assessment wave, adding risk-taking and trust as additional cultural-trait controls, and excluding education-related interests when deriving the patience measure from Facebook interests. Moreover, results are consistent for six additional countries where regional achievement data cover fewer grades or regions. The positive association between regional student achievement and Facebook-derived patience holds in the aggregate and is separately significant in five of the additional countries – Brazil, Canada, Germany, Kazakhstan, and Mexico – excluding only Spain.

Our main contribution is showing how the well-established theoretical and empirical relationship between patience and human capital accumulation can explain important parts of subnational differences in skills.[2] Prior analyses have largely left persistent regional differences in skills and income unexplained.[3] Our derivation of the regional patience measures further supports the use of social-media data in economic analysis of culture and social networks.[4] With our better quantitative measures of more fundamental preference differences, it becomes possible to understand portions of educational and economic outcomes that were previously listed as unexplained heterogeneity.

---

[2] Time preferences are important for economic development (Galor and Özak (2016); Sunde et al. (2022)), and previous analyses have shown that they are an important determinant of individual educational outcomes (Sutter et al. (2013); Golsteyn, Grönqvist, and Lindahl (2014); Polachek, Das, and Thamma-Apiroam (2015); De Paola and Gioia (2017); Castillo, Jordan, and Petrie (2019); Galor, Özak, and Sarid (2020); Angerer et al. (2023)) and of international achievement differences (Figlio et al. (2019); Hanushek et al. (2022)).

[3] Past studies consider proximate causes of regional skill differences such as family background, school spending, and institutional settings (e.g., Hanushek and Raymond (2005); Woessmann (2010); Dee and Jacob (2011)). Yet, most stop without providing convincing explanations of more fundamental causes of the substantial on-going geographical variations in outcomes (e.g., Hanushek (2016)).

[4] E.g., Wilson, Gosling, and Graham (2012); Obradovich et al. (2022); Chetty et al. (2022); Bailey et al. (2022); Marty and Duhaut (2024).

2. Methods: Deriving Regional Patience Measures from Facebook Interests

With 2.9 billion monthly active users, Facebook is the world's largest social network. Facebook's core business consists of selling advertising space which provides 97.5 percent of its revenues. Hence, Facebook's business model depends on its ability to keep users engaged on the platform while advertisers promote their products and services. To this purpose, Facebook puts considerable effort into inferring users' interests (Thorson et al. (2021)), which is critical to our analysis.

2.1 Extracting Facebook Interests

Facebook determines users' interests using a variety of sources, both inside the Facebook platform and on external websites (Cabañas, Cuevas, and Cuevas (2018); Obradovich et al. (2022)). Sources inside the Facebook platform include personal information that users share on Facebook as well as users' activity on Facebook, such as page likes, group memberships, and content with which users engage. Outside the platform, Facebook tracks users' visited websites, installed apps, and purchasing behavior. Facebook uses these data to deliver content and recommendations based on users' interests and to allow advertisers to target users whose interests are relevant for their products and services.

The hundreds of thousands of interests classified by Facebook are organized in nine main categories: business/industry, entertainment, family/relationships, fitness/wellness, food/drink, hobbies/activities, shopping/fashion, sports/outdoors, and technology. Interests can be very broad, such as "Entertainment" or "Music", or very narrow, such as "Caribbean Stud Poker", a casino game. Appendix Figure A1 depicts the 1,000 Facebook interests with the largest worldwide audience.

Following Obradovich et al. (2022), we retrieve Facebook interest data in two steps. First, we obtain a comprehensive list of Facebook interests by querying Facebook's marketing API, the interface that allows advertisers to configure their advertisement campaigns. Per text query, the API returns a collection of closely related Facebook interests with their estimated worldwide audience. We iteratively feed this function with all 25,322 terms of an English dictionary and 2,000 randomly selected titles of Wikipedia articles. This procedure produces 41,513 unique interests from which we select 1,000 with the largest worldwide audience.[5]

Second, for each of the 1,000 interests, we again use Facebook's marketing API to obtain the estimated audience size separately for each country in which Facebook has a presence, as well as for each Italian region and U.S. state. The audience size reflects the entire user population, rather than any specific subgroup such as parents or students, because we attempt to develop proxies for prevailing national or regional preferences. The population of Facebook users may not be representative for the entire population, but our method builds on the Facebook data's ability to predict a patience measure collected in representative samples. For each geographical entity, the process yields a vector of estimated audiences for each of the 1,000 interests. We standardize the estimated audience to mean zero and SD one across the 1,000 interests in each geographical entity.[6]

2.2 Predicting Country Patience from Facebook Interests

We first construct a country-level Facebook measure that we use to validate the overall approach of going from Facebook interests to patience. In the next section, we follow a

[5] We use 1,000 interests to make the data collection manageable. During data collection between April 2022 and May 2023, the API allowed a maximum of 300 queries per hour. For example, the over 50,000 queries for the U.S. states take more than seven days of uninterrupted queries.

[6] Dividing the Facebook audience counts by population or Facebook users in each geographic entity yields the same qualitative results.

conceptually similar approach to developing the subnational measures that are the heart of our analysis.

We build our Facebook measure of patience through country-level preferences developed in the Global Preference Survey (GPS) from representative population samples in 76 countries (Falk et al. (2018)). The GPS collected experimentally validated measures of patience (and other preferences) by combining a qualitative survey item and a hypothetical choice scenario that were chosen based on their capacity to predict incentivized choices in a laboratory setting.

We begin by reducing the dimensionality of Facebook interests by a principal component analysis (PCA) over the sample of all 216 Facebook countries/entities. The first 10 principal components (PCs) capture 70 percent of the total cross-country variance contained in Facebook interests, 20 PCs capture 80 percent, and 48 PCs capture 90 percent (Appendix Figure A2).

We then train a machine-learning model to characterize the relationship between the country-level PCs of Facebook interests and the GPS measure of patience across the 74 countries that have both GPS and Facebook data (Appendix Table A1).[7] Using a 10-fold cross-validated least absolute shrinkage and selection operator (LASSO) model, the R^2 of the in-sample prediction of patience by the reduced-dimensionality Facebook interests is quite stable between 0.65 and 0.70, independent of whether 10, 20, or 50 PCs are used (Appendix Figure A3).

We use the parameters from the machine-learning model to predict patience for all 80 countries with both PISA and Facebook data (Appendix Table A1 and Appendix Figure A4). Given the limited size of the training sample used, we rely on the parsimonious specification with 10 PCs for the out-of-sample predictions to avoid overfitting.

[7] The GPS measure is standardized to have mean zero and SD one across individuals in the GPS countries, so that estimates in our subsequent analyses can be interpreted in terms of SDs.

The beauty of this approach is its ability to capture latent components of users' underlying degree of patience. The procedure does not lend itself to identifying specific interests as main predictors of the GPS patience measure, because any single Facebook interest can load positively or negatively on different PCs that enter the patience prediction.[8] However, if we correlate our Facebook-derived patience measure with the individual Facebook interests, terms such as dogs, outdoor recreation, adventure, holiday, painting, wildlife, and garden consistently show up among the strongest positive correlations, whereas free software, WhatsApp, Facebook Messenger, massively multiplayer online role-playing games, and the like show up as negative. Among leisure activities (which are broadly represented by Facebook), the former patience-related interests may indicate valuation of longer-term future rewards, whereas the latter interests may proxy for instant gratification. Still, not all interests with strong correlations are easily interpretable, which aligns with the idea of picking up latent factors that are not well measured by individual Facebook interests. The Facebook-derived measure of patience can obviously be associated with regional differences in all sorts of other measures such as money, intelligence, or motivation – but, by construction, only to the extent that these are associated with patience as measured in the GPS.

We perform the same training and prediction models for risk-taking, another preference with relevance for intertemporal decision-making contained in the GPS and previously found to enter international student achievement (Hanushek et al. (2022)). The R^2 of the in-sample prediction for risk-taking is lower than for patience (Appendix Figure A3), indicating that risk-taking is harder to predict from Facebook interests.[9]

[8] Obradovich et al. (2022) show that direct prediction of cultural measures from individual Facebook interests generally provides worse predictions than using PCs.

[9] For other GPS preferences, the R^2 of the in-sample prediction is even lower at about 0.35 for trust and at most 0.2 for altruism, positive reciprocity, and negative reciprocity (see also Obradovich et al. (2022)).

2.3 Predicting Regional Patience from Facebook Interests

Our primary analysis builds on the development of subnational variations in patience. In parallel to the cross-country analysis, we reduce the dimensionality of Facebook interests using a PCA fit across regions *within* a given country. Fitting the PCA at the subnational level ensures that the PCs capture country-specific dimensions of Facebook interests. For both Italian regions and U.S. states, the first 4 PCs capture over 70 percent of the regional variance in Facebook interests (Appendix Figures A5 and A6); 90 percent of variance is captured by 10 PCs in Italy and 15 PCs in the United States.

Separately for Italy and the United States, we train a 10-fold cross-validated LASSO model to learn the relationship between the GPS measure of patience across countries and the Facebook interests aggregated by the parameters of within-country PCs.[10] A small number of PCs capture a considerable portion of the variation in the GPS patience measure. With 10 PCs, the R^2 of the in-sample prediction reaches 0.5 for PCs fitted across Italian regions and over 0.6 for PCs fitted across U.S. states (Appendix Figures A7 and A8).

We use the parameter estimates from the two internationally trained models to construct patience measures from the subnational Facebook interests of Italian regions and U.S. states. Figure 1 shows maps of the regional variation of patience in the two countries. In Italy, the regions with the lowest patience measure are Sicily and Campania in the South. The region with the highest level of patience is Trentino-Alto-Adige in the North-East.[11] In the United States, the

[10] We focus on all 74 countries with GPS and Facebook data to maximize the sample size for the machine-learning algorithm. However, results are quite similar if we instead only use the 25 OECD countries in the training model: the correlation among the two versions is 0.981 for Italian regions and 0.974 for U.S. states, and qualitative results do not change for our main regression analysis.

[11] Interestingly, parts of Trentino-Alto-Adige belonged to Austria and the former Austro-Hungarian empire for long periods of time, and large parts of the population speak German as their first language. The high level of predicted patience in Trentino-Alto-Adige is consistent with the fact that neighboring Austria has much greater

states that exhibit the highest level of patience are Vermont and Maine in the Northeast. Both countries tend to show a North-South gradient in the Facebook-derived measure of patience.

A similar prediction model for risk-taking performs substantially worse. The R^2 of the in-sample prediction is well below 0.2 for all models with up to 10 PCs in both Italy and the United States (see Appendix Figures A7, A8, and A9). We include risk-taking as a control variable throughout given its interrelatedness with patience, but its poor measurement at the regional level means that the estimates for patience are likely lower bounds.[12]

One way of directly validating our Facebook-derived patience measure is to compare it to regional GPS data. The GPS data contain regional identifiers that allow construction of non-representative regional GPS measures of patience (Sunde et al. (2022)). These are very noisy due to the small regional GPS sample sizes, averaging 50 individuals per Italian region and 20 per U.S. state. Nonetheless, they are positively correlated with our measure (weighted by the number of GPS observations per region): 0.49 (significant at the 5 percent level) across Italian regions and 0.23 (significant at the 10 percent level) across U.S. states.

3. Cross-country Results

The cross-country patience measures based on Facebook data allow us to validate whether Facebook interests can provide reliable estimates of geographically varying degrees of patience. They also provide some insights into the causal structure of the cross-country pattern of patience and achievement.

patience than Italy according to the country-level GPS measures, adding qualitative support for the Facebook-derived measure.

[12] In the cross-country analysis, patience and risk-taking are positively associated, and risk-taking is negatively associated with achievement, leading to negative bias (Hanushek et al. (2022)).

3.1 Cross-country Validation of Using Facebook Interests to Measure Patience

To validate our Facebook-derived measures of patience and risk-taking, we estimate their relationship with standardized student math achievement across countries for all seven available waves of the Programme for International Student Assessment (PISA) from 2000-2018 (see Hanushek et al. (2022)):

$$T_{ict} = \beta_1 Patience_c + \beta_2 Risk_c + \alpha_1 B_{ict} + \mu_t + \varepsilon_{ict} \tag{1}$$

where T_{ict} is the standardized PISA test score of student i in country c in year t; B is a vector of controls (student gender, age, and migration status); μ_t is a fixed effect for test wave to account for time trends and idiosyncrasies of individual tests; and ε_{ict} is an error term. β_1 and β_2 characterize the relationship of patience and risk-taking at the country level with student achievement. OLS regressions are weighted by students' sampling probability, giving equal weight to each country. Standard errors are clustered at the country level.

The comparison model from Hanushek et al. (2022) uses the original GPS measures and shows a strong positive relationship between patience and student achievement and a strong negative relationship with risk-taking (column 1 of Table 1, Panel A). Substituting our Facebook-derived preference measures (column 2) produces slightly larger preference parameters and corroborates the validity of the Facebook-derived measures.[13]

Out-of-sample predictions allow us to extend the analysis of the Facebook-derived measures of patience and risk-taking from 48 to 80 countries – all countries that participated in PISA and have Facebook data – encompassing over 2.6 million student observations. Results generalize very well to the extended sample, with increased precision and without significantly different

[13] The estimates rely on 10 PCs of Facebook interests but results are very similar when using additional (20-50) PCs (Appendix Table A2).

estimates (column 3). Even in the 32 countries that were not part of the original GPS analysis, results are qualitatively the same and statistically highly significant (column 4).

3.2 Exploration into Causality: Migrant Analysis

We also validate our measures with an analysis of migrant students that aims to get closer to a causal interpretation of the cross-country relationship between patience and student achievement. The most significant threats to identification of the preference effects are that the relationships are driven by reverse causation or by other factors of the country of schooling. We restrict the PISA analysis to students with a migrant background and assign them the values of patience and risk-taking of their home countries – an approach that avoids bias from reverse causation (see Figlio et al. (2019); Hanushek et al. (2022)).[14] By observing migrant students from different origin countries but schooled in the same residence country, we can include fixed effects for residence countries that control for other resident country factors that could bias the baseline cross-country analysis.

Migrant results in Panel B of Table 1 show that the prior positive patience relationship and the negative risk-taking relationship again replicate well.[15] When restricting the sample to non-GPS countries (column 4), estimates become quite imprecise (and larger), indicating limited power of the migrant analysis in the smaller sample.[16]

The cross-country migrant analysis further validates the informational content of the Facebook-derived measures and suggests a causal interpretation.

---

[14] Prior analysis suggests that migrant children obtain some of the preferences of their country of origin through family linkages (e.g., Guiso, Sapienza, and Zingales (2006); Bisin and Verdier (2011); Alesina and Giuliano (2014)).

[15] With the Facebook data, we expand the countries of origin considered in the migrant analysis from 56 to 93 (see Appendix Table A3). The destination countries increase only from 46 to 50 because some PISA countries do not report students' and parents' country of birth required to determine migrants' country-of-origin preferences.

[16] Results of the migrant analysis are again stable for patience when using 20-50 PCs (Appendix Table A4).

4. Subnational Results: Italian Regions and U.S. States

Our subnational empirical model follows the cross-country model introduced in the previous section. We think of patience as a deep determinant of student achievement, leading us to employ very parsimonious specifications of achievement differences. From this perspective, proximate inputs often included in education production functions such as parental education or school resources would be bad controls as they are endogenous to a region's patience.

Regionally representative data on student achievement come from student-level INVALSI test data for Italy and from state-level NAEP data for the United States and refer to the last waves before the COVID-19 pandemic (see Appendix A for details).[17] We initially focus on math achievement in eighth grade, the oldest cohort available in both countries and closest in age to PISA. To allow for interpretation in terms of standard deviations (SDs), we divide test scores by each country's student-level SD.

4.1 Achievement across Italian Regions

The longstanding North-South divide among Italian regions invites investigation of fundamental driving forces. Because the schooling system is regulated mostly at the country level, test score variations across regions are unlikely to be driven by the institutional structure of schools.

Regional differences in patience are strongly and significantly associated with student achievement. Student-level results in Panel A of Table 2 show that a one SD increase in regional patience is associated with an increase in math test scores of 1.40-1.61 SD, which is close to the

[17] INVALSI stands for Istituto Nazionale per la Valutazione del Sistema Dell'Istruzione and NAEP for National Assessment of Educational Progress.

cross-country estimates reported in Table 1. Measuring patience with 4, 7, or 10 PCs of Facebook interests has little impact.

Differences in patience account for over three-quarters of the aggregate regional variation in student achievement (Panel B of Table 2; see Appendix Figure A10).[18] Point estimates of the region-level analysis are very similar, albeit slightly smaller than in the student-level analysis.

The results are also consistent with a cumulative impact process: coefficient estimates at the student level increase continuously from an insignificant 0.31 SD in second grade to a highly significant 1.88 SD in tenth grade (Appendix Table A5). Region-level estimates are again quite similar. While other mechanisms may additionally be at play, the role of patience suggestively adds up across grades (see also Figlio et al. (2019)).

4.2 Achievement across U.S. States

Patience is also significantly associated with higher student achievement at the U.S. state level and accounts for slightly more than one-third of the state NAEP variation (Panel C of Table 2).[19] A one SD increase in the Facebook-derived measure of patience is associated with an increase of 0.17-0.29 SD in test scores.

The cross-state impact of patience is only about one quarter of that estimated for Italian regions. One possible explanation is that the population in the United States is substantially more mobile and mixed. In 2019, 42 percent of the U.S. population lives in a state different from their state of birth, suggesting that state cultural differences likely lessen over time. But the lower explanatory power of patience for U.S. states also suggests a wider range of other factors affecting achievement, such as compensatory state education policies.

[18] The R^2 is virtually unchanged when wave fixed effects and risk-taking are excluded from the model.

[19] The sample for Table 2 includes Washington, DC, but results are similar when it is excluded (not shown).

The impact of patience in fourth grade is only about half the size as in eighth grade (Panel C of Appendix Table A5). This is again consistent with a cumulating impact of patience, but a variety of other reasons could also be at play.

4.3 Robustness Analysis

Results prove stable for Italy and the United States in a series of robustness analyses. We summarize the various tests here and provide details in Appendix B.

Both in Italy and the United States, there are no significant overall gender differences. The impact of patience on reading is very similar to that for math, albeit with slightly smaller point estimates. All prior results hold not only for the different pre-COVID assessment waves separately but also for the post-COVID period. Additionally, when we consider other preferences from GPS, only trust is reliably estimated by Facebook interests, but including trust in the models leaves the patience results unchanged.

A potential methodological concern is that the 1,000 Facebook interests used to construct the patience measures contain interests that are directly related to education, possibly introducing endogeneity in the Facebook-derived patience measure to educational outcomes. We exclude 17 Facebook interests directly related to education from the construction of our patience measure. The alternative subnational patience measures are correlated above 0.998 both for Italian regions and U.S. states, and the estimated impact of measured patience on achievement hardly changes.

Using the individual-level data in Italy, we can confirm that the estimated impact of patience is larger for native students than for migrant students. Results also hold when excluding Trentino-Alto-Adige whose sample is not representative for the entire region and whose German-language population might limit comparability. Italy also participated with a regionally representative sample in the international PISA test in 2012, and we find that results hold equally well with this alternative achievement test. Finally, results are also robust in an analysis of unobservable selection and coefficient stability proposed by Oster (2019).

4.4 Exploration into Causality: Patience in Ancestry Countries of U.S. State Populations

The previous analysis of patience and subnational differences in human capital has been descriptive. Concerns about causal interpretation could emerge if patience of the regional population is affected by the region's human capital (reverse causation) or if regional patience is jointly determined with other regional factors that are important for educational achievement (omitted variables). For the cross-country analysis, our investigation of migrants in Section 3.2 addresses these main concerns of endogeneity. The within-country analysis is less prone to bias that may arise from national factors such as languages, laws, and institutional settings, and our robustness analysis showed that the subnational results are robust to conditioning on other cultural traits. But other threats to identification remain.

To explore further the empirical relevance of any remaining endogeneity bias, we look into the ancestry of U.S. residents. The ancestors of current U.S. state populations migrated from different countries, suggesting that time preferences of current residents have varying historical roots (Galor and Özak (2016); Becker, Enke, and Falk (2020)) that may partly be handed down from generation to generation within families (Bisin and Verdier (2011); Alesina and Giuliano (2014)). This suggests that the levels of patience of the countries from which the ancestors migrated can potentially proxy the patience of each state's current population while neither being affected by nor jointly determined with other current characteristics of U.S. states. We therefore employ an identification strategy that uses patience in the ancestry countries of U.S. states' populations as an instrumental variable (IV) for our Facebook-derived measure of patience in the U.S. states.

Demographic data from the American Community Survey (ACS) include ancestry or ethnic origin information for the 2022 population of each U.S. state (IPUMS, Ruggles et al. (2023)). The ancestry-based measure of patience for each U.S. state is developed by assigning respondents the patience level of the respective ancestry country from the GPS. We calculate shares for each ancestry country by U.S. state and weight the ancestry patience levels accordingly.[20]

The reduced-form estimate shows that states whose population migrated from ancestry countries with higher patience levels on average have significantly higher test scores (Table 3, column 1). In the first stage of the IV model, the ancestry-based patience instrument (measured directly from the GPS data) strongly predicts patience as measured by our Facebook-derived measure, independent of the number of PCs of Facebook interests used (columns 2-4). This first-stage association provides further credence to our derivation of patience measures from Facebook interests.

The second stage of the IV model provides statistically significant estimates of the impact of patience that are of the same order of magnitude as the OLS estimates. For the different numbers of PCs, the IV point estimates range from 0.22 to 0.37 (columns 5-7), quite similar to our baseline model (0.17 to 0.29).

While fully eliminating all concerns of causal interpretation is difficult in settings without experimental manipulation and while the IV model cannot rule out all potential biases, its results provide no indication that the baseline estimates are biased upwards by endogeneity from reverse causation or joint determination with other current state characteristics.

[20] For about a third of ACS respondents, the ancestry information is either missing or cannot be assigned to a specific country because it refers to an ethnic group (e.g., Kurdish) or a larger region (e.g., Eastern Europe). We assign these respondents the average U.S. patience level from the GPS and control for the share of missing ancestry information in the regressions. Results are very similar when coding missing ancestry information as missing, which is equivalent to assigning the average of the observed ancestries of the respective state (not shown).

4.5 Regional Analysis in Additional Countries

While we have focused on Italy and the United States as two countries with interesting regional variation and consistent test data at different grades for a substantial number of regions, we can assess the stability of our results by extending the analysis to six additional countries with publicly available subnational test data. We leverage regional indicators in the PISA data since 2012 for all countries with at least ten regions: Canada and Spain in 2012, 2015, and 2018, Brazil and Mexico in 2012, and Kazakhstan in 2018. Also, the Institut zur Qualitätsentwicklung im Bildungswesen (IQB) provides regionally representative math achievement data for German ninth-grade students in 2012 and 2018. For each country, we implement the method described in Section 2.3 to obtain regional measures of patience from Facebook interests, consistently using only 3 PCs because of the small number of regions in some countries.[21]

The consistency of results across these additional countries supports the methodology for investigating achievement differences within countries. In the pooled model of 190 regions in eight countries, the highly significant patience coefficient suggests that a one SD increase in patience is associated with a 0.34 SD increase in math scores (Table 4, column 1). Country-specific results are more tentative due to the limited regional information in several countries, but separate regressions show a positive regional association between patience and achievement that is statistically significant in each country except Spain (Table 4, columns 2-9 and Appendix Figure A11). The magnitude of coefficient estimates varies considerably across countries, suggesting that the strength of the relationship might depend on country-specific features, but there are too few country observations to analyze these differences systematically.

[21] Pooled results are similar using more PCs, but country-specific results are not stable at higher numbers of PCs in Brazil, Canada, and Germany.

5. Conclusions

Time preferences, while clearly important to individual investment decisions, have an even broader impact on education decisions. Aggregate preferences, which are a component of cultural identities, also affect political perspectives and community decisions about educational institutions such as the definition and importance of school quality.

This analysis investigates the importance of patience in determining historically significant but largely unexplained regional differences in student skills. These skill differences in turn have lasting consequences for incomes and for regional growth patterns.

We use the extensive compilations of social media information by Facebook to estimate preference differences for subnational regions within Italy and the United States. The measures of patience constructed from Facebook interests are validated by international comparisons where direct measures of time preferences are available.

Differences in patience across regions in Italy and across states in the United States provide a powerful explanation of human capital outcomes. This new perspective on student performance helps to explain why, for example, North-South differences in student outcomes in both countries have been very stable over time even in the face of national efforts to equalize performance. Causal identification of the preference-achievement relationship across subnational regions is particularly challenging, but an instrumental-variable approach that exploits historical ancestry variation across U.S. states supports a causal interpretation.

Our findings imply that similar educational inputs can lead to substantially different outcomes due to differences in patience. When addressing within-country differences in student achievement, policymakers might look beyond such proximate factors as school spending or even family educational background to take possible differences in patience into account.

Institutional features of schooling, such as reliance on parental choice or test-based accountability, appear less tied to aggregate preferences (Hanushek et al. (2022)). Thus, institutional reforms of school systems appear to be a viable policy mechanism for improvement that does not necessarily depend on changing preferences (Woessmann (2016)). Moreover, while cultural traits are considered hard to change (e.g., Guiso, Sapienza, and Zingales (2006); Bisin and Verdier (2011)), evidence shows that traits such as patience are malleable, especially at a young age, and can be improved through specific interventions (e.g., Bird (2001); Alan and Ertac (2018); Jung, Bharati, and Chin (2021)). Hence, policies aimed at increasing patience may also be an avenue for addressing educational investments and regional deficits in student outcomes.

References

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

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

Angerer, Silvia, Jana Bolvashenkova, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter (2023). Children's Patience and School-Track Choices Several Years Later: Linking Experimental and Field Data. *Journal of Public Economics* 220: 104837.

Bailey, Michael, Drew M. Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel, Johannes Stroebel (2022). The Social Integration of International Migrants: Evidence from the Networks of Syrians in Germany. NBER Working Paper 29925. Cambridge, MA: National Bureau of Economic Research.

Becker, Anke, Benjamin Enke, Armin Falk (2020). Ancient Origins of the Global Variation in Economic Preferences. *AEA Papers and Proceedings* 110: 319-323.

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

Benjamin, Daniel J., Sebastian A. Brown, Jesse M. Shapiro (2013). Who Is 'Behavioral'? Cognitive Ability and Anomalous Preferences. *Journal of the European Economic Association* 11 (6): 1231-1255.

Bigoni, Maria, Stefania Bortolotti, Marco Casari, Diego Gambetta (2019). At the Root of the North–South Cooperation Gap in Italy: Preferences or Beliefs? *Economic Journal* 129 (619): 1139-1152.

Bigoni, Maria, Stefania Bortolotti, Marco Casari, Diego Gambetta, Francesca Pancotto (2016). Amoral Familism, Social Capital, or Trust? The Behavioural Foundations of the Italian North–South Divide. *Economic Journal* 126 (594): 1318-1341.

Bird, Edward J. (2001). Does the Welfare State Induce Risk-Taking? *Journal of Public Economics* 80 (3): 357-383.

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

Cabañas, José González, Ángel Cuevas, Rubén Cuevas (2018). Facebook Use of Sensitive Data for Advertising in Europe. arXiv:1802.05030.

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

Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, et al. (2022). Social Capital I: Measurement and Associations with Economic Mobility. *Nature* 608 (7921): 108-121.

De Paola, Maria, Francesca Gioia (2017). Impatience and Academic Performance. Less Effort and Less Ambitious Goals. *Journal of Policy Modeling* 39 (3): 443-460.

Dee, Thomas S., Brian A. Jacob (2011). The Impact of No Child Left Behind on Student Achievement. *Journal of Policy Analysis and Management* 30 (3): 418-446.

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

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

Falorsi, Piero Demetrio, Patrizia Falzetti, Roberto Ricci (2019). *Le Metodologie Di Campionamento E Scomposizione Della Devianza Nelle Rilevazioni Nazionali Dell'invalsi: Le Rilevazioni Degli Apprendimenti A.S. 2018-2019*. Milano: Franco Angeli.

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

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

Galor, Oded, Ömer Özak, Assaf Sarid (2020). Linguistic Traits and Human Capital Formation. *AEA Papers and Proceedings* 110: 309-313.

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

Guiso, Luigi, Paola Sapienza, Luigi Zingales (2004). The Role of Social Capital in Financial Development. *American Economic Review* 94 (3): 526-556.

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

Hanushek, Eric A. (2016). What Matters for Achievement: Updating Coleman on the Influence of Families and Schools. *Education Next* 16 (2): 22-30.

Hanushek, Eric A., Lavinia Kinne, Philipp Lergetporer, Ludger Woessmann (2022). Patience, Risk-Taking, and Human Capital Investment across Countries. *Economic Journal* 132 (646): 2290-2307.

Hanushek, Eric A., Margaret E. Raymond (2005). Does School Accountability Lead to Improved Student Performance? *Journal of Policy Analysis and Management* 24 (2): 297-327.

Hanushek, Eric A., Jens Ruhose, Ludger Woessmann (2017). Knowledge Capital and Aggregate Income Differences: Development Accounting for U.S. States. *American Economic Journal: Macroeconomics* 9 (4): 184-224.

Hanushek, Eric A., Ludger Woessmann (2015). *The Knowledge Capital of Nations: Education and the Economics of Growth*. Cambridge, MA: MIT Press.

Ichino, Andrea, Giovanni Maggi (2000). Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm. *Quarterly Journal of Economics* 115 (3): 1057-1090.

Jung, Dawoon, Tushar Bharati, Seungwoo Chin (2021). Does Education Affect Time Preference? Evidence from Indonesia. *Economic Development and Cultural Change* 69 (4): 1451-1499.

Marty, Robert, Alice Duhaut (2024). Global Poverty Estimation Using Private and Public Sector Big Data Sources. *Scientific Reports* 14 (1): 3160.

Michaeli, Moti, Marco Casari, Andrea Ichino, Maria De Paola, Ginevra Marandola, Vincenzo Scoppa (2023). Civicness Drain. *Economic Journal* 133 (649): 323-354.

Obradovich, Nick, Ömer Özak, Ignacio Martín, Ignacio Ortuño-Ortín, Edmond Awad, et al. (2022). Expanding the Measurement of Culture with a Sample of Two Billion Humans. *Journal of The Royal Society Interface* 19 (190): 20220085.

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

Polachek, Solomon W., Tirthatanmoy Das, Rewat Thamma-Apiroam (2015). Micro- and Macroeconomic Implications of Heterogeneity in the Production of Human Capital. *Journal of Political Economy* 123 (6): 1410-1455.

Putnam, Robert D. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rogers, Megan Schouweiler (2023). IPUMS USA: Version 14.0 [Dataset]. https://doi.org/10.18128/D010.V14.0. Minneapolis, MN: IPUMS.

Sunde, Uwe, Thomas Dohmen, Benjamin Enke, Armin Falk, David Huffman, Gerrit Meyerheim (2022). Patience and Comparative Development. *Review of Economic Studies* 89 (5): 2806-2840.

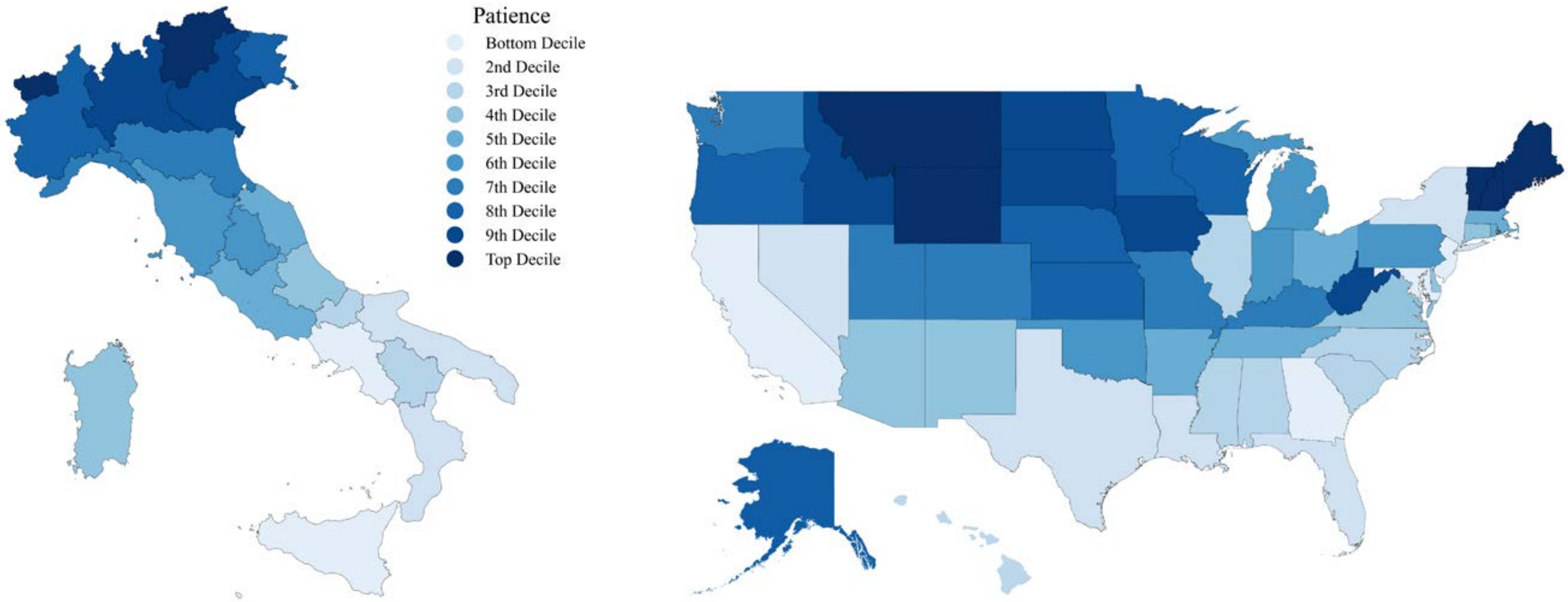
Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rützler, Stefan T. Trautmann (2013). Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior. *American Economic Review* 103 (1): 510-531.

Thorson, Kjerstin, Kelley Cotter, Mel Medeiros, Chankyung Pak (2021). Algorithmic Inference, Political Interest, and Exposure to News and Politics on Facebook. *Information, Communication & Society* 24 (2): 183-200.

Wilson, Robert E., Samuel D. Gosling, Lindsay T. Graham (2012). A Review of Facebook Research in the Social Sciences. *Perspectives on Psychological Science* 7 (3): 203-220.

Woessmann, Ludger (2010). Institutional Determinants of School Efficiency and Equity: German States as a Microcosm for OECD Countries. *Journal of Economics and Statistics* 230 (2): 234-270.

Figure 1: Measure of Patience Derived from Facebook Interests for Italian Regions and U.S. States



A. Italian Regions **B. U.S. States**

Notes: The figure shows maps of the Facebook-derived measure of patience obtained with 4 PCs for Italian regions (Panel A) and U.S. states (Panel B), respectively. Each color corresponds to a decile of the distribution of patience within each country. Darker colors denote higher levels of patience.

Table 1: Patience, Risk-taking, and Student Achievement: Cross-Country Validation Exercise

	GPS measure	Facebook measure (10 PCs)		
		Original sample	Extended sample	Non-GPS sample
	(1)	(2)	(3)	(4)
A. Cross-country analysis				
Patience	1.225***	1.684***	1.722***	1.771***
	(0.132)	*(0.135)*	*(0.119)*	*(0.210)*
Risk-taking	-1.229***	-1.359***	-1.537***	-1.660***
	(0.188)	*(0.310)*	*(0.254)*	*(0.388)*
Control variables	Yes	Yes	Yes	Yes
Observations	1,954,840	1,954,840	2,660,408	705,568
Residence countries	48	48	80	32
R^2	0.200	0.210	0.220	0.241
B. Migrant analysis				
Patience	0.957***	0.805***	0.902***	1.766***
	(0.115)	*(0.182)*	*(0.205)*	*(0.481)*
Risk-taking	-0.315**	-0.677**	-1.221***	-3.531***
	(0.124)	*(0.278)*	*(0.350)*	*(0.549)*
Control variables	Yes	Yes	Yes	Yes
Residence-country by wave fixed effects	Yes	Yes	Yes	Yes
Observations	78,403	78,403	90,983	12,580
Countries of origin	56	56	93	37
Residence countries	46	46	50	34
R^2	0.280	0.272	0.298	0.310

Notes: Dependent variable: PISA math test score. Least squares regressions. Panel A: all PISA waves 2000-2018; weighted by students' sampling probability. Panel B: waves 2003-2018; students with both parents not born in the country where the student attends school; including 180 fixed effects for each residence-country by wave cell. Control variables: Panel A: student gender, age, and migration status; imputation dummies; and wave fixed effects; Panel B: student gender, age, dummy for OECD country of origin, imputation dummies. Robust standard errors adjusted for clustering at the country level (migrant analysis: country of origin) in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2000-2018; Falk et al. (2018); own elaboration of Facebook data.

Table 2: Patience and Student Achievement: Subnational Analysis for Italy and the United States

	4 PCs	7 PCs	10 PCs
	(1)	(2)	(3)
A. Italy (individual level)			
Patience	1.614***	1.398***	1.495***
	(0.173)	*(0.107)*	*(0.107)*
Control variables	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes
Observations	59,034	59,034	59,034
Regions	20	20	20
R^2	0.095	0.100	0.101
B. Italy (regional level)			
Patience	1.370***	1.201***	1.284***
	(0.169)	*(0.086)*	*(0.089)*
Wave fixed effects	Yes	Yes	Yes
Observations	40	40	40
Regions	20	20	20
R^2	0.771	0.843	0.852
C. United States (state level)			
Patience	0.293***	0.172*	0.291**
	(0.089)	*(0.096)*	*(0.131)*
Wave fixed effects	Yes	Yes	Yes
Observations	153	153	153
Regions	51	51	51
R^2	0.360	0.348	0.364

Notes: Dependent variable: Panels A and B: INVALSI eighth-grade math test score in waves 2018 and 2019; Panel C: NAEP eighth-grade math test score in all NAEP waves 2015-2019. Least squares regressions with wave fixed effects. Unit of observation: Panel A: student; Panel B: region-wave combination; Panel C: state-wave combination. Patience measured at the regional/state level throughout. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. Additional control variables (Panel A): student gender, age, and migration status; imputation dummies. The Italian region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional (state) level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI mathematics achievement test, 2017-2019; NAEP mathematics achievement test, 2015-2019; own elaboration of Facebook data.

Table 3: Patience and Student Achievement across U.S. States: Using Ancestry Information as Instruments

	Reduced form	First stage			Second stage		
		4 PCs	7 PCs	10 PCs	4 PCs	7 PCs	10 PCs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ancestry patience	0.945***	3.617***	3.605***	2.143***			
	(0.349)	*(0.306)*	*(0.283)*	*(0.229)*			
Patience					0.261***	0.221**	0.372**
					(0.096)	*(0.090)*	*(0.156)*
Wave fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap Wald *F* statistic		139.321	161.733	87.387			
Observations	153	153	153	153	153	153	153
Regions	51	51	51	51	51	51	51
(Centered) R^2	0.518				0.503	0.502	0.469

Notes: Dependent variable: col. 1 and 5-7: NAEP eighth-grade math test score in all NAEP waves 2015-2019; col. 2-4: patience. Unit of observation: state-wave combination. Patience measured at the state level throughout. Column headers indicate number of principal components (PCs) used to compute patience measure. Regressions control for risk-taking computed with the equivalent number of PCs and for the share of missing ancestry information. Robust standard errors adjusted for clustering at the state level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: NAEP mathematics achievement test, 2015-2019; ACS; own elaboration of Facebook data.

Table 4: Patience and Student Achievement: Subnational Analysis in Eight Countries

	Eight countries pooled	Brazil	Canada	Germany	Italy	Kazakhstan	Mexico	Spain	United States
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Patience (3 PCs)	0.337***	1.556***	0.383**	2.093***	1.527***	0.459**	0.666***	0.060	0.218*
	(0.085)	*(0.206)*	*(0.155)*	*(0.624)*	*(0.205)*	*(0.194)*	*(0.160)*	*(0.108)*	*(0.113)*
Grade/age	–	Age 15	Age 15	Grade 9	Grade 8	Age 15	Age 15	Age 15	Grade 8
Wave fixed effects	Yes	No	Yes	Yes	Yes	No	No	Yes	Yes
Country fixed effects	Yes	No	No	No	No	No	No	No	No
Observations	383	27	30	32	40	16	32	51	153
Regions	190	27	10	16	20	16	32	17	51
R^2	0.288	0.719	0.726	0.577	0.747	0.178	0.517	0.362	0.300

Notes: Dependent variable: math test scores. Least squares regressions with wave fixed effects. Unit of observation: region-wave combination. Test and wave information: Brazil and Mexico: PISA 2012; Canada and Spain: PISA 2012, 2015, and 2018; Germany: IQB 2012 and 2018; Italy: INVALSI 2018 and 2019; Kazakhstan: PISA 2018; United States: NAEP 2015, 2017, and 2019. Regressions control for risk-taking computed with 3 PCs. Robust standard errors adjusted for clustering at the state level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA, IQB, INVALSI, and NAEP mathematics achievement tests; own elaboration of Facebook data.

Online Appendix

Appendix A: Data on Regional Student Achievement

This appendix describes the regionally representative assessment data used for Italy and the United States: INVALSI and NAEP, respectively.

A.1 Italy: INVALSI

Since 2007, the Istituto Nazionale per la Valutazione del Sistema Dell'Istruzione (INVALSI) assesses a random sample of Italian students in math and Italian every year. Furthermore, INVALSI administers student, teacher, and principal questionnaires to collect background information about the educational environment. We use data on math achievement in the school years 2017-2018 and 2018-2019, the last years before the COVID-19 pandemic. In our main analysis, we focus on the sample of 59,034 eighth-grade students because they are closest in age to the students in PISA and NAEP, but we subsequently expand the analysis to include students in grades 2, 5, and 10, yielding a sample size of 235,661 students.

The sample of students is drawn following a two-step procedure, where a varying number of classes is randomly selected within a random sample of schools stratified at the regional level. Crucially for our analysis, the sample is representative at the regional level for 19 of the 20 regions in Italy (Falorsi, Falzetti, and Ricci (2019)). The exception is Trentino-Alto-Adige, where only students in the autonomous municipalities of Bolzano and Trento are tested. The difference between the lowest and highest performing region in Italy in eighth-grade math amounts to roughly three quarters of a SD, equivalent to the average learning of almost three school years.

In robustness checks, we complement the INVALSI analysis using Italian data from PISA 2012 where Italy oversampled students in each region to obtain a representative sample.

A.2 United States: NAEP

We use data from the National Assessment of Educational Progress (Main-NAEP), the largest nationally representative assessment of students in the United States. In our primary analysis, we focus on NAEP mathematics test scores in grade eight, using data from the last three waves of NAEP before the COVID-19 pandemic, namely NAEP 2015, 2017, and 2019. The resulting dataset consists of state-level test scores for the 50 U.S. states and the federal district of Washington, DC. Approximately 140,000 students take part in a typical NAEP assessment.[22] In additional analyses, we also use data on fourth-grade students. In the United States, the difference between the lowest and highest performing state in eighth-grade math is equivalent to almost three years of schooling, similar to what was found in Italy.

We divide both INVALSI and NAEP test scores by the student-level SD in the respective country, so that regression coefficients can be interpreted in terms of SDs.

[22] Source: https://nces.ed.gov/nationsreportcard/guides/statsig.aspx#:~:text=A%20NAEP%20national%20assessment%20typically,samples%20of%20approximately%20140%2C000%20students (accessed 23 September 2024).

Appendix B: Robustness Analysis

Results prove stable in a series of robustness analyses. Both in Italy and the United States, we find similar results for reading achievement, albeit with slightly smaller point estimates. Results are robust in the separate assessment waves available in each country, as well as in post-COVID waves. Baseline results do not differ significantly by gender. They are also robust in models conditioning on trust as another cultural trait, as well as when excluding education-related Facebook interests from the derivation of the patience measure. From the individual-level data for Italy, we confirm that the estimates are larger for native students than for migrant students. The availability of individual-level data for Italy allows a more in-depth analysis than for the United States, where the analysis is constrained by the region-level data.

B.1 Italy

We exploit the within-country nature and the richness of the INVALSI data to replicate our analysis using reading outcomes. Our main analysis focuses on math achievement, which is generally considered the most comparable subject across countries, while student reading outcomes are inherently language-specific and less suitable for cross-country analysis. But, within our separate countries, the significant positive association of patience with student achievement also holds for reading. Results show that a one SD increase in patience is associated with a 1.04-1.33 SD increase in student reading achievement in the Italian individual-level analysis (Appendix Table A6). At the regional level, a one SD increase in patience is associated with an increase of 0.80-1.04 SD in reading scores. The magnitude of the coefficients in reading is slightly smaller than in math but results clearly show in both subjects.

Results are also very robust across subsamples of waves and gender. Results do not depend on the year in which the assessment was conducted, suggesting that they are not driven by the specific timing of the achievement observation (Appendix Table A7). While our main analysis refers to the last waves before the COVID-19 pandemic, qualitative results are the same for the available post-COVID waves 2021, 2022, and 2023 (Appendix Table A8). Results also hold similarly for girls and boys (Appendix Table A7), and the gender difference is not statistically significant.[23]

In line with a leading role of cultural traits as a deep determinant of student achievement, results are stronger for native students than for migrant students. A one SD increase in patience is associated with a 1.47-1.68 SD increase in achievement for native students, but a 0.71-0.97 SD (0.76-0.99 SD) increase for students with a second- (first-) generation migrant background (Appendix Table A9). This pattern would be expected if it were indeed patience as a cultural trait that drives the achievement results, as the culture of the residence region is presumably less important for migrant students who have been less exposed to the regional culture.[24]

To confirm that the association captures patience and not other cultural traits, we also condition on additional preference measures. Risk-taking, the other preference with intertemporal bearing, is already included in the baseline model. Among the other four preference parameters contained in the GPS, only trust is reasonably well predicted by Facebook interests (see footnote 9 in section 2.2 of the main text). Results are robust to including trust as an additional control variable (Appendix Table A10).

A potential concern with the analysis is that the 1,000 Facebook interests used to construct the patience measures contain some interests that are directly related to education, thereby possibly introducing endogeneity of the Facebook-derived patience measure to educational

[23] Reported results are based on Facebook-derived measures obtained with 4 PCs, but results are qualitatively the same with 7 and 10 PCs (not shown).

[24] Hanushek et al. (2022) find a similar pattern in their analysis of international student achievement.

outcomes. To test for this possibility, we identify 17 interests from the total list of employed Facebook interests (see Appendix Figure A1) that are related to education and exclude them from the derivation of the patience measure.[25] It turns out that the patience measures derived with and without the education-related interests are very strongly correlated (above 0.998 for different PCs both for Italian regions and U.S. states), indicating that the baseline measure is not strongly dependent on the inclusion of education-related interests. Consequently, results hardly change when using the alternative measure (Appendix Table A11), showing that the baseline analysis is not driven by endogeneity from education-related Facebook interests.

Results also hold when excluding Trentino-Alto-Adige whose sample is not representative for the entire region and whose German-language population might limit comparability. In the INVALSI test of this region, only students in the autonomous municipalities of Bolzano and Trento are tested (see Appendix A.1). This sampling restricted to municipal areas may bias our estimates, not least because Trentino-Alto-Adige is the Italian region with the highest estimated level of patience (see Section 2.3). Furthermore, we want to be sure that results are not driven by the Austrian history and the partially German-speaking population of the region. When omitting Trentino-Alto-Adige from the analysis, results are qualitatively the same and, if anything, slightly larger in magnitude (Appendix Table A12).

Results are also robust in an analysis of unobservable selection and coefficient stability proposed by Oster (2019). We compare our baseline model to a restricted model without control variables, following the standard procedure of setting $R_{max} = 1.3\tilde{R}$. Assuming an equal degree of selection between observables and unobservables, $\delta = 1$, the estimated bias-adjusted

[25] The 17 education-related Facebook interests are: books, college, e-books, education, educational technology, English language, high school, higher education, kindergarten, learning, literature, mathematics, reading, school, science, student, and university.

coefficient β^* for patience is between 1.53 and 1.76 (Appendix Table A13). In all cases, the bias-adjusted coefficient β^* is larger than our main estimates. The values δ for which $\beta = 0$ lie between -2.86 and -4.84. In all cases, these values are much larger (in absolute terms) than using the standard cutoff $\delta = 1$. These results imply that the selection on unobservables would need to be more than 2.6 times larger (and of opposite sign) than the selection on observables to push the coefficient of patience to 0.

Finally, we make use of the fact that Italy participated with a regionally representative sample in the international PISA test in 2012 to show that results hold equally well in this alternative achievement test. Intriguingly, the PISA results shown in Appendix Table A14 are very similar to the INVALSI results shown in Panel A of Table 2, indicating that a one SD increase in patience is associated with a 1.47-1.57 SD increase in the PISA math score.

B.2 United States

The results for U.S. states also replicate well for reading outcomes. The results closely mirror the findings for Italy: the magnitude of the coefficient of patience is slightly smaller compared to the analysis of math achievement (Appendix Table A15). A one SD increase in patience is associated with an increase of 0.14-0.23 SD in reading achievement. Again, this analysis confirms that results do not depend on a particular subject.

Results also do not depend on the specific year in which student achievement is observed. Results are qualitatively the same for all analyzed NAEP waves – 2015, 2017, and 2019 (Appendix Table A16). The magnitude of the patience coefficient tends to be smaller in the most recent wave, although not statistically significantly so. Similarly, results hold in the post-COVID wave 2022 (Appendix Table A8).

The U.S. results are also similar across genders. Patience is significantly positively associated with student achievement of both boys and girls (Appendix Table A17). The coefficient estimates are somewhat larger for boys than for girls, but not significantly so.

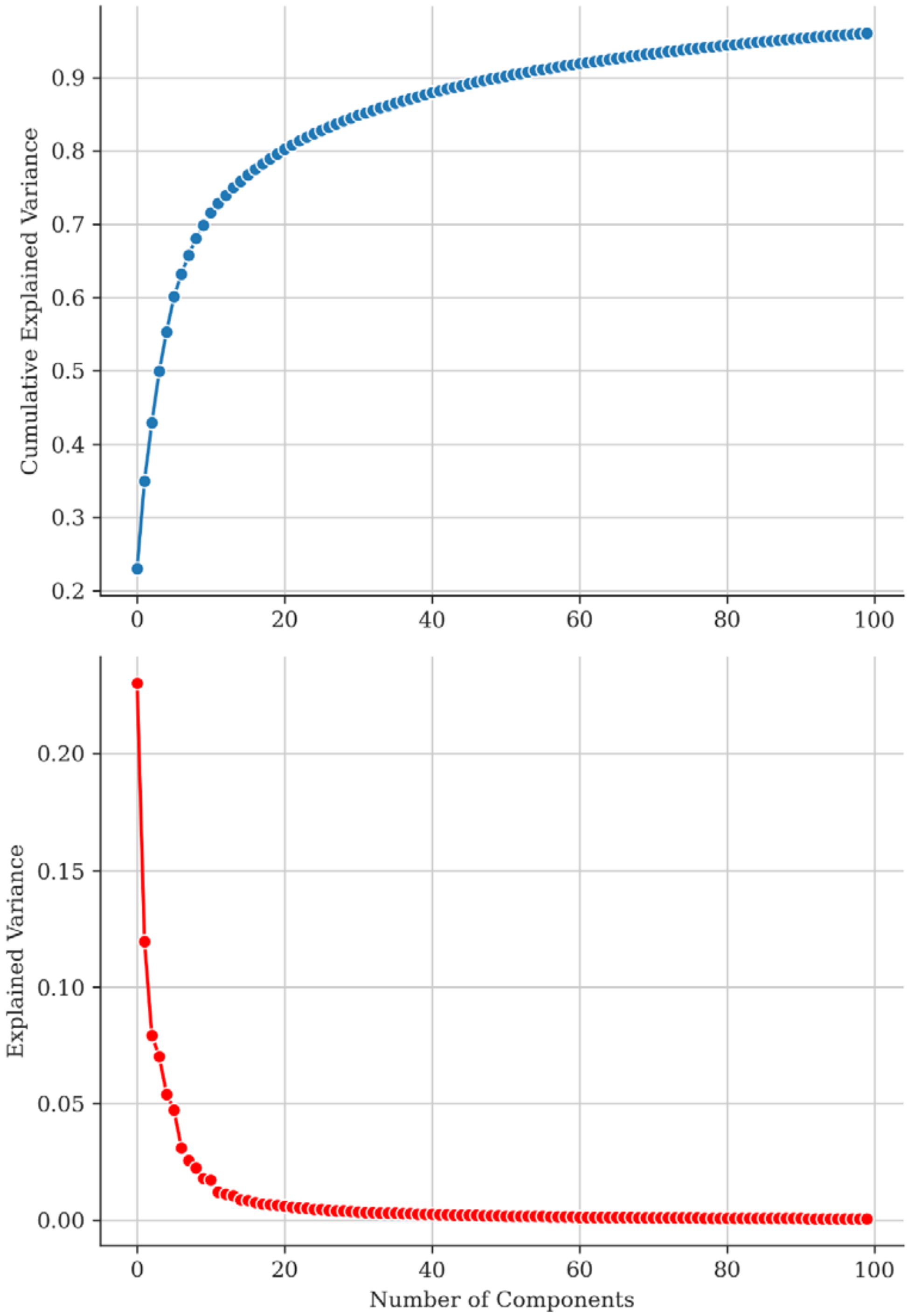
Finally, the U.S. results are also robust to conditioning on trust (Appendix Table A10) and to excluding education-related Facebook interests from the analysis (Appendix Table A11).

Appendix Figures and Tables

Figure A1: Word Cloud of Facebook Interests

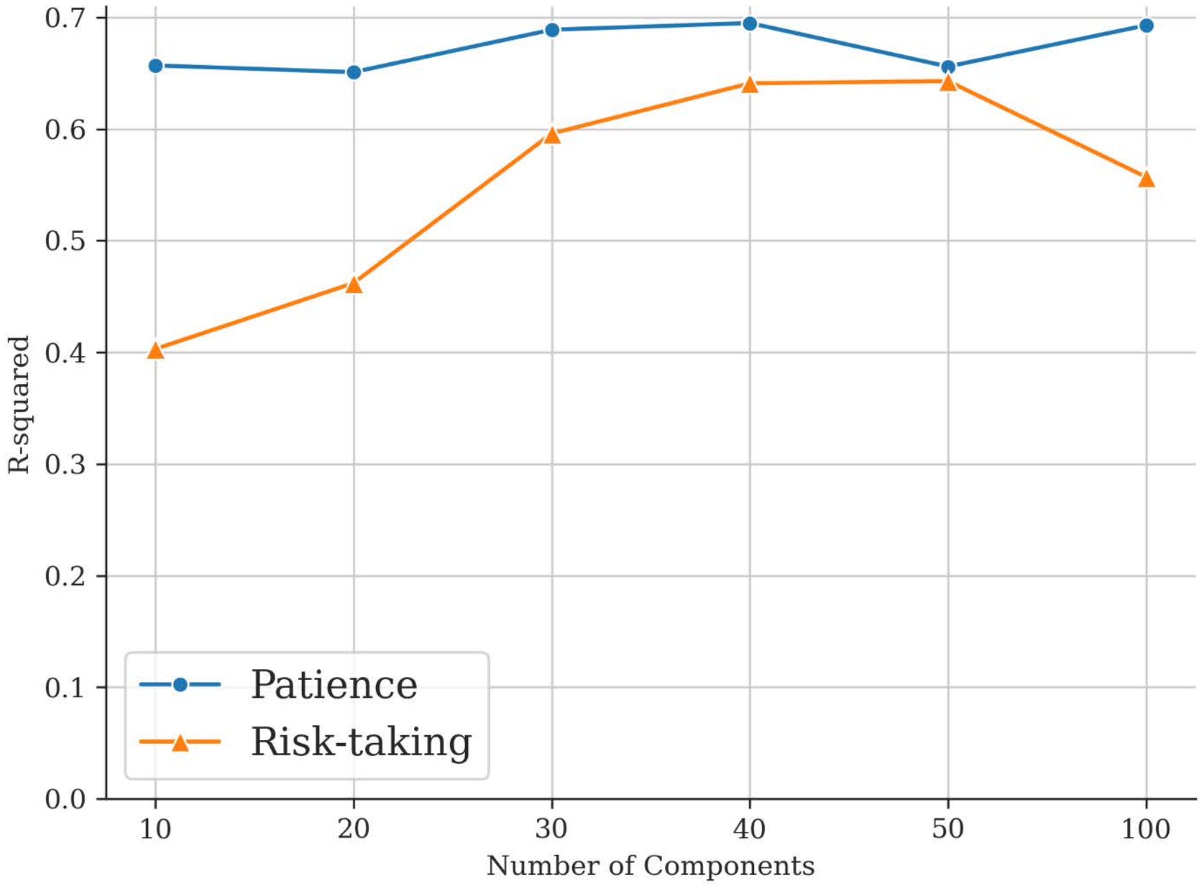
Notes: Word cloud of 1,000 Facebook interests with largest worldwide audience. Larger font sizes indicate larger audience.

Figure A2: Variance in Facebook Interests Captured by PCs: International Sample



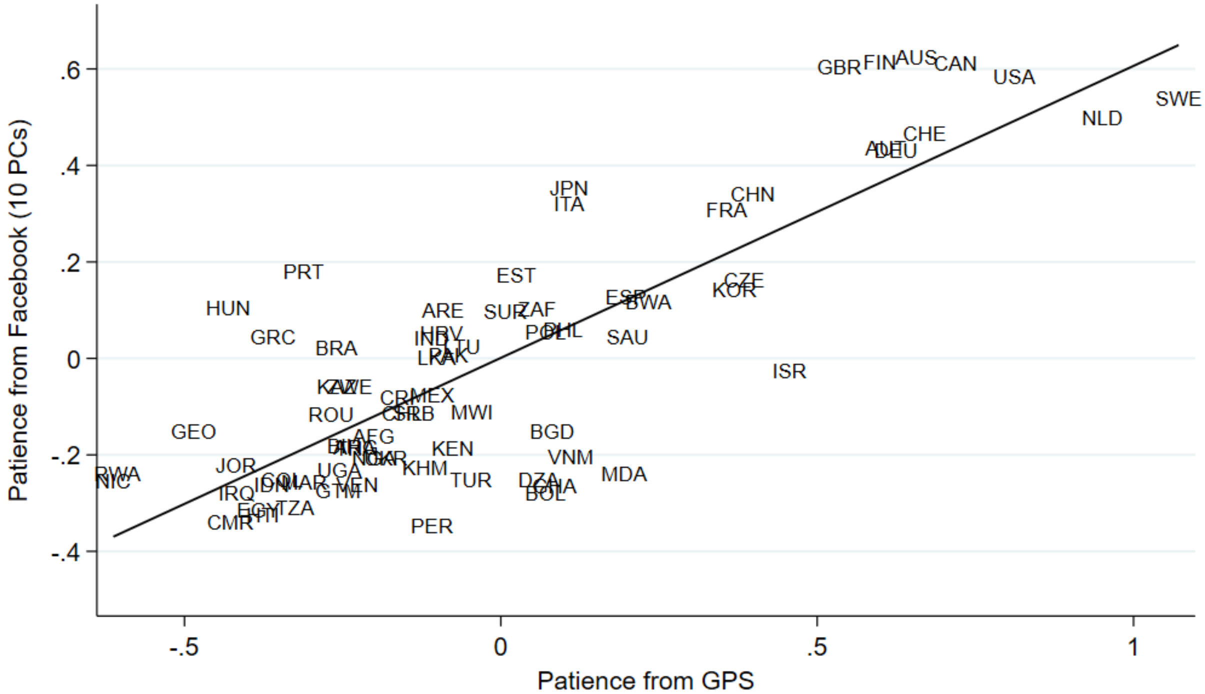
Notes: The top figure shows the cumulative variance in Facebook interests captured by the PCs of the Facebook interests in the international sample, the bottom figure shows the variance captured by each component.

Figure A3: Performance of GPS Prediction with Facebook Interests: International Sample

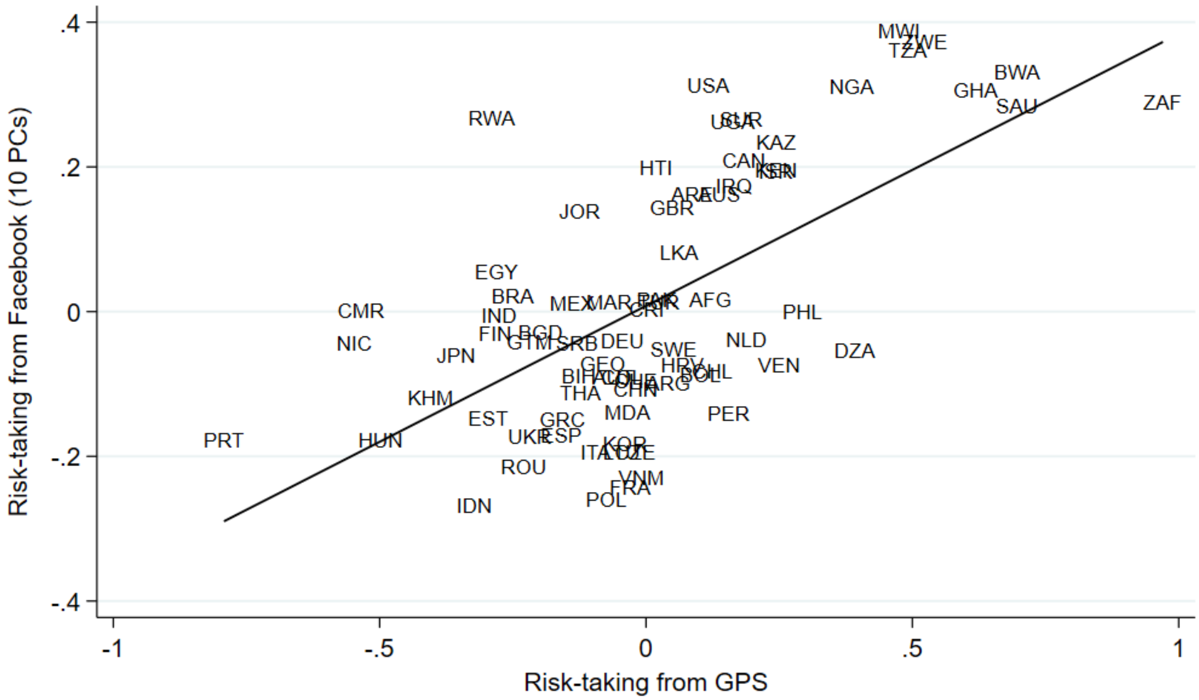


Notes: The figure shows the R^2 of regressions of the GPS measures of patience and risk-taking, respectively, on the PCs of Facebook interests (obtained with PC loadings of country-level Facebook interests) for different numbers of PCs used in the regression. 10-fold cross-validated LASSO model. Sample: all 74 countries for which GPS and Facebook data are available.

Figure A4: Measures of Patience and Risk-Taking: GPS and Prediction from Facebook



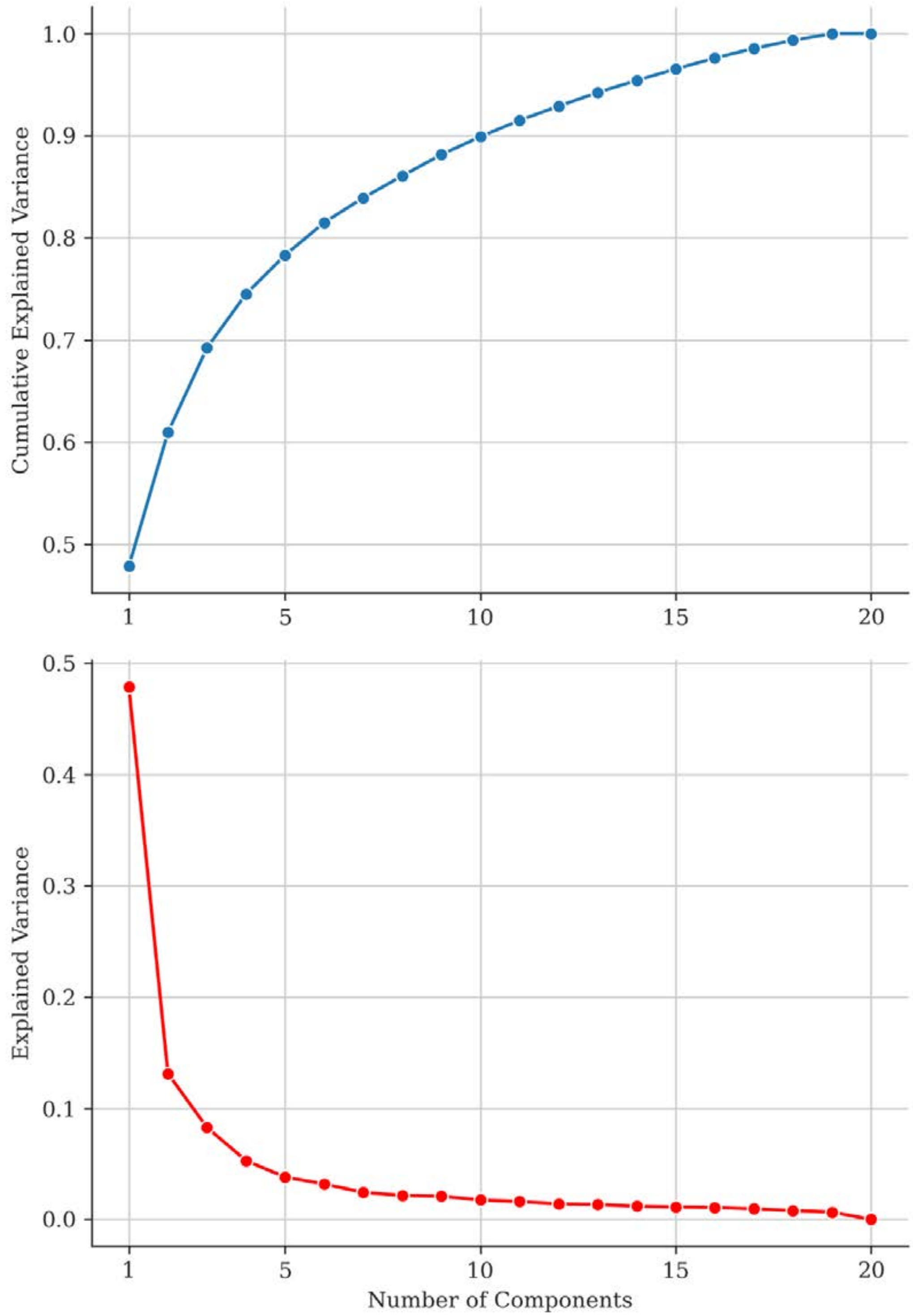
A. Patience



B. Risk-Taking

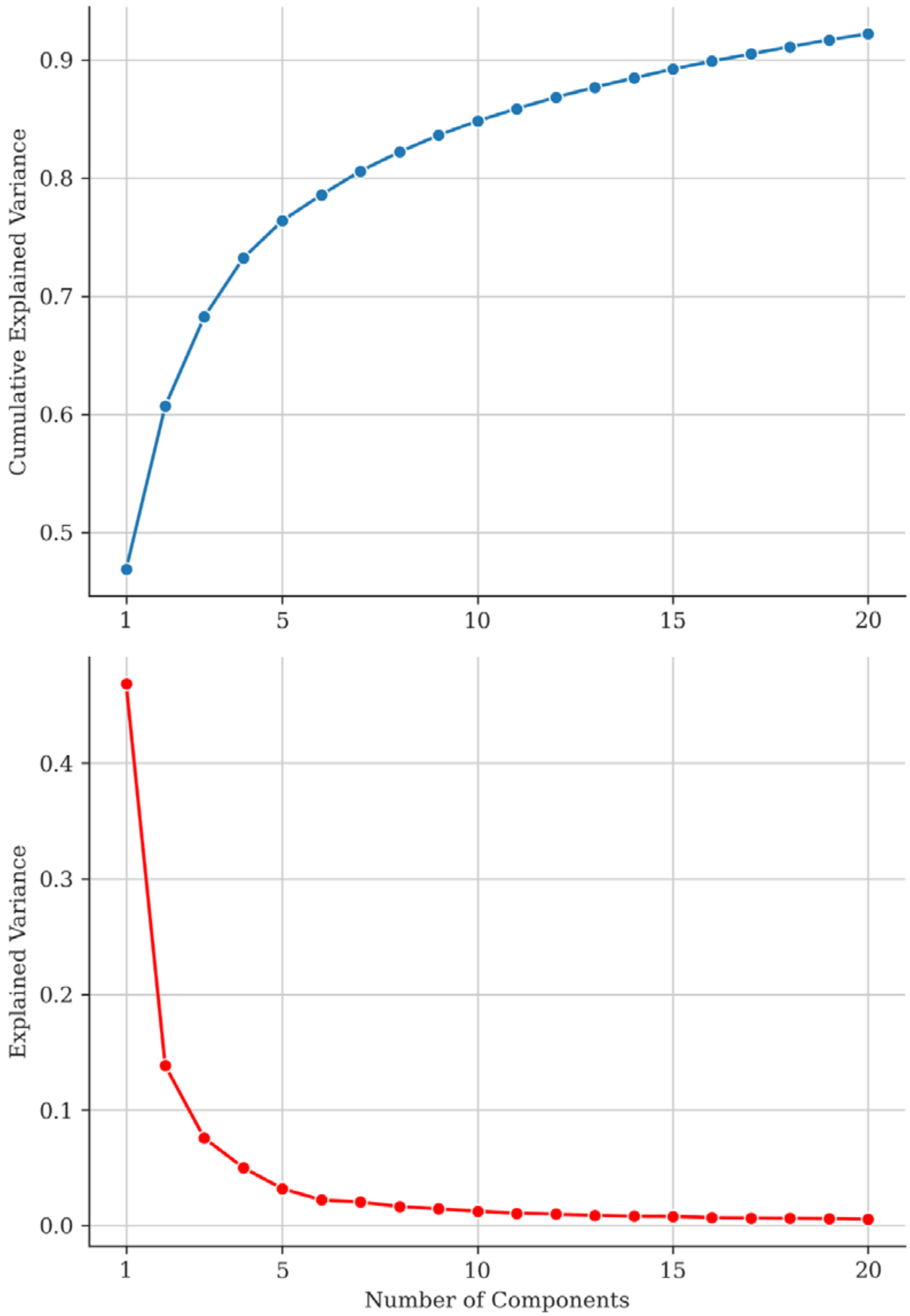
Notes: The figure shows scatterplots of the GPS measures of patience (Panel A) and risk-taking (Panel B), respectively, against the predictions of these measures from Facebook interests (10 PCs). Sample: all 74 countries for which GPS and Facebook data are available.

Figure A5: Variance in Facebook Interests Captured by PCs: Italian Regions



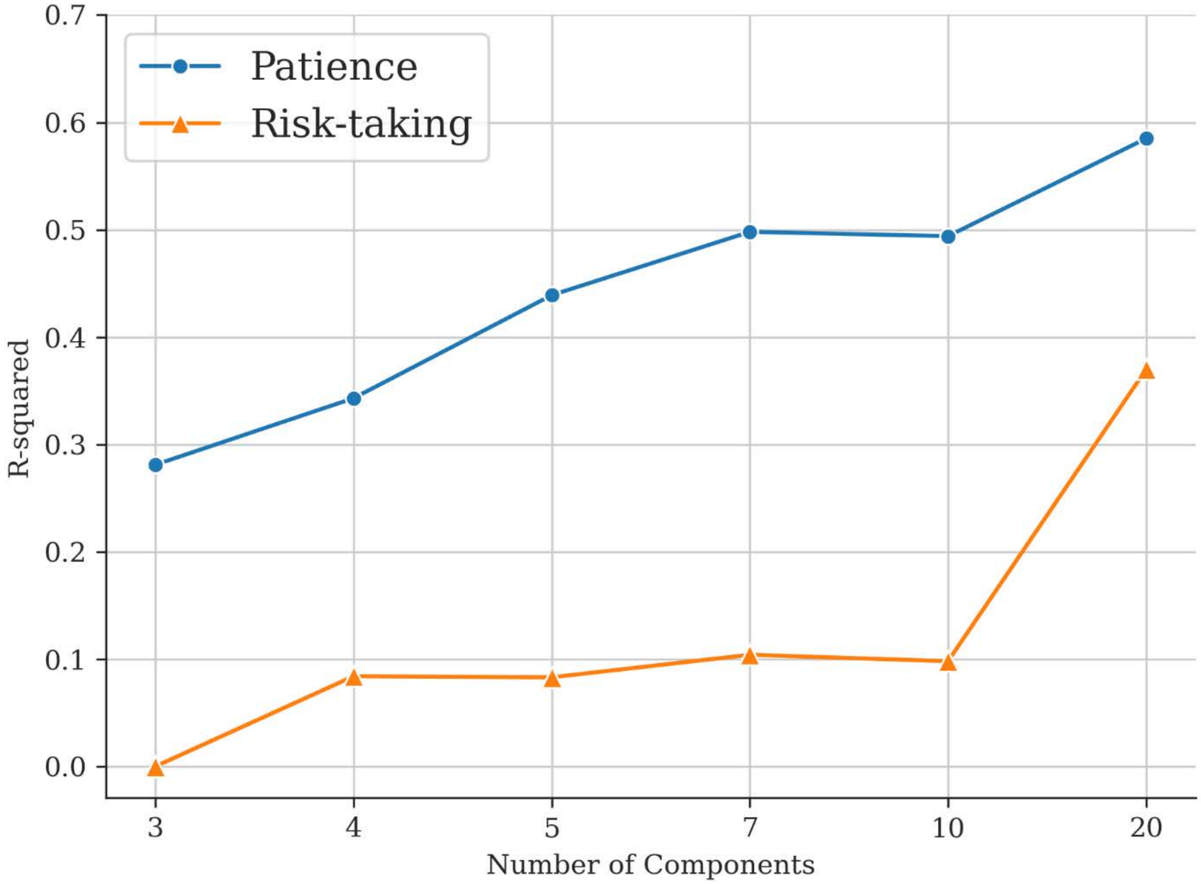
Notes: The top figure shows the cumulative variance in Facebook interests captured by the PCs of the Facebook interests in the Italian regions, the bottom figure shows the variance captured by each component.

Figure A6: Variance in Facebook Interests Captured by PCs: U.S. States



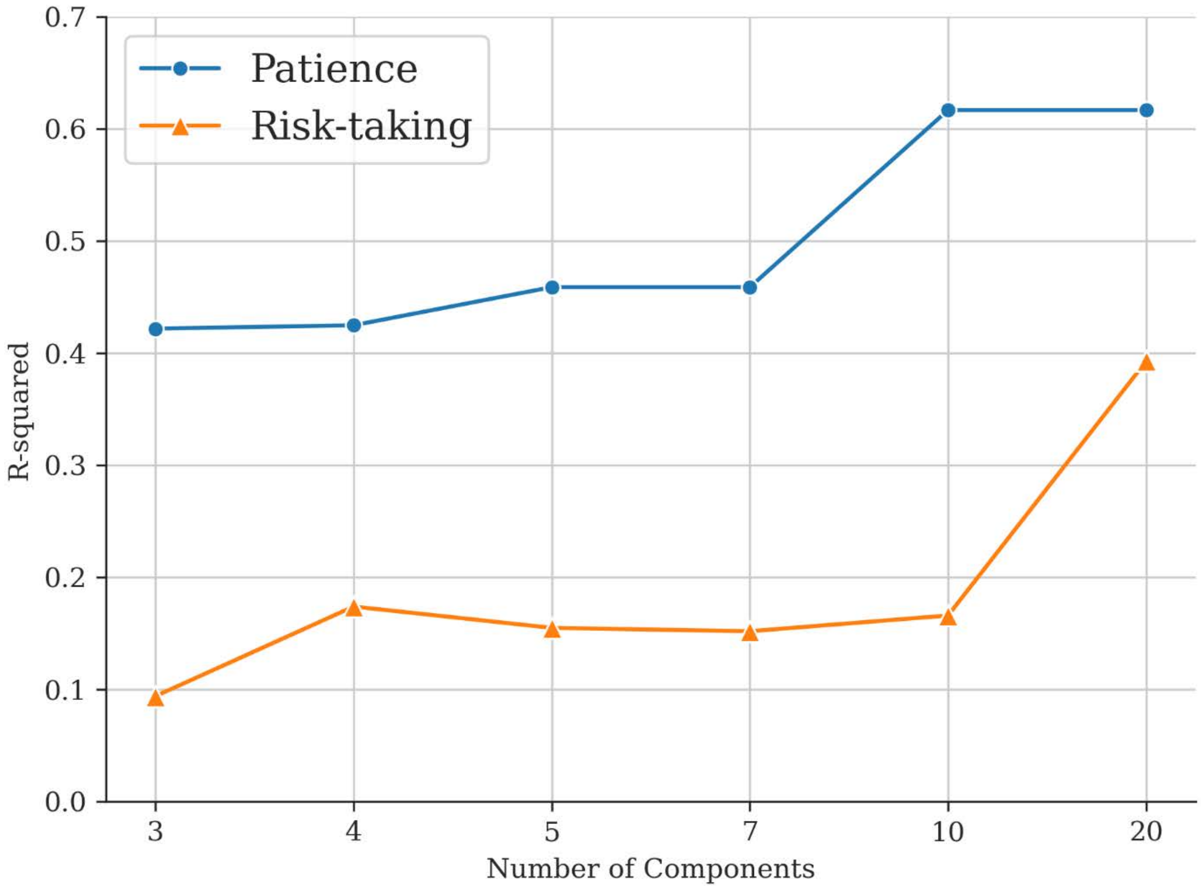
Notes: The top figure shows the cumulative variance in Facebook interests captured by the PCs of the Facebook interests in the U.S. states, the bottom figure shows the variance captured by each component.

Figure A7: Performance of GPS Prediction with Facebook Interests: PC Loadings from Italian Regions



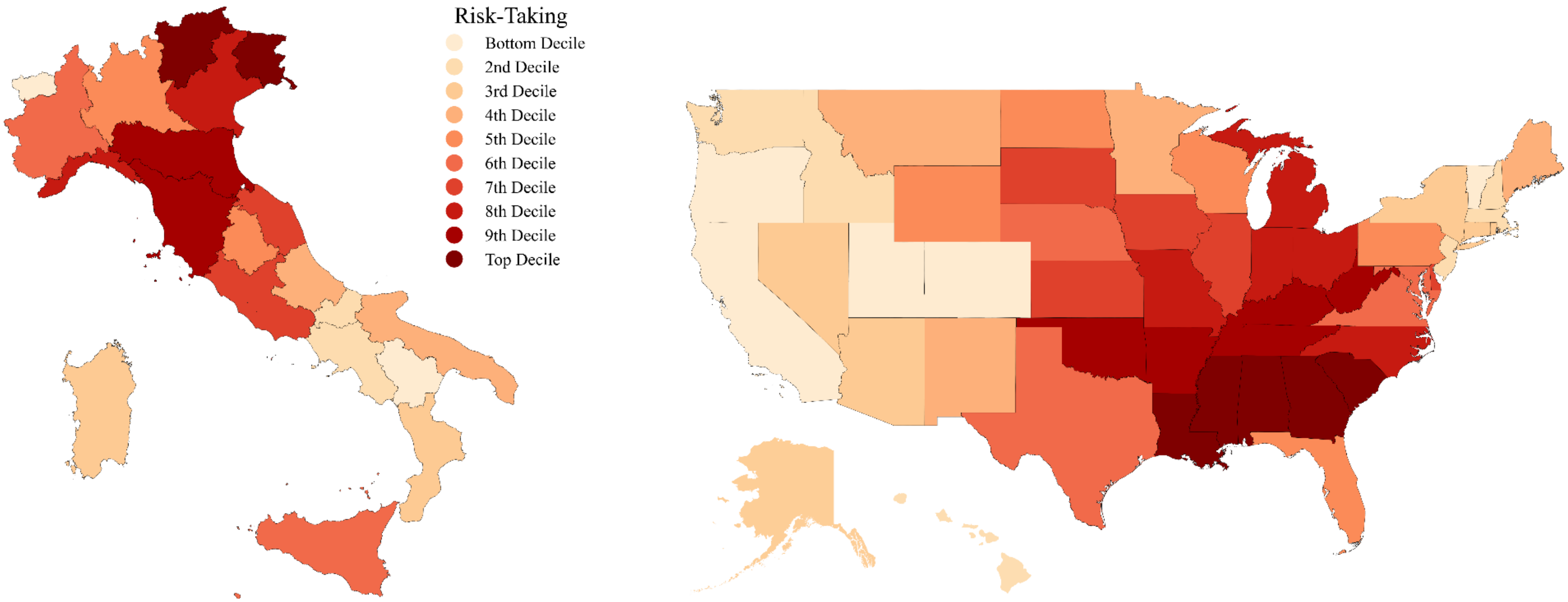
Notes: The figure shows the R^2 of regressions of the GPS measures of patience and risk-taking, respectively, on the PCs of Facebook interests (obtained with PC loadings of Italian-region-level Facebook interests) for different numbers of PCs used in the regression. 10-fold cross-validated LASSO model. Sample: all 74 countries for which GPS and Facebook data are available.

Figure A8: Performance of GPS Prediction with Facebook Interests: PC Loadings from U.S. States



Notes: The figure shows the R^2 of regressions of the GPS measures of patience and risk-taking, respectively, on the PCs of Facebook interests (obtained with PC loadings of U.S. state-level Facebook interests) for different numbers of PCs used in the regression. 10-fold cross-validated LASSO model. Sample: all 74 countries for which GPS and Facebook data are available.

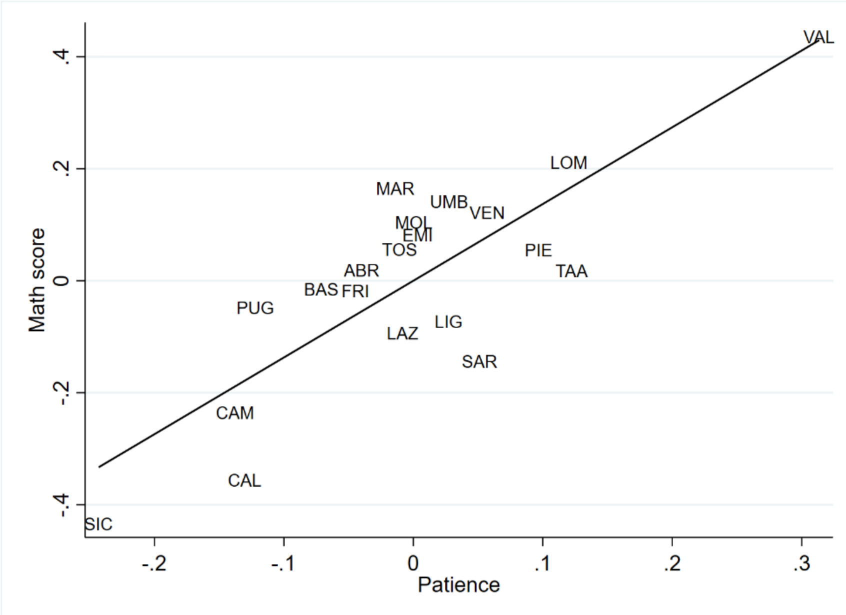
Figure A9: Measure of Risk-Taking Derived from Facebook Interests for Italian Regions and U.S. States



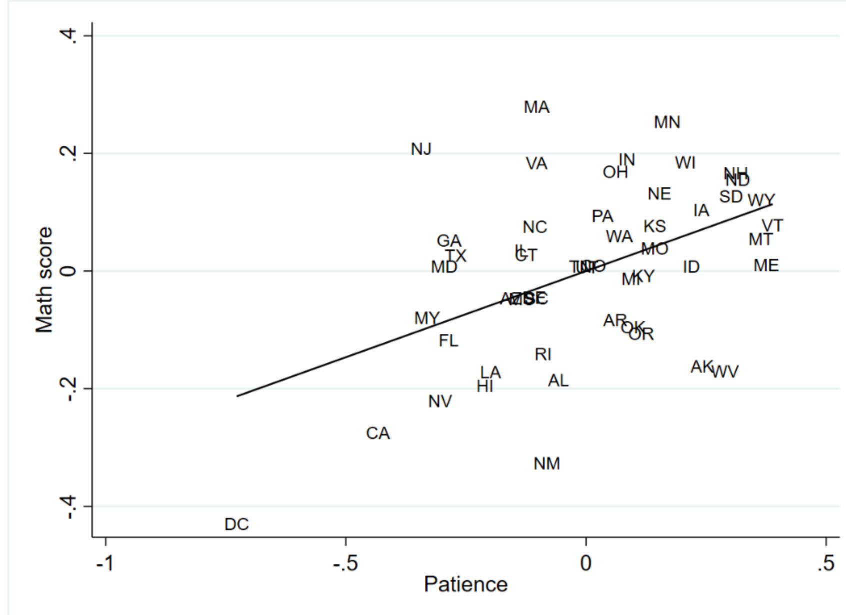
A. Italian Regions **B. U.S. States**

Notes: The figures show maps of the Facebook-derived measure of risk-taking obtained with 4 PCs for Italian regions (Panel A) and U.S. states (Panel B), respectively. Each color corresponds to a decile of the distribution of risk-taking within each country. Darker colors denote higher levels of risk-taking.

Figure A10: Patience and Student Achievement across Italian Regions and U.S. States



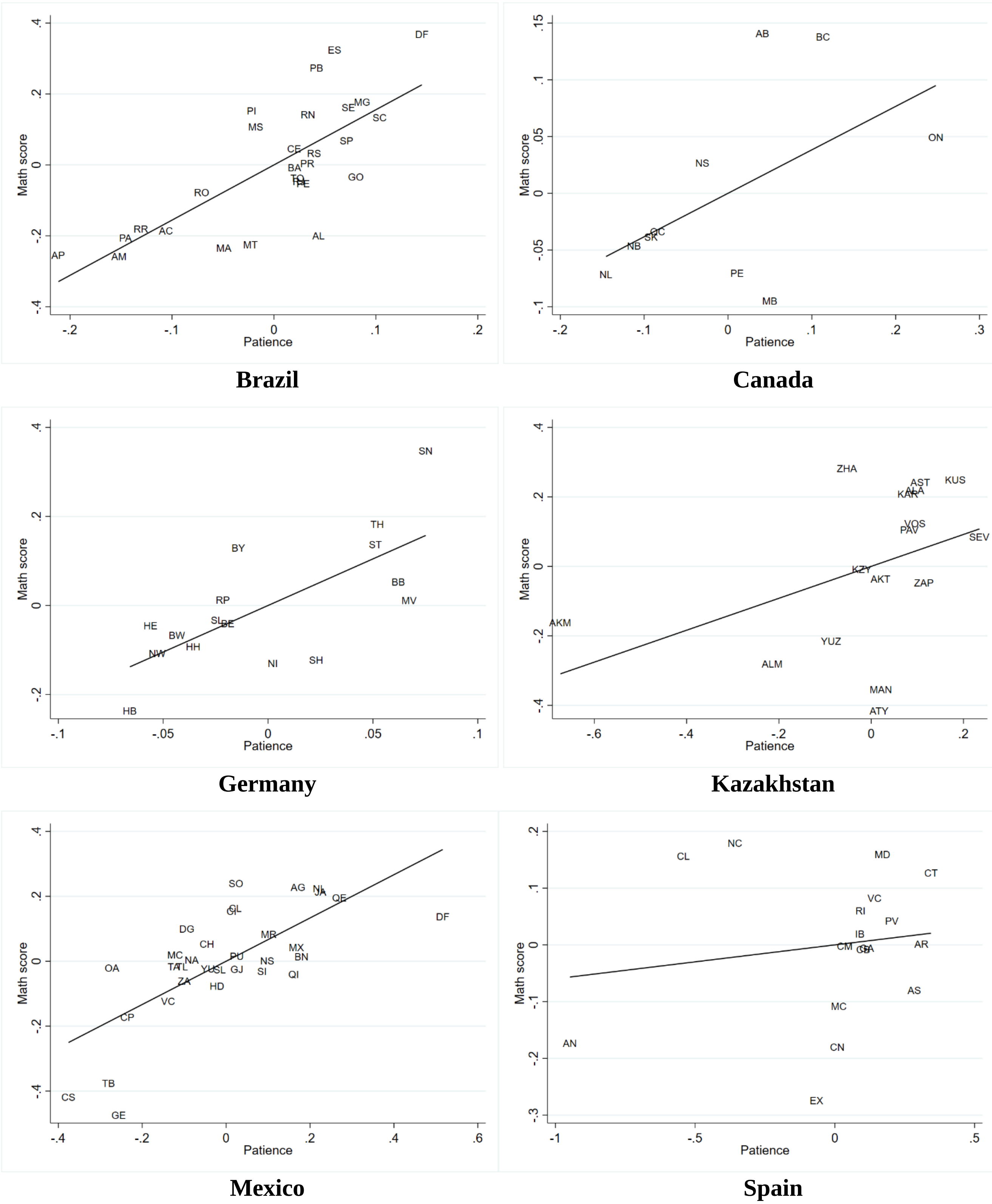
Italy



United States

Notes: Scatterplots of math test scores (wave averages) against patience across regions, both conditional on risk-taking. See Table 2 for underlying measures.

Figure A11: Patience and Student Achievement across Subnational Regions



Notes: Scatterplots of math test scores (wave averages) against patience across regions, both conditional on risk-taking. See Table 4 for underlying measures.

Table A1: Countries in the Cross-country Validation Exercise

	PISA countries			Training sample
	Only Facebook	Only GPS	Facebook and GPS	Facebook and GPS
	(1)	(2)	(3)	(4)
Afghanistan				x
Albania	x			
Algeria			x	x
Argentina			x	x
Australia			x	x
Austria			x	x
Azerbaijan	x			
Bangladesh				x
Belarus	x			
Belgium	x			
Bolivia				x
Bosnia and Herzegovina			x	x
Botswana				x
Brazil			x	x
Brunei Darussalam	x			
Bulgaria	x			
Cambodia				x
Cameroon				x
Canada			x	x
Chile			x	x
China				x
Colombia			x	x
Costa Rica			x	x
Croatia			x	x
Czech Republic			x	x
Denmark	x			
Dominican Republic	x			
Egypt				x
Estonia			x	x
Finland			x	x
France			x	x
Georgia			x	x
Germany			x	x
Ghana				x
Greece			x	x
Guatemala				x
Haiti				x
Hong Kong	x			
Hungary			x	x
Iceland	x			
India				x
Indonesia			x	x
Iraq				x
Ireland	x			
Israel			x	x
Italy			x	x
Japan			x	x
Jordan			x	x

(continued on next page)

Table A1 (continued)

	PISA countries			Training sample
	Only Facebook	Only GPS	Facebook and GPS	Facebook and GPS
	(1)	(2)	(3)	(4)
Kazakhstan			x	x
Kenya				x
Korea			x	x
Kyrgyzstan	x			
Latvia	x			
Lebanon	x			
Liechtenstein	x			
Lithuania			x	x
Luxembourg	x			
Macao	x			
Malawi				x
Malaysia	x			
Malta	x			
Mauritius	x			
Mexico			x	x
Moldova			x	x
Montenegro	x			
Morocco			x	x
Netherlands			x	x
New Zealand	x			
Nicaragua				x
Nigeria				x
North Macedonia	x			
Norway	x			
Pakistan				x
Panama	x			
Peru			x	x
Philippines			x	x
Poland			x	x
Portugal			x	x
Qatar	x			
Romania			x	x
Russia		x		
Rwanda				x
Saudi Arabia			x	x
Serbia			x	x
Singapore	x			
Slovakia	x			
Slovenia	x			
South Africa				x
Spain			x	x
Sri Lanka				x
Suriname				x
Sweden			x	x
Switzerland			x	x
Tanzania				x
Thailand			x	x
Trinidad and Tobago	x			

(continued on next page)

Table A1 (continued)

	PISA countries			Training sample
	Only Facebook	Only GPS	Facebook and GPS	Facebook and GPS
	(1)	(2)	(3)	(4)
Tunisia	x			
Turkey			x	x
Uganda				x
Ukraine			x	x
United Arab Emirates			x	x
United Kingdom			x	x
United States			x	x
Uruguay	x			
Venezuela				x
Vietnam			x	x
Zimbabwe				x
Total: 107 countries	32	1	48	74

Notes: Sample of countries: Col. 1-3: countries included in the cross-country validation exercise (Panel A of Table 1). Col. 4: countries included in training the machine-learning model. Country names are as reported in PISA codebooks or Facebook/GPS data and do not represent any political views of the authors.

Table A2: Validation of Cross-Country Analysis: Different Numbers of Principal Components (PCs)

	20 PCs	30 PCs	40 PCs	50 PCs
	(1)	(2)	(3)	(4)
A. Original country sample (GPS countries)				
Patience	1.638***	1.588***	1.593***	1.660***
	(0.136)	*(0.141)*	*(0.138)*	*(0.146)*
Risk-taking	-1.525***	-1.089***	-0.989***	-0.940***
	(0.443)	*(0.382)*	*(0.332)*	*(0.332)*
Control variables	Yes	Yes	Yes	Yes
Observations	1,954,840	1,954,840	1,954,840	1,954,840
Residence countries	48	48	48	48
R^2	0.208	0.199	0.198	0.197
B. Extended country sample (all Facebook countries)				
Patience	1.682***	1.583***	1.588***	1.689***
	(0.125)	*(0.128)*	*(0.129)*	*(0.135)*
Risk-taking	-1.531***	-1.422***	-1.224***	-1.089***
	(0.330)	*(0.331)*	*(0.281)*	*(0.279)*
Control variables	Yes	Yes	Yes	Yes
Observations	2,660,408	2,660,408	2,660,408	2,660,408
Residence countries	80	80	80	80
R^2	0.204	0.203	0.198	0.194

Notes: Dependent variable: PISA math test score in all PISA waves 2000-2018. Least squares regressions weighted by students' sampling probability. Control variables: student gender, age, and migration status; imputation dummies; and wave fixed effects. Robust standard errors adjusted for clustering at the country level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2000-2018; own elaboration of Facebook data.

Table A3: Countries in the Migrant Analysis

	GPS/Facebook country of origin			PISA destination country	
	Only GPS	Only Facebook	Both	GPS analysis	Facebook analysis
	(1)	(2)	(3)	(4)	(5)
Afghanistan			x		
Albania		x			
Algeria					
Argentina			x	x	x
Armenia		x			
Australia			x	x	x
Austria			x	x	x
Azerbaijan		x			
Bangladesh			x		
Belarus		x		x	x
Belgium		x		x	x
Bolivia			x		
Bosnia and Herzegovina			x	x	x
Brazil			x		
Brunei Darussalam				x	x
Bulgaria		x			
Cape Verde		x			
Canada			x	x	x
Chile			x		
China			x		
Colombia			x		
Costa Rica				x	x
Croatia			x	x	x
Czech Republic			x	x	x
Denmark		x		x	x
Dominican Republic		x		x	x
Egypt			x		
Estonia			x		
Ethiopia		x			
Fiji		x			
Finland			x	x	x
France			x		
Georgia			x		x
Germany			x	x	x
Greece			x		x
Haiti			x		
Hong Kong				x	x
Hungary			x		
Iceland		x			
India			x		
Indonesia			x	x	x
Iran	x				
Iraq			x		
Ireland		x		x	x
Israel				x	x
Italy			x		
Japan					
Jordan			x	x	x

(continued on next page)

Table A3 (continued)

	GPS/Facebook country of origin			PISA destination country	
	Only GPS	Only Facebook	Both	GPS analysis	Facebook analysis
	(1)	(2)	(3)	(4)	(5)
Kazakhstan			x		
Kuwait		x			
Kyrgyzstan					x
Latvia				x	x
Lebanon		x			
Libya		x			
Liechtenstein		x		x	x
Lithuania			x		
Luxembourg				x	x
Macao		x		x	x
Malaysia		x			
Mauritius				x	x
Mexico				x	x
Moldova			x	x	x
Montenegro		x		x	x
Morocco			x	x	x
Netherlands			x	x	x
New Zealand		x		x	x
Nicaragua			x		
Nigeria			x		
North Macedonia				x	x
Norway		x		x	x
Pakistan			x		
Palestine		x			
Panama		x		x	x
Paraguay		x			
Peru					
Philippines			x	x	x
Poland			x		
Portugal			x	x	x
Qatar		x		x	x
Romania			x		
Russia	x				
Samoa		x			
Saudi Arabia			x	x	x
Serbia			x		x
Singapore		x			
Slovakia		x		x	x
Slovenia		x		x	x
Somalia		x			
South Africa			x		
South Korea			x	x	x
Spain			x		
Suriname			x		
Sweden			x		
Switzerland			x	x	x
Tajikistan		x			
Thailand			x		

(continued on next page)

Table A3 (continued)

	GPS/Facebook country of origin			PISA destination country	
	Only GPS	Only Facebook	Both	GPS analysis	Facebook analysis
	(1)	(2)	(3)	(4)	(5)
Tonga		x			
Turkey			x	x	x
Ukraine			x	x	x
United Arab Emirates			x		
United Kingdom			x	x	x
United States			x		
Uruguay		x		x	x
Uzbekistan		x			
Venezuela			x		
Vietnam			x		
Yemen		x			
Zambia		x			
Total: 108 countries	2	37	56	46	50

Notes: Sample of countries that serve as countries of origin (col. 1-3) or destination countries (col. 4-5) in the migrant analysis (Panel B of Table 1). Country names are as reported in PISA codebooks or Facebook/GPS data and do not represent any political views of the authors.

Table A4: Validation of Migrant Analysis: Different Numbers of Principal Components (PCs)

	20 PCs	30 PCs	40 PCs	50 PCs
	(1)	(2)	(3)	(4)
A. Original sample (GPS countries of origin)				
Patience	0.783***	0.876***	0.885***	0.875***
	(0.193)	*(0.197)*	*(0.192)*	*(0.216)*
Risk-taking	-0.676**	0.008	0.087	0.156
	(0.306)	*(0.367)*	*(0.322)*	*(0.371)*
Control variables	Yes	Yes	Yes	Yes
Residence-country by wave fixed effects	Yes	Yes	Yes	Yes
Observations	78,403	78,403	78,403	78,403
Countries of origin	56	56	56	56
Residence countries	46	46	46	46
R^2	0.271	0.271	0.272	0.270
B. Extended sample (all Facebook countries of origin)				
Patience	0.838***	1.027***	1.033***	0.995***
	(0.211)	*(0.198)*	*(0.191)*	*(0.211)*
Risk-taking	-1.155***	-0.067	0.064	0.154
	(0.422)	*(0.357)*	*(0.297)*	*(0.341)*
Control variables	Yes	Yes	Yes	Yes
Residence-country by wave fixed effects	Yes	Yes	Yes	Yes
Observations	90,983	90,983	90,983	90,983
Countries of origin	93	93	93	93
Residence countries	50	50	50	50
R^2	0.295	0.294	0.294	0.291

Notes: Dependent variable: PISA math test score, waves 2003-2018. Least squares regressions, including 180 fixed effects for each residence-country by wave cell. Sample: students with both parents not born in the country where the student attends school. Control variables: student gender, age, dummy for OECD country of origin, imputation dummies. Robust standard errors adjusted for clustering at the country-of-origin level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA international student achievement test, 2003-2018; own elaboration of Facebook data.

Table A5: Patience and Student Achievement at Different Grade Levels

	Grade 2	Grade 4/5	Grade 8	Grade 10
	(1)	(2)	(3)	(4)
A. Italy (individual level)				
Patience	0.308	0.581*	1.614***	1.885***
	(0.197)	*(0.292)*	*(0.173)*	*(0.209)*
Control variables	Yes	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes	Yes
Observations	48,812	50,608	59,034	77,207
Regions	20	20	20	20
R^2	0.028	0.032	0.095	0.155
B. Italy (regional level)				
Patience	0.204	0.416*	1.370***	1.590***
	(0.209)	*(0.244)*	*(0.169)*	*(0.226)*
Wave fixed effects	Yes	Yes	Yes	Yes
Observations	40	40	40	40
Regions	20	20	20	20
R^2	0.045	0.097	0.771	0.743
C. United States (state level)				
Patience	–	0.156**	0.293***	–
		(0.064)	*(0.089)*	
Wave fixed effects		Yes	Yes	
Observations		153	153	
Regions		51	51	
R^2		0.158	0.360	

Notes: Dependent variable: Panels A and B: INVALSI math test score in waves 2018 and 2019; Panel C: NAEP math test score in all NAEP waves 2015-2019. Grade level indicated in column headers (col. 2 refers to fifth grade in Italy and fourth grade in the United States). Least squares regressions with wave fixed effects. Unit of observation: Panel A: student; Panel B: region-wave combination; Panel C: state-wave combination. Patience measured at the regional/state level throughout. Patience measure computed with 4 principal components (PCs). Regressions control for risk-taking computed with 4 PCs. Additional control variables (Panel A): student gender, age, and migration status; imputation dummies. The Italian region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional (state) level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI mathematics achievement test, 2017-2019; NAEP mathematics achievement test, 2015-2019; own elaboration of Facebook data.

Table A6: Patience and Reading Achievement: Analysis of Italian Regions

	4 PCs (1)	7 PCs (2)	10 PCs (3)
A. Individual level			
Patience	1.330***	1.044***	1.118***
	(0.184)	*(0.119)*	*(0.122)*
Control variables	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes
Observations	59,441	59,441	59,441
Regions	20	20	20
R^2	0.108	0.110	0.111
B. Regional level			
Patience	1.039***	0.799***	0.856***
	(0.145)	*(0.080)*	*(0.083)*
Wave fixed effects	Yes	Yes	Yes
Observations	40	40	40
Regions	20	20	20
R^2	0.677	0.740	0.753

Notes: Dependent variable: INVALSI eighth-grade reading test score in waves 2018 and 2019. Least squares regressions with wave fixed effects. Unit of observation: Panel A: student; Panel B: region-wave combination. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. Additional control variables (Panel A): student gender, age, and migration status; imputation dummies. The region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI reading achievement test, 2017-2019; own elaboration of Facebook data.

Table A7: Patience and Math Achievement: Analysis of Italian Regions by Wave and Gender

	2018	2019	Males	Females
	(1)	(2)	(3)	(4)
A. Individual level				
Patience (4 PCs)	1.711***	1.516***	1.689***	1.540***
	(0.164)	*(0.201)*	*(0.191)*	*(0.172)*
Control variables	Yes	Yes	Yes	Yes
Wave fixed effects	No	No	Yes	Yes
Observations	29,359	29,675	30,325	28,504
Regions	20	20	20	20
R^2	0.099	0.092	0.101	0.087
B. Regional level				
Patience (4 PCs)	1.474***	1.266***	1.434***	1.321***
	(0.237)	*(0.252)*	*(0.181)*	*(0.167)*
Wave fixed effects	No	No	Yes	Yes
Observations	20	20	40	40
Regions	20	20	20	20
R^2	0.804	0.739	0.753	0.774

Notes: Dependent variable: INVALSI eighth-grade math test score in waves 2018 and 2019. Least squares regressions with wave fixed effects. Unit of observation: Panel A: student; Panel B: region-wave combination. Patience measure computed with 4 principal components (PCs). Regressions control for risk-taking computed with 4 PCs. Additional control variables (Panel A): student gender, age, and migration status; imputation dummies. The region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI reading achievement test, 2017-2019; own elaboration of Facebook data.

Table A8: Patience and Math Achievement: Post COVID-19 Waves

		4 PCs	7 PCs	10 PCs
		(1)	(2)	(3)
A. Italy (regional level)				
2021	Patience	1.332***	1.098***	1.173***
		(0.259)	*(0.129)*	*(0.134)*
	Observations (regions)	20	20	20
	R^2	0.713	0.812	0.821
2022	Patience	1.409***	1.077***	1.151***
		(0.270)	*(0.136)*	*(0.141)*
	Observations (regions)	20	20	20
	R^2	0.689	0.793	0.804
2023	Patience	1.220***	0.981***	1.050***
		(0.260)	*(0.134)*	*(0.139)*
	Observations (regions)	20	20	20
	R^2	0.658	0.763	0.774
Pooled (2021-2023)	Patience	1.320***	1.052***	1.125***
		(0.147)	*(0.074)*	*(0.077)*
	Wave fixed effects	Yes	Yes	Yes
	Observations (regions)	60	60	60
	R^2	0.686	0.788	0.799
B. United States (state level)				
2022	Patience	0.246***	0.153*	0.208*
		(0.083)	*(0.087)*	*(0.120)*
	Observations (states)	51	51	51
	R^2	0.295	0.289	0.282

Notes: Dependent variable: Panel A: INVALSI eighth-grade math test score; Panel B: NAEP eighth-grade math test score. Least squares regressions. Unit of observation: Panel A: region-wave combination; Panel B: state-wave combination. Patience measured at the regional/state level throughout. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. The Italian region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI mathematics achievement test, 2021-2023; NAEP mathematics achievement test, 2022; own elaboration of Facebook data.

Table A9: Patience and Math Achievement: Analysis of Italian Regions by Migrant Status

	4 PCs	7 PCs	10 PCs
	(1)	(2)	(3)
A. Native students			
Patience	1.677***	1.472***	1.571***
	(0.172)	*(0.112)*	*(0.113)*
Control variables	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes
Observations	51,691	51,691	51,691
Regions	20	20	20
R^2	0.086	0.091	0.092
B. Second-generation migrant students			
Patience	0.967***	0.715***	0.787***
	(0.261)	*(0.225)*	*(0.232)*
Wave fixed effects	Yes	Yes	Yes
Observations	3,572	3,572	3,572
Regions	20	20	20
R^2	0.034	0.035	0.036
C. First-generation migrant students			
Patience	0.757***	0.919***	0.986***
	(0.178)	*(0.084)*	*(0.092)*
Wave fixed effects	Yes	Yes	Yes
Observations	1,719	1,719	1,719
Regions	20	20	20
R^2	0.093	0.097	0.097

Notes: Dependent variable: INVALSI eighth-grade math test score in waves 2018 and 2019. Least squares regressions with wave fixed effects. Unit of observation: student. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. Additional control variables: student gender and age; imputation dummies. The region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI mathematics achievement test, 2017-2019; own elaboration of Facebook data.

Table A10: Patience and Math Achievement: Conditioning on Trust

	4 PCs (1)	7 PCs (2)	10 PCs (3)
A. Italy (individual level)			
Patience	1.614***	1.369***	1.462***
	(0.173)	*(0.098)*	*(0.102)*
Control variables	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes
Observations	59,034	59,034	59,034
Regions	20	20	20
R^2	0.095	0.103	0.103
B. Italy (regional level)			
Patience	1.237***	1.171***	1.251***
	(0.169)	*(0.075)*	*(0.093)*
Wave fixed effects	Yes	Yes	Yes
Observations	40	40	40
Regions	20	20	20
R^2	0.771	0.887	0.891
C. United States (state level)			
Patience	0.312	0.167*	0.274**
	(0.966)	*(0.092)*	*(0.133)*
Wave fixed effects	Yes	Yes	Yes
Observations	153	153	153
Regions	51	51	51
R^2	0.360	0.362	0.372

Notes: Dependent variable: Panels A and B: INVALSI eighth-grade math test score in waves 2018 and 2019; Panel C: NAEP eighth-grade math test score in all NAEP waves 2015-2019. Least squares regressions with wave fixed effects. Unit of observation: Panel A: student; Panel B: region-wave combination; Panel C: state-wave combination. Patience measured at the regional/state level throughout. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking and trust computed with the equivalent number of PCs. Additional control variables (Panel A): student gender, age, and migration status; imputation dummies. The Italian region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional (state) level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI mathematics achievement test, 2017-2019; NAEP mathematics achievement test, 2015-2019; own elaboration of Facebook data.

Table A11: Patience and Math Achievement: Excluding Education-Related Interests from the Construction

	4 PCs	7 PCs	10 PCs
	(1)	(2)	(3)
A. Italy (individual level)			
Patience	1.596***	1.344***	1.438***
	(0.168)	*(0.107)*	*(0.108)*
Control variables	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes
Observations	59,034	59,034	59,034
Regions	20	20	20
R^2	0.095	0.099	0.100
B. Italy (regional level)			
Patience	1.355***	1.153***	1.234***
	(0.168)	*(0.086)*	*(0.090)*
Wave fixed effects	Yes	Yes	Yes
Observations	40	40	40
Regions	20	20	20
R^2	0.771	0.832	0.839
C. United States (state level)			
Patience	0.287***	0.156	0.280**
	(0.089)	*(0.098)*	*(0.128)*
Wave fixed effects	Yes	Yes	Yes
Observations	153	153	153
Regions	51	51	51
R^2	0.365	0.353	0.373

Notes: Dependent variable: Panels A and B: INVALSI eighth-grade math test score in waves 2018 and 2019; Panel C: NAEP eighth-grade math test score in all NAEP waves 2015-2019. Least squares regressions with wave fixed effects. Unit of observation: Panel A: student; Panel B: region-wave combination; Panel C: state-wave combination. Patience measured at the regional/state level throughout. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. Additional control variables (Panel A): student gender, age, and migration status; imputation dummies. The Italian region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional (state) level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI mathematics achievement test, 2017-2019; NAEP mathematics achievement test, 2015-2019; own elaboration of Facebook data.

Table A12: Patience and Math Achievement: Analysis of Italian Regions Excluding Trentino-Alto-Adige

	4 PCs	7 PCs	10 PCs
	(1)	(2)	(3)
A. Individual level			
Patience	1.716***	1.413***	1.520***
	(0.158)	*(0.122)*	*(0.124)*
Control variables	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes
Observations	55,437	55,437	55,437
Regions	19	19	19
R^2	0.095	0.098	0.098
B. Regional level			
Patience	1.462***	1.221***	1.314***
	(0.171)	*(0.094)*	*(0.097)*
Wave fixed effects	Yes	Yes	Yes
Observations	38	38	38
Regions	19	19	19
R^2	0.784	0.836	0.846

Notes: Dependent variable: INVALSI eighth-grade math test score in waves 2018 and 2019. Least squares regressions with wave fixed effects. Unit of observation: Panel A: student; Panel B: region-wave combination. Students in the autonomous municipalities of Trento and Bolzano are dropped from the estimation sample. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. Additional control variables (Panel A): student gender, age, and migration status; imputation dummies. Robust standard errors adjusted for clustering at the regional level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI mathematics achievement test, 2017-2019; own elaboration of Facebook data.

Table A13: Analysis of Unobservable Selection and Coefficient Stability following Oster (2019): Analysis of Italian Regions

	4 PCs		7 PCs		10 PCs	
	Restricted	Extended	Restricted	Extended	Restricted	Extended
	(1)	(2)	(3)	(4)	(5)	(6)
Patience	1.375***	1.614***	1.202***	1.398***	1.285***	1.495***
	(0.173)	*(0.174)*	*(0.102)*	*(0.107)*	*(0.104)*	*(0.108)*
Control variables	No	Yes	No	Yes	No	Yes
Wave fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	59,034	59,034	59,034	59,034	59,034	59,034
Regions	20	20	20	20	20	20
R^2	0.049	0.095	0.053	0.100	0.054	0.101
Oster (2019) diagnostics						
Bound β^* for $\delta = 1$	1.761		1.526		1.633	
δ to match $\beta = 0$	-4.838		-2.895		-2.856	

Notes: Dependent variable: INVALSI eighth-grade math test score in waves 2018 and 2019. Least squares regressions with wave fixed effects. Unit of observation: student. Patience measure computed with number of principal components (PCs) indicated in column header. Regressions control for risk-taking computed with the equivalent number of PCs. Odd columns: restricted model with wave fixed effects. Even columns: baseline models with wave fixed effects, student gender, age, and migration status; imputation dummies. Oster statistics computed using $R_{max} = 1.3\tilde{R}$, where $\tilde{R}$ denotes the R^2 reported in even columns. The region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: INVALSI mathematics achievement test, 2017-2019; own elaboration of Facebook data.

Table A14: Patience and Math Achievement: Analysis of Italian Regions using PISA Data

	4 PCs	7 PCs	10 PCs
	(1)	(2)	(3)
Patience	1.484***	1.473***	1.570***
	(0.263)	*(0.132)*	*(0.138)*
Control variables	Yes	Yes	Yes
Observations	31,073	31,073	31,073
States	20	20	20
R^2	0.106	0.113	0.113

Notes: Dependent variable: PISA 2012 math test score. Least squares regressions. Unit of observation: student. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. Additional control variables: student gender, age, and migration status; imputation dummies. The region of Trentino-Alto-Adige is represented by the two municipalities of Bolzano and Trento. Robust standard errors adjusted for clustering at the regional level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: PISA student achievement test, 2012; own elaboration of Facebook data.

Table A15: Patience and Reading Achievement: Analysis of U.S. States

	4 PCs (1)	7 PCs (2)	10 PCs (3)
Patience	0.228***	0.141*	0.228**
	(0.074)	*(0.077)*	*(0.103)*
Wave fixed effects	Yes	Yes	Yes
Observations	153	153	153
States	51	51	51
R^2	0.385	0.375	0.396

Notes: Dependent variable: NAEP eighth-grade reading test score in all NAEP waves 2015-2019. Least squares regressions with wave fixed effects. Unit of observation: state-wave combination. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. Robust standard errors adjusted for clustering at the state level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: NAEP mathematics achievement test, 2015-2019; own elaboration of Facebook data.

Table A16: Patience and Math Achievement: Analysis of U.S. States by Wave

	4 PCs	7 PCs	10 PCs
	(1)	(2)	(3)
A. 2015			
Patience	0.335***	0.194**	0.349***
	(0.081)	*(0.082)*	*(0.119)*
States	51	51	51
R^2	0.426	0.410	0.433
B. 2017			
Patience	0.309***	0.179**	0.293**
	(0.084)	*(0.085)*	*(0.125)*
States	51	51	51
R^2	0.373	0.361	0.374
C. 2019			
Patience	0.235***	0.142*	0.231**
	(0.077)	*(0.078)*	*(0.114)*
States	51	51	51
R^2	0.277	0.267	0.280

Notes: Dependent variable: NAEP eighth-grade math test score in all NAEP waves 2015-2019. Least squares regressions with wave fixed effects. Unit of observation: state-wave combination. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. Robust standard errors adjusted for clustering at the state level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: NAEP mathematics achievement test, 2015-2019; own elaboration of Facebook data.

Table A17: Patience and Math Achievement: Analysis of U.S. States by Gender

	4 PCs (1)	7 PCs (2)	10 PCs (3)
A. Males			
Patience	0.322***	0.193*	0.310**
	(0.101)	*(0.108)*	*(0.147)*
Wave fixed effects	Yes	Yes	Yes
Observations	153	153	153
States	51	51	51
R^2	0.388	0.377	0.388
B. Females			
Patience	0.263***	0.147*	0.263**
	(0.079)	*(0.086)*	*(0.118)*
Wave fixed effects	Yes	Yes	Yes
Observations	153	153	153
States	51	51	51
R^2	0.319	0.305	0.324

Notes: Dependent variable: NAEP eighth-grade math test score in all NAEP waves 2015-2019. Least squares regressions with wave fixed effects. Unit of observation: state-wave combination. Col. 1-3 use the patience measure computed with 4, 7, and 10 principal components (PCs), respectively. Regressions control for risk-taking computed with the equivalent number of PCs. Robust standard errors adjusted for clustering at the state level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent. Data sources: NAEP mathematics achievement test, 2015-2019; own elaboration of Facebook data.