



Schooling Ain't Learning in Europe: A Club Convergence Perspective

Linda Glawe¹ · Carlos Mendez²

Accepted: 21 May 2022 / Published online: 21 June 2022
© The Author(s) 2022

Abstract

Convergence in human capital is a key precondition to achieve income convergence in the European Union; however, so far research on this topic has nearly been absent. Our paper contributes to the literature by investigating the human capital convergence dynamics within the EU over the period 1990–2016 using a nonlinear dynamic factor model. While we find evidence of absolute convergence with respect to the average years of schooling, we identify four convergence clubs with respect to learning outcomes, and the divergence across those four clubs is increasing over time. A subsequent analysis of the determinants of the learning clubs reveals that institutional and learning spillovers are decisive for whether an EU Member State is on a high or low human capital trajectory.

Keywords Human capital · Learning · Club convergence · European Union · Institutional quality · Spillover effects

JEL Classification R10 · R11 · R58

Introduction

Real convergence across the European Union (EU) Member States is a key goal of the European integration process, as stated for example in the EC Treaty of 1957 and the Maastricht Treaty of 1992. However, recent research indicates that especially after the Global Financial Crisis, there appear to be increasing income divergence tendencies within the EU (see Glawe and Wagner 2021a,

✉ Linda Glawe
linda.glawe@fernuni-hagen.de

Carlos Mendez
carlos@gsid.nagoya-u.ac.jp

¹ Faculty of Economics, University of Hagen, Universitätsstr. 11, 58084 Hagen, Germany

² Graduate School of International Development, Nagoya University, Furocho, Chikusa Ward, Nagoya, Aichi 464-8601, Japan



for an overview). In addition, various studies report that the EU Member States converge to different income clubs; these clusters usually follow a specific geographical pattern with many south-eastern countries being caught in a bad income equilibrium (see, e.g., Von Lyncker and Thoennesen 2017; Bartkowska and Riedl 2012). It is of particular importance to identify the driving factors of this development since persistent income disparities could endanger the stability and smooth functioning of the EU. In particular, it appears logical to study the dynamics of the underlying growth determinants to gain more insights into the income convergence process itself; however, so far, little research has been conducted in this respect. Despite an uprising literature branch that focuses on institutional convergence within the EU (see Glawe and Wagner 2021b), research on the convergence of other growth determinants is nearly absent (not only for the EU, but also with respect to other countries and regions).

Our paper aims at contributing to close this research gap by focusing on human capital. There is a significant body of literature that identifies human capital as one of the most important factors driving economic growth (see, e.g., Lucas 1988; Mankiw et al. 1992; De la Fuente and Doménech 2006; Hanushek and Woessmann 2008) and real convergence, also within the EU (see, e.g., Von Lyncker and Thoennesen 2017; Glawe and Wagner 2021b, who identify human capital as a key driver of income and, respectively, institutional convergence clubs within the EU). Moreover, a sound human capital endowment is critical for switching from an imitation-based to an innovation-driven growth strategy which is in turn decisive for sustaining growth at higher income levels (cf. the literature on innovation-based endogenous growth models, e.g., Aghion and Howitt 1992). This is of particular relevance for the EU since the region as a whole is already comparatively well developed, but region-specific income disparities are persistent and even widening, particularly between the core EU and the periphery. For some upper-middle-income EU countries, upgrading human capital might be also a necessary precondition to overcome an (upper-) middle-income trap (MIT), since education is usually identified as an important MIT triggering factor. Overall, studying the convergence dynamics of human capital within the EU seems to be an important task.

In our paper, we analyze whether there is human capital convergence across EU Member States or whether there exist multiple human capital clubs within the EU over the period 1990 to 2016 by using the log- t test proposed by Phillips and Sul (2007a, 2009). Importantly, we employ not only quantitative human capital data (namely, the average years of schooling) but also focus on the quality of human capital by exploiting a new database on harmonized measures of learning. Thus, we take into account that recent research emphasizes that schooling (“quantity”) does not necessarily imply learning (“quality”) (see, e.g., Filmer et al. 2020; Glawe and Wagner 2022). Our results indicate that while there is only one club in schooling, there exist multiple learning clubs. Interestingly, neighboring EU countries tend to cluster together, indicating that physical location and geographical spillovers are important to the formation of learning clusters within Europe. A subsequent ordered probit analysis of the determinants of the learning clubs reveals that there exist indeed strong spatial spillover effects. Our results suggest that the spatial lag of the initial level of learning is decisive for learning club membership, trumping institutional



quality (which still has a small positive effect once we control for learning) and culture measured by trust.

The remainder of the paper is structured as follows: The subsequent “[Related Literature](#)” section briefly presents the related literature. “[Data and Methods](#)” section then introduces our data and estimation strategy. Our regression results are discussed in “[Results](#)” section. In particular, “[Convergence Test and Club Identification](#)” sub-section presents the results of the convergence test, and “[Determinants of Club Membership](#)” sub-section describes our findings with respect to the analysis of the club determinants. “[Concluding Remarks and Policy Implications](#)” section concludes and draws some policy implications.

Related Literature

There are several literature strands related to our paper, most importantly the literature on the impact of human capital on growth, the literature on the nature of human capital development, and the literature on the concept of club convergence. Each of them will be briefly discussed in the following.

Impact of Human Capital on Growth

It is by now a well-established fact that human capital is decisive for explaining differences in economic performance across countries and regions. At the macro-level, the economic growth literature identifies three main channels through which education can affect growth, namely via (i) increases in labor productivity (cf. Mankiw et al. 1992), (ii) a rise in the innovative capacity (cf. Nelson and Phelps 1966; Benhabib and Spiegel 2005), as well as (iii) a facilitated diffusion of knowledge (cf. e.g., Lucas 1988; Romer 1990; see also Hanushek and Woessmann 2008). There are also numerous studies that analyze empirically the association between education and economic growth. While the initial studies focused primarily on measures of *school quantity* such as enrollment ratios or the percentage of the population with completed secondary education (see Topel 1999; Temple 2002, for overviews), more recently the focus has shifted to the *quality of education* (usually measured by cognitive test scores, see, e.g., Shultz and Hanushek 2012).

Club Convergence

Within the economic growth literature, it can be distinguished between different convergence concepts: σ -convergence can be understood as a decline in the dispersion of per capita income over time, whereas β -convergence is defined as the negative correlation between the growth rate of per capita income and the initial income level, implying that poor countries tend to grow faster than rich ones (Barro et al. 1991; Barro and Sala-i-Martin 1992). It can be further differentiated between absolute (unconditional) and conditional β -convergence (Sala-i-Martin 1996; Islam 2003). Absolute β -convergence implies that economies will converge toward the



same steady-state over time given the variability of the initial conditions. The main idea behind conditional β -convergence is that countries will approach their own unique steady state which may depend on country-specific characteristics such as the savings rate, factor endowment, or institutions. The idea of club convergence is in turn based on models that allow for multiple equilibria (cf. Baumol 1986; Durlauf and Johnson 1995; Galor 1996). In particular, a group of countries may converge toward a particular equilibrium if they have the same initial location or attribute that correspond to that equilibrium (cf. Islam 2003).

There is already a considerable body of literature that analyzes the convergence behavior of per capita income within the EU. An extensive overview of this literature can be found in Glawe and Wagner (2021a). Especially since 2010, there is an increasing number of studies that apply the log- t test of Phillips and Sul (2007a, 2009) to investigate the formation of income convergence clubs across E(M)U countries (see, e.g., Fritsche and Kuzin 2011; Monfort et al. 2013; Glawe and Wagner 2021b) and regions (see, e.g., Bartkowska and Riedl 2012; Von Lyncker and Thoennesen 2017). Most of these studies find that there exist multiple income clubs within the E(M)U. In addition, there is evidence that the convergence clubs follow a certain geographic pattern; in particular, there appears to be a kind of northwest-southeast divide; various studies also detect a core-periphery divide.

The Nature of Human Capital Development

If human capital is an important driver of economic growth, human capital convergence should be necessary precondition for convergence in per capita incomes. However, there are rather few studies that analyze empirically the convergence behavior of human capital across countries. These studies focus mostly on the traditional convergence tests such as σ - and β -convergence and neglect the possibility of club convergence. Moreover, they usually only analyze measures of human capital quantity. Examples include the studies of O'Neill (1995), Cohen (1996), and Sab and Smith (2001) which all report human capital convergence. So far, there is no research that analyzes the dynamics of human capital development within the European Union (or the Euro Area).

Data and Methods

Schooling and Learning Data

In the economic growth literature, mean years of total schooling across all education levels is one of the most commonly used indicators of education quantity across countries (Barro and Lee 2013). Consistent with this literature, our indicator of schooling (or education quantity) is the mean years of schooling. These data are originally compiled and harmonized by Barro and Lee (2013) in a five-year interval basis. Roser and Ortiz-Ospina (2016) build upon these data by adding information from the United Nation's human development report and then interpolate the



missing years. Table 1 succinctly summarizes this dataset in the context of our sample and time horizon. The overall message of these descriptive statistics is fairly optimistic. Over the 1990–2016 period, the mean years of schooling increased from 8.66 to 12.07 years and the cross-country disparities decreased from 1.28 to 1.11. The log values of these numbers are reassuring in the sense that outliers do not appear to be the source of these changes.

Pritchett (2013) strongly argues that years of schooling do not represent actual learning. In other words, the quantity of education (measured in years of schooling) is often a poor indicator of the quality of education (measured in terms of test scores). Harmonized dataset about learning scores, however, are usually unavailable for many countries and years. The recent study of Lim et al. (2018) provides a new dataset that includes harmonized measures of learning for 195 countries over the 1990–2016 period. The estimates of learning in this dataset are based on student testing data from major international and national assessments. Specifically, learning scores are based on 1894 tests, covering mathematics, science, and reading across 132 countries and 163 subnational locations.

In the Lim et al. (2018) dataset, the learning scores are available for school-aged children between 5 to 19 years. They are constructed from multiple international assessment programs such as the Program for International Student Assessment, Progress in International Reading Literacy Study (PIRLS), and the International Association for the Evaluation of Educational Achievement.¹ To estimate test scores for all countries, years, and ages, a spatiotemporal Gaussian regression model was used. With math, science, and reading tests given equal weight, a combined learning measure ranging from 0 to 1000 has been estimated.

In this article, we exploit this new dataset of Lim et al. (2018) to evaluate the evolution of learning scores for the age group between 15 to 19 years. The lower panel of Table 1 provides a first overview of this dataset in the context of our sample and time horizon. The overall message of the learning data is less optimistic than that of schooling. Although the mean score of learning has improved, learning disparities have increased over the 1990–2016 period.

Figure 1 illustrates the sharp contrast between schooling and learning. In panel (a), the evolution of the standard deviation of years of schooling indicates that there is a systematic decrease in cross-country disparities. In contrast, panel (b) shows that the cross-country disparities in learning have systematically increased.

¹ In addition to the international programs, regional assessment programs were also considered, including the Southern and Eastern Africa Consortium for Monitoring Educational Quality, the Latin American Laboratory for Assessment of the Quality of Education, and the Programme d'Analyse des Systèmes Educatifs de la Confem.



Convergence Test and Club Identification

The nonlinear dynamic factor model proposed by Phillips and Sul (2007a, 2009)² decomposes a panel-data variable, y_{it} , as follows:

$$y_{it} = \delta_{it}\mu_t \tag{1}$$

where δ_{it} is an idiosyncratic component and μ_t is a common component. Equation 1 represents a simple dynamic factor model where δ_{it} is a factor-loading coefficient that represents the individual distance between a common long-run trend, δ_{it} , and the observed variable, y_{it} . More intuitively, δ_{it} describes the distance (transition path) of an individual economy toward its long-run equilibrium, μ_t .

To test the existence of a common long-run equilibrium and the convergence hypothesis, Phillips and Sul (2007a) define a relative transition parameter, h_{it} , as

$$h_{it} = \frac{y_{it}}{\frac{1}{N} \sum_{i=1}^N y_{it}} = \frac{\delta_{it}}{\frac{1}{N} \sum_{i=1}^N \delta_{it}}. \tag{2}$$

By dividing the observed variable, y_{it} , by the panel average, the relative transition parameter helps remove the common component, μ_t , from Eq. 1. Next, as $t \rightarrow \infty$, convergence is defined as

$$H_t = \frac{1}{N} \sum_{i=1}^N (h_{it} - 1)^2. \tag{3}$$

In other words, when the relative transition parameter converges to unity, $h_{it} \rightarrow 1$, the cross-sectional variance converges to zero, $H_t \rightarrow 0$. To test this hypothesis, Phillips and Sul (2007a) propose the following log- t regression model

$$\log\left(\frac{H_1}{H_t}\right) - 2\log\{\log(t)\} = a + b\log(t) + \epsilon_t \tag{4}$$

for $t = [rT], [rT] + 1, \dots, T$ with $r > 0$,

where $[rT]$ is the initial observation in the regression, which implies that the first fraction of the data, r , is discarded. Based on Monte Carlo experiments, Phillips and Sul (2007a) suggest to set $r = 0.3$ when the sample is small (that is, $T \leq 50$). An inferential procedure based on a one-sided t -test with HAC standard errors is suggested for Eq. 4. In this setting, the null hypothesis of convergence is rejected when $t_b < -1.65$.

An appealing feature of the model of Eq. 4 is that the magnitude and sign of coefficient b indicates different convergence patterns. When $b < 0$, the model suggests divergence. In contrast, when $0 \leq b < 2$, the model suggests convergence in growth rates (also known as relative convergence). When $b \geq 2$, the model suggests

² A growth-theoretical foundation of Phillips and Sul's (2007a) log- t test is provided by Phillips and Sul (2007b). In particular, they show that augmented Solow growth regressions allowing for heterogenous technology can explain multiple equilibria.



Table 1 Descriptive statistics for schooling and learning

	Mean	SD	Min	Max
Schooling in 1990	8.66	1.28	5.86	10.99
Schooling in 2016	12.07	1.11	9.16	14.6
Log Schooling in 1990	2.15	0.15	1.77	2.4
Log Schooling in 2016	2.49	0.1	2.21	2.68
Learning in 1990	513.28	15.8	476.94	543.3
Learning in 2016	519.96	17.95	473.78	553.82
Log Learning in 1990	6.24	0.03	6.17	6.3
Log Learning in 2016	6.25	0.04	6.16	6.32

Schooling is measured as mean years of total schooling across all education levels. Learning is measured using test scores for the age group between 15 to 19 years. For both indicators, the sample includes 28 European countries

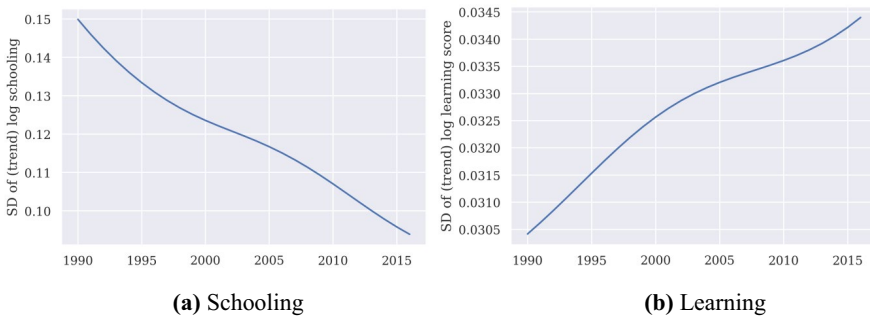


Fig. 1 Evolution of disparities in schooling and learning, 1990–2016. Notes: Schooling is measured as mean years of total schooling across all education levels. Learning is measured using test scores for the age group between 15 to 19 years. For both indicators, the sample includes 28 European countries. Note that in Fig. 1 we use the *trend* of (the logarithm of) schooling and learning, respectively. Thus, the standard deviations are not identical to those reported in Table 1

convergence in levels (also known as absolute convergence). Finally, a measure of the speed of convergence can be calculated as $b/2$.

When the null hypothesis of overall convergence is rejected, there may be possible convergence clubs in the panel. Phillips and Sul (2007a) propose the following clustering algorithm in order to identify local convergence clubs. The five steps of the clustering algorithm are briefly summarized as follows.

1. **Cross-sectional ordering:** The individuals in the panel are ordered decreasingly according to their last observation.
2. **Core group formation:** The second step is to select the first k highest units in the panel form the subgroup G_k for some $2 \leq k < N$ and apply the log- t regression to obtain the convergence test statistic $t_k = t(G_k)$ for this subgroup. Then, we choose the core group size k^* by maximizing t_k over k according to the criterion



- $k^* = \operatorname{argmax}_k \{t_k\}$ subject to $\min \{t_k\} > -1.65$. t_k indicates the one-side t -statistic that is needed to evaluate the statistical significance of the convergence test. If $t_k > -1.65$ is not valid for $k = 2$, then the highest unit is dropped from the core group and the algorithm can be repeated again for the rest of the sample.
3. **Sieving individuals for club membership:** In this step, one of the remaining individuals is added to the core group G_k and then we execute log- t test again. A new group is formed when the t -statistic is greater than zero.
 4. **Recursion and stopping:** In this step, we form the new group consisting of all individuals that could not be selected in Step 3, and apply log- t test for this subgroup. If $t_k > -1.65$, it indicates that two convergence subgroups exist. Otherwise, if the null hypothesis of convergence is rejected, Steps 1 to 3 are repeated. If no core group is found, then the remaining provinces are labeled as divergent provinces and the algorithm stops.
 5. **Club merging:** After following steps of clustering algorithm, if we identify more than one convergence club, we test whether these clubs might merge into larger clubs. We run the log- t test for all pairs of initial clubs. The merging procedure is iterative. That is, we run the log- t test for the initial Clubs 1 and 2, and if they fulfill the convergence test jointly, they should be merged into a new Club 1. We repeat this merging procedure for the other two clubs until the convergence test is rejected.

Determinants of Club Membership

Institutional Determinants

The literature on the deep determinants of economic development, coined among others by Acemoglu et al. (2001, 2014) and Rodrik et al. (2004), argues that (formal) institutions are decisive for the economic development process of a country since they shape the incentives of key economic actors in a society. Among others, institutions influence investment decisions with respect to human capital (see Acemoglu et al. 2005; Dias and Tebaldi 2012). Thus, it appears important to consider the effects of institutional quality on the formation of human capital clusters. In order to control for differences in institutions across countries, we include the first principal component of the six Worldwide Governance Indicators (WGIs), namely government effectiveness, regulatory quality, rule of law, control of corruption, voice and accountability, as well as political stability and violence avoidance.³ Descriptive statistics are provided in Table 2.⁴ The mean of our institutional variable is about zero

³ Please note that the set of institutions the matter for the accumulation of human capital (and more broadly economic development) is complex and cannot be fully captured by an index. Keeping this limitation in mind, we still think that our choice to use the WGIs is reasonable. The WGIs are considered as one the most reliable and accurate measures of national institutional quality and frequently used in empirical studies investigating the impact of institutions economic development (cf. Dollar and Kraay 2003).

⁴ Please note that our sample only includes 26 countries (and not 28) since two countries are diverging.



with a relatively high standard deviation of 2.30. The Netherlands is the country with the highest level of institutional quality (2.88), and Croatia brings up the rear (with a score of -5.38).

Culture/Trust Factors

Also, cultural aspects might play a role for the accumulation of human capital, as pointed out by a considerable body of literature (see, e.g., Knack and Keefer 1995; La Porta et al. 1997; Dearmon and Grier 2011). We focus here particularly on the effect of culture on beliefs about trust (Guiso et al. 2006) which in turn influence, among others, education and the formation of human capital. For instance, higher levels of trust lower transaction costs associated with employing educated labor and increase the efficiency of information sharing, thus raising the efficiency of human capital (cf. e.g., Dearmon and Grier 2009; Bjørnskov 2009; Coleman 1988). In controlling for cultural differences, particularly differences in trust, we include the variable “trust in others” published by Eurostat census data. This variable ranges from 0 (“You do not trust any other person.”) to 10 (“Most people can be trusted.”). The average trust level of EU countries is about 6, ranging from 4.20 for Bulgaria to 8.30 for Denmark (cf. Table 2).

Geographical Spillovers

Geographical proximity facilitates the diffusion and adoption of technological innovations (Ertur and Koch 2007; Basile et al. 2012; Caragliu and Nijkamp 2016). There is a growing literature that documents the effects of geographical spillovers on economic and social development.⁵ In particular, Quah (1996) points out the importance of informational and geographical externalities for explaining the formation of convergence clubs. Information flows diffuse more easily across regions that belong to the same state or share a border, hence spatial dependence and spillovers can occur at both subnational and national levels.

The Moran’s *I* statistic is used in geographical sciences to quantify the overall degree of spatial dependence in a regional system. Figure 2 shows that spatial dependence in learning outcomes has been systematically increasing within the EU. Specifically, a positive value in Moran’s *I* statistic indicates that neighboring countries tend to cluster together, thus geographical location and spillovers are likely to play an important role in the convergence process among European countries. Descriptive statistics of the spatially lagged variables used in our ordered probit regressions can be found in Table 2.

In order to capture the role of spatial dependence and geographical spillovers in the convergence process, we first identify the four nearest neighbors of each country; then, we compute a weighted average of their performance in learning outcomes,

⁵ See Karlsson and Gråsjö (2019) for a recent survey on knowledge externalities and regional economic development.



Table 2 Descriptive statistics, ordered probit model

	Mean	SD	Min	Max
<i>Dependent variable</i>				
Learning club	2.11	1.07	0	4
<i>Initial level of the variable</i>				
Learning	6.24	0.03	6.17	6.30
Institutions	0.00	2.30	-5.38	2.88
Trust	5.93	0.85	4.20	8.30
<i>Spatial lag of the initial level of the variable</i>				
Learning	-0.11	1.73	-3.69	2.49
Institutions	5.93	0.85	4.20	8.30
Trust	5.87	0.38	5.15	6.50

The spatial lags are computed using the four nearest neighbors approach. Institutional quality is measured as the first principal component of the 6 WGI

institutional quality, and trust levels.⁶ In the spatial econometrics literature, these new variables are known as spatial lags. They indicate the performance of the geographical neighbors of each country. In this paper, we use the spatial lag of learning, institutions, and trust to capture the influence of the neighbors' performance in the convergence process.

Modeling Framework

We employ an ordered probit model to identify the factors that explain the formation of learning clubs across EU Member States. Bartkowska and Riedl (2012) are among the first to use an ordered regression model to examine the determinants of convergence clusters.

The ordered probit model assigns each country to one convergence club, denoted by $c = 1, \dots, C$. In our case, the endogenous grouping procedure described in “[Convergence Test and Club Identification](#)” section indicates that c takes values from 1 to 4. The club alternatives can be ranked in a logical way according to the steady-state level of each club. Assuming that club membership is driven by a continuous latent variable denoted by y_i^* that indicates the country's individual steady-state, our empirical model can be specified as follows:

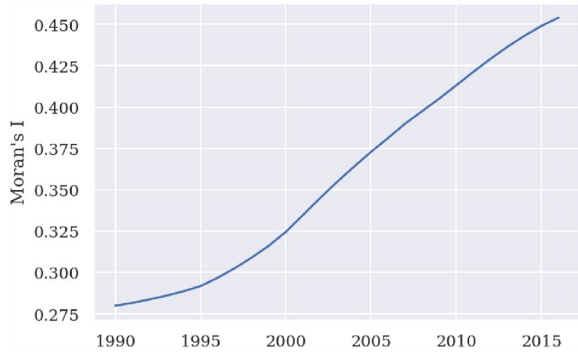
$$y_i^* = \beta X_i + \varepsilon_i, \tag{5}$$

where X_i contains the explanatory variables, including the initial levels of learning, institutional quality and trust as well as their respective spatial lags, i indicates the country, ε_i has a standard normal distribution with mean 0 and variance 1.

⁶ Using instead the six nearest neighbors does not change our key findings.



Fig. 2 Evolution of spatial dependence in learning. Notes: Moran's I is computed using a four nearest neighbors approach



Since y_i^* is unobserved, the model is estimated using a maximum likelihood approach. The observed variable is an ordinal variable $y_i = c$ if $\gamma_{c-1} < y_i^* \leq \gamma_c$ where γ are unknown cut points (or threshold parameters). Since the magnitude of β has no immediate economic interpretation, we report the marginal effects of the predicted probabilities of each variable. The predicted probability that a country belongs to a certain learning club is given by Eq. 6.

$$p_{ic} = p(y_i = c) = \Phi_i(\gamma_c - \beta X_i) - \Phi_i(\gamma_{c-1} - \beta X_i) \tag{6}$$

The marginal effect of an increase in a regressor on the probability of belonging to club c can then be formulated as:

$$\partial p_{ic} / \partial x_{ri} = \{ \Phi_i(\gamma_{c-1} - \beta X_i) - \Phi_i(\gamma_c - \beta X_i) \} \beta_r, 0 < c < C \tag{7}$$

More precisely, the marginal effect indicates how a unit increase in the independent variable changes the probability of being assigned to a specific club while holding all other variables fixed at their sample means.

Results

Convergence Test and Club Identification

Figure 3 presents the relative transition paths for the variables schooling (Panel a) and learning (Panel b) of the 28 EU countries included in our sample.^{7 8} With respect to schooling (our measure of human capital quantity), the heterogeneity across the EU Member States appears to decrease over time, whereas such a convergence tendency is not apparent for the achieved learning outcome (our measure of human capital quality).

⁷ The dynamics of learning are highly consistent with the dynamics of per capita income. A figure of the latter is available in Appendix A.

⁸ Please note that all interactive figures are also available on the following website: <https://carlos-mendez.org/publication/20220616-jces/>



As expected, the $\log-t$ test of convergence of Phillips and Sul (2007a) shows that the null hypothesis of absolute convergence among the EU Member States is not rejected for the “schooling” measure (see first row of Table 3). The fitted coefficient \hat{b} is negative but insignificant, suggesting—according to Phillips and Sul (2009)—that convergence is rather weak. Visual inspection reveals that this tendency might be driven by Germany which is clearly visible diverging from the other EU countries. This intuition is confirmed when excluding Germany from our sample and applying the $\log-t$ test again. As shown in the second row of Table 3, the coefficient turns positive and significant as soon as we exclude Germany from our sample.

The magnitude of \hat{b} (positive but below 2) now suggests convergence in growth rates (that is, relative convergence) but not convergence in levels (that is, absolute convergence). The convergence speed can be computed as $b/2$ and amounts to 6%. We next apply the $\log-t$ test to our measure of human capital quality to investigate whether the EU Member States also converge to a common long-run learning equilibrium. The results are shown in the last row of Table 3. The null hypothesis of overall convergence is rejected since \hat{b} is significantly less than 0 with a t -statistic of -58.2 (that is, far below the critical threshold of -1.65). As already mentioned above, this finding is also illustrated in Fig. 3, Panel (b) since the gaps in learning remain at the end of sample period (in fact, they have even slightly increased).

Even though we do not find evidence for a single long-run equilibrium, we can perform the Phillips and Sul clustering procedure to check whether there exist multiple local equilibria. Our results are presented in Table 4. We identify four significant learning clubs and a diverging group composed by Bulgaria and Romania. The t -statistics for all clubs are greater than -1.65 , suggesting that the null hypothesis of convergence cannot be rejected within each subgroup.⁹ Figure 4 depicts the relative transition paths of the learning clubs which are not converging among themselves but rather moving away from each other instead.¹⁰ In particular, the relative transition paths of Clubs 1 and—to a somewhat lesser extent—also for Club 2 show a clear upward trend, whereas the paths of Clubs 3 and 4 are downward sloping.

Figure 5 shows the spatial distribution of the convergence clubs and the diverging countries. The convergence speed for the first three clubs ranges from 7 to 11%. Club 4 is the smallest cluster (comprising three EU Member States, namely Greece, Malta, and Slovakia) and shows the strongest convergence behavior with a convergence speed of 50%. Strikingly, the lower two clubs comprise mostly countries that joined the EU after 2004 (the so-called New Member States, NMS). There are some exceptions though. For instance, Estonia and Poland managed to join the top learning cluster, and Latvia as well as Slovenia are both members of the second-highest human capital cluster. Overall, the convergence clubs seem to follow (at least to some extent) a geographical pattern since many Western and Northern countries

⁹ We also check whether the original clubs can be merged to shape bigger clubs. The merging analysis (see Table 8 in Appendix C) indicates that the clubs cannot be merged together to form larger convergence clusters. Hence, the four original subgroups are confirmed as the final learning convergence clubs.

¹⁰ Figure 7 in Appendix B offers a more detailed view of the evolution of learning differences within each convergence club.



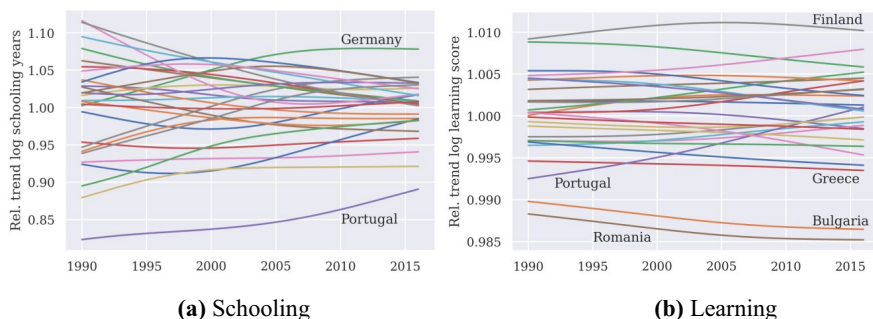


Fig. 3 Transitional paths for schooling and learning. Notes: Schooling is measured as mean years of total schooling across all education levels. Learning is measured using test scores for the age group between 15 to 19 years. For both indicators, the sample includes 28 European countries. Interactive versions of these figures are available at <https://embed.deepnote.com/dfc1465f-493d-4a8b-b11c-8412ba77aed7/6f6d59a2-f068-4848-a39c-8c7021e7a7ad/8b4728031b814d3c8dbdcedd301d10930> and <https://embed.deepnote.com/dfc1465f-493d-4a8b-b11c-8412ba77aed7/6f6d59a2-f068-4848-a39c-8c7021e7a7ad/9e27558c6b28411bb7323b0821742362>

Table 3 log- t convergence test

	Coeff	SE	T -stat
Schooling	-0.02	0.04	-0.5
Schooling(*)	0.12	0.05	2.66
Learning	-0.83	0.01	-58.2

Schooling is measured as mean years of total schooling across all education levels. Learning is measured using test scores for the age group between 15 to 19 years. The symbol (*) indicates that Germany is not included in the sample

converge toward the upper two clusters (again, there are some exceptions). These findings are relatively similar to those obtained by studies analyzing income clubs or institutional clubs within the EU or EA (cf. Bartkowska and Riedl 2012; Von Lyncker and Thoennesen 2017; Glawe and Wagner 2021a, b). However, some Southern EU countries such as Spain, Portugal, and Italy perform on average better with respect to learning than for instance institutional quality (cf. Glawe and Wagner 2021a).¹¹

¹¹ It is important to note that the Phillips and Sul (2007a) algorithm focuses primarily on the growth rate and less on the level of a variable (in our case, the learning score). This also explains why Sweden is found to be in Club 3 while Portugal is assigned to the highest club even though the former records a higher learning score for most of the sample period. It is because of Portugal's strong upward tendency in learning that it is assigned to Club 1 (it in fact records the highest learning score growth rate among all countries in our sample, cf. also Fig. 7 (the interactive online version)). Further discussion on the educational reform efforts in Portugal and Sweden is provided in "Further Discussion: Educational Reforms to Improve Students' Learning Results—What Can We Learn from the Most Successful Countries?" section.

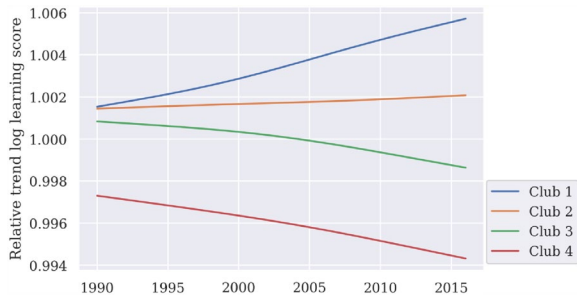


Table 4 Learning convergence clubs and non-convergent group

	Club1	Club2	Club3	Club4	Group5
Coeff	0.22	0.15	0.2	1	-0.34
T-stat	2.94	2.05	2.65	7.68	-7.53

The null hypothesis of convergence is rejected with the *T*-stat is smaller than -1.65

Fig. 4 Convergence clubs in learning: average tendencies



Determinants of Club Membership

Tables 5, 6 and 7 present our baseline regression results. As shown in Panel A of Table 5, the initial level of learning is insignificant; however, the spatial lag of learning is an important driver of club membership (see Panel B). For instance, a higher initial value of learning in its four nearest EU neighboring countries increases a country’s probability of belonging to Club 1, whereas it decreases the probability of belonging to Club 3. That is, geographical location is decisive for whether a country is on a high or low human capital development trajectory.

This is particularly interesting since spatial effects usually only play (at best) a subordinate role for the formation of income clusters within the EU, even at the regional level (see, e.g., Bartkowska and Riedl 2012). Also when adding the non-spatial initial level of learning (in Panel C), the spatial lag stays statistically significant.

Table 6 displays our results regarding the initial level of institutional quality. As expected, the sign is positive for the upper two clubs and negative for the lower two clubs, implying that a higher initial level of institutional quality increases the likelihood of an EU Member State to converge to one of the upper human capital clusters. As shown in Panel B, also the spatial lag of institutional quality is statistically significant. That is, if a country’s neighbors exhibit rather high levels of institutional quality, this increases the probability that the country itself will converge to the highest learning club. Including both institutional variables simultaneously (as shown in Panel C), only the spatial lag of institutional quality remains significant; however, the other institutional variable still has the correct sign.

In a robustness check, we include the six individual WGIs separately in order to gain more insights regarding which dimension of institutional quality matters most for whether a country is on a low or high human capital trajectory. Our findings



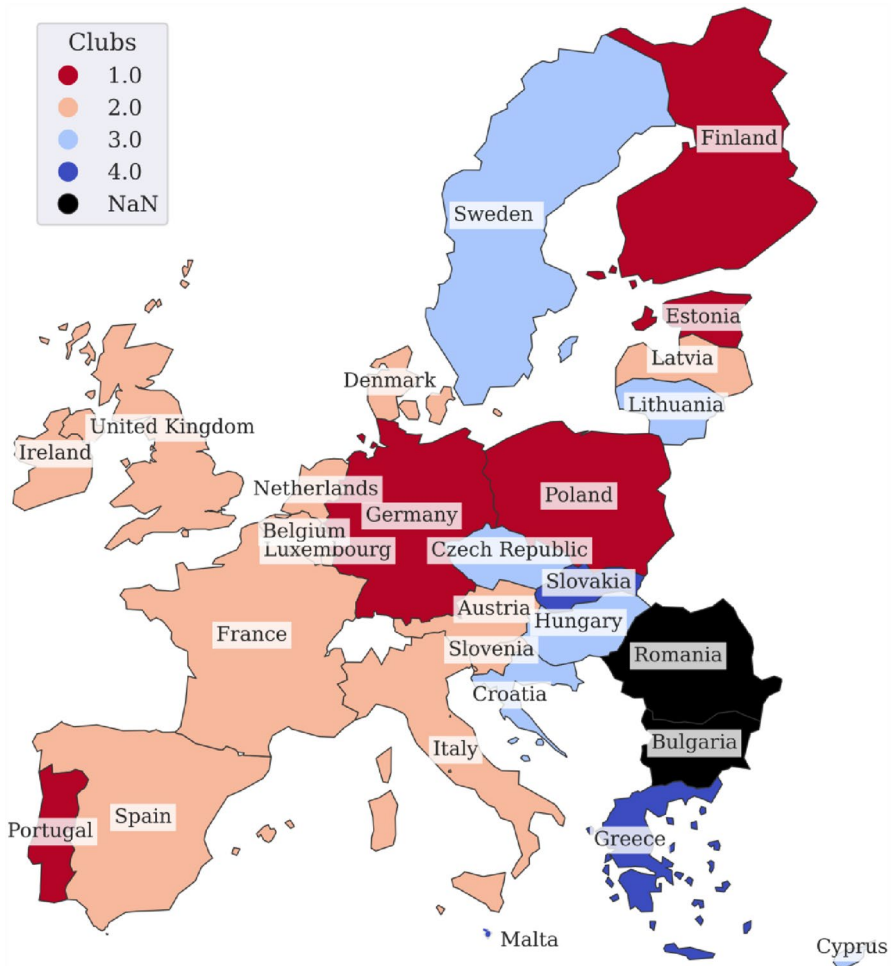


Fig. 5 Spatial distribution of the convergence clubs in learning. Notes: NaN indicates divergent countries

are presented in Appendix D, Tables 9–14. The institutional dimensions “voice and accountability” and “regulatory quality” appear to be most decisive for learning club membership, if we focus on the significance and magnitude of the marginal effects, followed by the indicators “control of corruption”,¹² “government effectiveness”, and “rule of law.” As before, only the spatial lags of the respective institutional indicators remain statistically significant when including the level of institutions and its spatial lag simultaneously. Only for the indicator “voice and accountability” also the non-spatial variable stays significant at the 10-percent level for Club 1.

¹² A scarce corruption control could be perceived as “anti-meritocratic” and may thus reduce the incentive to human capital accumulation.



Table 5 Ordered probit, initial level of human capital

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Initial learning	3.1402 (2.7530)	1.2041 (1.8704)	-2.1994 (2.1956)	-2.1449 (2.2801)
<i>Panel B</i>				
Spatial lag of initial learning	9.5798** (3.9905)	5.4585 (4.4778)	-9.3927** (4.7895)	-5.6457 (3.4639)
<i>Panel C</i>				
Initial learning	1.7978 (2.3478)	1.0732 (1.7522)	-1.8344 (2.5601)	-1.0366 (1.5164)
Spatial lag of initial learning	8.9834** (4.0454)	5.3624 (4.2485)	-9.1663* (4.7941)	-5.1796 (3.2210)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis

Table 6 Ordered probit, institutional quality

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Institutions	0.0572** (0.0230)	0.0259 (0.0282)	-0.0469* (0.0257)	-0.0363 (0.0228)
<i>Panel B</i>				
Spatial lag of institutions	0.0792*** (0.0274)	0.0401 (0.0439)	-0.0716* (0.0386)	-0.0477* (0.0278)
<i>Panel C</i>				
Institutions	0.0340 (0.0251)	0.0188 (0.0223)	-0.0333 (0.0274)	-0.0194 (0.0158)
Spatial lag of institutions	0.0596** (0.0287)	0.0329 (0.0381)	-0.0585 (0.0373)	-0.0341 (0.0240)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis

As displayed in Table 7, the initial level of trust has the expected sign; however, it is insignificant for all clubs/columns (see Panel A). The spatial lag of trust is marginally significant for Clubs 1 and 3 and stays significant when including both measures simultaneously (cf. Panels B and C of Table 7, respectively). Overall, cultural spillover effects appear to play only a minor role.

In a next step, we investigate in how far our results change when including various determinants simultaneously since this can provide us with more insights



Table 7 Ordered probit, culture (trust)

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Trust	0.0740 (0.0555)	0.0277 (0.0377)	-0.0505 (0.0474)	-0.0512 (0.0429)
<i>Panel B</i>				
Spatial lag of trust	0.3623* (0.1849)	0.1624 (0.1551)	-0.2874* (0.1525)	-0.2373 (0.1475)
<i>Panel C</i>				
Trust	0.0124 (0.0590)	0.0056 (0.0276)	-0.0099 (0.0486)	-0.0081 (0.0378)
Spatial lag of trust	0.3502* (0.2029)	0.1581 (0.1523)	-0.2796* (0.1550)	-0.2287 (0.1613)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis

regarding which variable matters most for determining club membership. The respective regression tables can be found in Appendix E.

In Table 15 in Appendix E, we use the spatial lag of learning in combination with the institutional variables. We can observe that as soon as we include the spatial lag of learning, the significance of the initial level of institutional quality decreases to the 10-percent level (cf. Panels A and B) while that of the spatial lag of institutional quality turns completely insignificant and has the wrong sign (cf. Panel B). Please note that both spatially lagged variables (learning and institutional quality) are highly correlated (0.77) which likely explains the change in sign of the institutional variable. It might also be that institutional quality has only a small direct impact on the learning clubs and additionally an indirect effect through the spatial lag of initial learning.¹³

When including simultaneously the spatial lag of learning and the trust variables, the latter are always insignificant (cf. Table 16 in Appendix E). The spatial lag of learning stays significant for Panel A (that is, when only additionally including the non-spatially lagged initial level of trust), but its significance also decreases when additionally including the spatial lag of trust. This is probably again due to the fact that these two spatial variables are highly correlated with each other (with a coefficient of 0.65).¹⁴

In Table 17 in Appendix E, we include the spatial lag of initial learning as well as the initial levels of institutions and trust simultaneously. Our main findings do not

¹³ When additionally including the non-spatial initial level of learning, the spatial lag of learning stays significant while institutional quality turns insignificant.

¹⁴ Additionally including the non-spatial initial level of learning does not change our main findings, the spatial lag of learning remains highly significant while trust is still insignificant (but has the wrong sign).



change: the spatial lag of learning is positively signed and significant at the 5-percent level in Columns (1) and (3), and institutional quality is also positively signed and significant at the 10-percent level. Trust remains insignificant.¹⁵ Overall, the spatial lag of human capital “trumps” institutions and trust, however, the latter two variables (especially institutional quality) probably have an indirect effect through the initial human capital endowment of (neighboring) countries.¹⁶

In a robustness test, we also control for the country-specific “return to education” using data from Montenegro and Patrinos (2014). Our results are reported in Appendix F, Table 18. Please note that our sample size slightly decreases to 25 because of data availability. We find that a higher return to education increases the likelihood that an EU Member State converges to one of the upper two human capital clubs whereas a membership in the lower two clubs becomes less likely. However, the marginal effects are insignificant for all clubs (cf. Panel A of Table 18). Also when adding the variable “return to education” to our previous key specifications, our main findings remain unchanged (cf. Panels B-D of Table 18).

Finally, we check in how far our results change if we control for characteristics of the schooling systems and related variables. In particular, we focus on the organization type of the educational system, the class size, the duration of compulsory education, as well as the (log of the) government education expenditure share in GDP.¹⁷ Please keep in mind that the main focus of our paper is, however, on the *underlying/deep* determinants of learning and not on the rather *proximate* ones (which might be in fact influenced by the institutional setting of a country). Having said this, in the following we briefly summarize our findings. As shown in Table 19 in Appendix G, a smaller class size and a longer duration of compulsory education appear to be in general conducive to learning (even though the marginal effects are insignificant, see, e.g., Panels A1 and B1). Moreover, as expected, a higher share of government education expenditures in GDP increases the likelihood of a Club 1 membership (cf. Panel C1). However, also this effect turns insignificant once we control for the initial level of education and spatial effects (cf. Panel C2). With respect to the organization of the educational system (“single structure”, “common core curriculum”, or “early

¹⁵ Again, additionally including the non-spatial lag of learning does not change these findings. The results are not presented here to save space but are available upon request.

¹⁶ Our key findings do not change when merging Clubs 3 and 4 together. Moreover, our results also remain valid in a probit model setting with Club 1 and Club 2 forming the new Club 1, and Club 3 and Club 4 forming the new Club 2. Finally, we also tested what happens if we include the diverging countries Bulgaria and Romania in Club 4. Again, our findings do not change (also in a probit setting). The results are not displayed here to save space but are available upon request.

¹⁷ Data on government expenditures on education are obtained from the World Bank’s (2022) World Development Indicators. The remaining data come from Eurydice (<https://eacea.ec.europa.eu/national-policies/eurydice/>) and the OECD.



tracking” system)¹⁸ we do not find any meaningful/significant impact on the learning trajectory of EU countries.¹⁹

As already mentioned above, variables depicting characteristics of the educational system might be correlated with our other key variables (such as institutional quality and the initial level of human capital).²⁰ Finally, it is worth noting that it is difficult to adequately measure characteristics of the school systems; it very often depends on the country specific context and a combination of several different factors (see also Andere 2015, for an interesting discussion on this topic). The organization type of the educational system is for instance a very broad measure. Thus, the above results should be interpreted with considerable caution. The next section provides a rather “qualitative” discussion of the educational reform bundles undertaken in educationally successful EU countries (particularly Finland) and elaborates on what other Member States can learn from these experiences.

Further Discussion: Educational Reforms to Improve Students’ Learning Results—What Can We Learn from the Most Successful Countries?

If we take a closer look at the education reforms of the best performing EU countries, we might draw some policy implications on how to improve school quality (keeping in mind that of course no educational system or set of reforms from one country can be easily transferred to another). **Finland** seems to be a particularly good example from which other EU Member States can learn from; it is often regarded as a kind of “poster child for school improvement” (cf. Darling-Hammond 2010). Also according to our club convergence analysis in “[Convergence Test and Club Identification](#)” section, Finland is the top performer of the highest learning cluster within the EU. However, back in the 1970s, Finland was not really succeeding educationally in international comparison, and it was not until comprehensive reforms of the educational system were implemented that Finland managed to secure its educational top position, not only within the EU but also worldwide (cf. Darling-Hammond 2009). Finland first launched reforms to equalize educational opportunity at the beginning of the 1970s.²¹ By the end of the decade, the reform focus then

¹⁸ “Single structure” and “common core structure” systems are both characterized by a general education program followed by all students; however, in a “single structure system” a single institution covers primary and lower secondary education, whereas in the “common core structure” system, primary education and secondary education are provided in two separate institutions. In an “early tracking” system, after completion of primary education, students follow distinct educational pathways or specific types of schooling and ultimately receive different levels of certificates.

¹⁹ We create dummy variables for each of these three types of systems and include them separately in our regressions; however, they all fail to be statistically significant (please note that few countries have “mixed” systems; we thus checked how our results change if we exclude these states or assign them to either one system or the other). These results are not depicted here to save space but are available upon request.

²⁰ As argued by Byrne and Plekhanov (2021), existing educational systems “may reflect behaviours and cultural norms in a given society”.

²¹ Educational equity is often considered as an important success factor for the performance of Finnish students and has remained an important component in the Finnish education policy (cf. Jaku-Sihvonen 2002; Välijärvi et al. 2007; Kumpulainen and Lankinen 2012).



shifted toward enhancing teacher quality.²² Among others, teacher education was improved and extended with the help of intensive investments. During the 1990s, the transfer of higher-order skills (e.g., problem solving capacity and critical thinking) became the center of reforms. One measure to reach this goal was the establishment of research-based master degree programs for teachers (cf. also Buchberger and Buchberger 2004). In general, Finland has a fairly competitive teacher selection process and only around 15% of the college graduates that apply to become teachers are eventually admitted. They subsequently receive a three-year graduate-level teaching preparation program at no cost (Darling-Hammond 2009). With teachers being “professionals who know how to improve learning for all” (cf. Laukkanen 2008, cited in Darling-Hammond 2010), Finland was also able to successfully adopt a more decentralized educational system which allows making use of local creativity (paired with thoughtful guidance on the general goals, cf. Darling-Hammond 2009). Finland’s well-developed institutions might be another important reason for why local autonomy appears to be rather beneficial for learning. For instance, Hanushek et al. (2013) show that local autonomy has an ambiguous or negative effect in countries with weak institutions.

Overall, the Finnish experience tells us that educational reforms targeted at upgrading teaching quality combined with an “empowerment of the teaching profession” appear to be promising strategies to improve the learning capacity and outcome of students (cf. Laukkanen 2006, p. 15).²³ This point is also argued by various empirical studies analyzing micro-data. For instance, using a longitudinal dataset of more than 6,000 students and their teachers, Lee (2018) finds that students who had been taught by a succession of high-performing and qualified teachers had greater short- and long-term educational success.^{24,25} Also teaching practices appear to play an important role for cognitive skill accumulation. Focusing on the USA, Bietenbeck (2014) finds that modern teaching practices (such as working in small groups, giving explanations, and relating the learning content to the students’ daily life) significantly improve pupils’ reasoning skills, whereas traditional teaching practices (such as listening and memorizing facts) only improve the factual knowledge and the solving of routine problems.

In fact, the Finnish reforms seem to have already inspired some countries with respect to their educational strategies. One example is **Sweden**. Sweden has seen a dramatic loss of knowledge among students within a relatively short amount of

²² See also Sahlberg (2006), Laukkanen (2008), Ministry of Education (2008), and Jakku-Sihvonen and Niemi (2006) on the importance of teacher quality for Finland’s education success. For a more critical discussion, see Andere (2015).

²³ However, even though teacher quality appears to be a particularly crucial (school-based) factor, Finland’s success is likely the result of a combination of many factors that are intertwined, as argued among others by Välijärvi et al. (2007) and Andere (2015).

²⁴ Lee (2018) suggests that schools should prioritize hiring teachers with a specialization in the courses they are supposed to teach. Moreover, he recommends that teachers should share students’ data and performance to the subsequent teacher.

²⁵ See also Hanushek et al. (2019) for a literature review on teacher quality.



time,²⁶ and a 2015 OECD report found that the PISA performance of Swedish students is low compared to other OECD countries. Ringarp (2016) argues that Sweden (similar to Germany, see also below) has experienced a so-called PISA-shock in the early 2000s. The key problem areas highlighted in the OECD (2015a) report are the unclear educational priorities, the unattractiveness of the teaching profession, the mismatch between local autonomy and public accountability, as well as learning environments that are not always conducive to learning and not challenging enough (cf. OECD 2015a: p. 8). Similar conclusions were made by the National Agency of Education (NAE) which highlights the number of unqualified teachers, the lacking awareness of teachers with regard to school curricula and grading criteria, as well as grade inflation as important reasons for the stagnation and even decline in knowledge development (Skolverket, NAE 2004a, b). Sweden implemented several reforms with the aim of improving students' educational performance, among others, by establishing teacher certifications and a new teaching training program in 2008, as well as by introducing a new curriculum and new grading criteria in 2011 (cf. Ringarp 2016). Particularly the former two reforms resemble Finland's efforts to enhance teacher quality. However, according to the OECD (2015a) report, more consistent and coherent efforts are necessary, particularly with respect to the further building up of adequate teaching and learning capacity as well as a clearer definition of learning priorities and the establishment of unambiguous responsibilities.

Also in **Germany**, the debate initiated by the "PISA-shock" in 2000 caused some major policy changes as well as a "shift in the idea of education" (cf. Davoli and Entorf 2018). The German reforms focused, among others, on reducing inequalities and improving the achievement of disadvantaged students. Subsequently, test scores increased considerably, as also confirmed by our analysis in "Convergence Test and Club Identification" section (Germany is part of the highest learning club and has seen an upward trend). It is important to note that in Germany, education is regulated by the individual "Bundesländer" (federal states). Davoli and Entorf (2018) highlight seven major points that are usually regarded as having (to a greater or lesser extent) contributed to the rise in learning scores, namely (1) rethinking streaming children at the age of 10 (which has however only been (partly) implemented in some states, namely in Berlin, Brandenburg, and Hesse), (2) softening segregation (by merging the two lower-level track schools into one school in some federal states), (3) standardization of curricula, (4) monitoring and ensuring comparability, (5) introduction of central examinations ("Zentralabitur"), (6) increasing school autonomy, and (7) expanding and strengthening the educational content of pre-primary schools. During the 2010s, the impact of socio-economic background on the learning performance of German students had indeed declined (cf. OECD

²⁶ This is in line with the findings of our club convergence analysis as Sweden belongs to the second-lowest learning cluster and even though it is among the best performing Club 3 countries (together with Czech Republic) it has seen a continuous downward trend. Still, the international perception is that Sweden is an educationally rather well performing country. This impression probably still stems from earlier achievements. In Sweden, education has always been a public priority and the promotion of learning for all students as well as lifelong learning are regarded as crucial (cf. OECD 2015a). Of course, Sweden's level of learning is not low; however, the country has seen a rather steep decline (see also Footnote 10).



2015b; Davoli and Entorf 2018) and the share of “resilient” students (that is, students that score among the worldwide top 25% despite a disadvantageous socio-economic background) had increased by 9 percentage points (cf. OECD 2016). Even though the German reforms emphasize educational equity and in this respect remind one of Finland’s focus on equalized educational opportunity, there still exist considerable differences between these two countries (for instance, Finland has a single structure, non-grade oriented educational system whereas Germany still relies on an early-tracking system, cf. “Determinants of Club Membership” section). How does the picture look with respect to teaching quality and teacher education, one of Finland’s key reform areas? Germany’s initial teacher education is also often assessed as being relatively strong (even though the German teacher education system lacks coherence, cf. Terhart 2008 and Kotthoff and Terhart 2013 for a more detailed discussion) and there exist rigorous tests that teacher candidates have to pass in order to enter teacher education programs. All German teachers have the equivalent of a bachelor and master’s degree in education and specialize in a content area (cf. Rothman 2017). After having completed a “first phase” of teacher education at a university, teacher candidates subsequently work for about one year at a German school with an experienced mentor during the so-called second phase which focuses primarily on practical training (“Referendariat”). However, the fragmentation of teacher education (cf. Kotthoff and Terhart 2013) as well as lacking ongoing professional learning for teachers are identified as problem areas that have to be addressed in the future (cf. Rothman 2017).

After having discussed the successful Finnish educational reforms as well as the experiences of Sweden and Germany, which both experienced a so-called PISA-shock in the early 2000s and subsequently installed comprehensive school reforms (with varying results), in the following, we will briefly focus on the reform efforts of Portugal, Estonia, and Poland. These three countries all managed to improve their students’ learning performance relative to the other EU countries until the middle of the 2010s—that is, the end of our study period—and they are all members of the highest learning cluster (even though they are probably less prominent examples of educationally well-performing countries compared to Finland).

Even though **Portugal** has the lowest learning score within its learning cluster, its catching up tendency since the 1990s has been remarkable (Crato 2020, even calls Portugal’s development a “quantum leap”, p. 209). From 2001 onward, Portugal has implemented several educational reforms. Reform efforts intensified particularly during the 2011–2015 period and were oriented toward quality learning. The key reform areas included a demanding and well-structured curriculum, frequent and reliable assessments, the provision of high-quality education for all, the promotion



of equal opportunity and demanding education, as well as ensuring teaching quality (through the introduction of entrance exams for teacher candidates, the regulation of access to teachers' colleges and education programs at universities, an extension of requirements for teacher preparation programs, and the inclusion of additional subject-related content in these programs, see Crato 2020, pp. 227–229 for more details). However, the policy course changed in 2016 and afterward, students' educational performance started to decrease according to some international test results; however, this period is not covered in our analysis. This more recent development in Portugal is a reminder that persistent and coherent reform efforts are necessary to ensure long-term success of the educational reforms.

Estonia has also seen a steady upward trend within its learning cluster, especially since 2000. During the 1990s, Estonia installed comprehensive reforms with respect to secondary and higher education which starkly resemble Finland's reforms of the 1980s. Reforms included, among others, additional teacher training under the independent school program of 1994, a high degree of decentralization and school autonomy, and the introduction of additional accountability measures (cf. OECD 2001; Lees 2016; and Byrne and Plekhanov 2021 for more detailed information). Byrne and Plekhanov (2021) exploit the fact that the extensive educational reforms in Estonia in the early 1990s were only implemented throughout Estonian-speaking schools (and only much later and less comprehensive in Russian-speaking schools) to study the effectiveness of reforms. Using difference-in-difference estimation, the authors find that the Estonian educational reforms resulted in improvement of around 15% of the standard deviation in terms of adult skills. Also our analysis has shown that Estonia is the one of the best-performing countries in the highest learning cluster, converging more and more toward Finland.

In **Poland**, a comprehensive structural reform of the educational system was introduced in 1999 (Jakubowski 2021). Key objectives of the reform included the enhancement of the quality and effectiveness of the educational system, the guaranteeing of equal educational opportunity, and a rise in educational attainment (Wiśniewski and Zahorska 2020). These goals should be achieved via the introduction of a redesigned national core curriculum, the introduction of external exams at the end of each cycle of schooling (that is, primary, lower secondary, and upper secondary), greater school-level autonomy which should promote innovative teaching methods, and, probably most importantly, the reduction of primary education from 8 to 6 grades and the related establishment of a new three-year general lower secondary school level (which in fact extended comprehensive general education by one year, cf. Jakubowski 2021). In the second half of the 2000s, the Ministry of Education then installed a curriculum reform with a focus on the development of cognitive, analytical, and non-routine problem-solving skills (Wiśniewski and Zahorska



2020). Even though Poland has seen considerably rising test scores since the onset of reforms (as also supported by our analysis), “politically minded commentators” were rather critical toward the reforms (Jakubowski 2021, p. 138) and there was a kind of “collective public opinion that schools were not working properly” (Wiśniewski and Zahorska 2020, p. 183), leading to a partial reversal of reforms around 2016.

Overall, our qualitative analysis of educational reforms in selected EU Member States has shown that various countries seem to have been—either intentionally or unintentionally—influenced by Finland’s reform process. Overall, teacher quality appears to be among the most important school-based factors affecting learning. Moreover, the reform experiences of the Club 1 members have shown that enhancing equal educational opportunity is likely another important feature of a successful educational reform program. It has also become clear that reforms often take time; if there are no coherent strategies and persistent efforts, the success of educational policies will very probably only be short lived (this can be to some extent currently experienced in Portugal). This makes it politically challenging to initiate and then uphold reforms aimed at improving school quality: they are costly and benefits might only unfold with a time lag. Furthermore, a country’s institutional framework is likely decisive for the success of educational reforms: It is questionable whether educational reforms that are conducted in a country with well-developed institutions would have the same effect in a country with a less favorable institutional environment.²⁷ Higher levels of political instability and/or populism may also make a reversal of reforms more likely. That is, educational reforms are important; however, they are less likely to unfold their benefits if the foundation, the underlying factors of development, are lacking. This is also confirmed by our ordered probit analysis.

Concluding Remarks and Policy Implications

Real convergence has always been an important—if not the most important—goal of the European integration process. In order to achieve a sustainable harmonization of per capita incomes across the EU Member States, it is decisive that there is convergence in the underlying factors of economic development (such as physical and human capital as well as institutional quality). However, the analysis of convergence dynamics in the factors driving income convergence has been neglected for a long time, especially compared to the vast literature dealing with income convergence.

²⁷ In this context, it is worth noting that Estonia is found to also converge to the highest institutional cluster within the EU (cf. Glawe and Wagner 2021a).



Only recently, studies have started to examine the convergence patterns of, for example, institutional quality, also within the EU (cf. Beyaert et al. 2019; Glawe and Wagner 2021a).

Our paper contributes to this literature by focusing on human capital which is considered as an important convergence factor by standard growth theory. Besides being among the very first papers concerned with human capital (club) convergence, another important contribution of our research is that we distinguish between convergence in school quantity and in school quality. Taking into account the quality of education rather than focusing solely on quantitative indicators appears to be particularly important to gain insights into income convergence processes within comparatively well-developed regions such as the EU which already have a relative high endowment with human capital compared to other geographical areas such as Africa or South Asia. In particular, a sufficiently high level of human capital quality might be decisive for successfully mastering an innovation-based growth strategy which, in turn, is crucial for sustaining growth at higher income levels.

Using a nonlinear dynamic factor model, we find that while there is absolute convergence with respect to the average years of schooling, there exist four convergence clusters with respect to learning outcomes (that is, school quality), and the heterogeneity between those four clubs is increasing over time. While the differences in learning are not extremely large in absolute terms (even though they are certainly non-negligible), the fact that the learning trajectories of the four clubs exhibit an extremely pronounced divergence tendency (manifesting in steep slopes of relative transition paths) is particularly worrisome. The increasing heterogeneity in learning scores might lead to a further bifurcation of the EU and could even trigger an upper middle-income trap in some Eastern European countries since human capital is found to be a decisive factor in the context of growth slowdowns at the middle-income range (cf. Glawe and Wagner 2016; Agénor 2017), even more so in the context of the ongoing digitalization process (cf. Glawe and Wagner 2020).

In the second part of our paper, we study the determinants of the learning clubs by using an ordered probit model. Our results show that, besides the initial spatial lag of learning, institutional quality is another determinant of learning clubs. We also find evidence of significant institutional spillover effects in some specifications. This could imply that the currently observable institutional deterioration in some EU Member States might hamper the human capital accumulation of neighboring countries and thus, ultimately endanger real convergence and the stability of the European Union. The worsening institutional situation in Poland and Hungary could thus lead to a kind of self-reinforcing “downward spiral” in Eastern Europe. Between



2010 and 2020, Hungary has seen a considerable downward trend regarding the institutional indicators “voice and accountability” as well as “regulatory quality”, and Poland regarding the institutional dimensions “voice and accountability” and “political stability/violence avoidance”, diverging more and more from Western European states such as Germany (cf. the WGI database; Kaufmann et al. 2010).

In general, policy measures should aim at improving institutional quality all over the EU but with a particularly strong emphasis on the New Member States and other countries that are located in one of the lower learning clubs. One problem in this respect is that violations of common law principles etc. in individual EU Member States (like recently observable in Hungary and Poland) are difficult to sanction since these countries can in turn threaten with a veto on important other votes (for instance, regarding the Corona reconstruction fund). One solution is to abrogate the unanimity rule, which, in turn, is afflicted with serious other problems (e.g., the potential “exploitation” of some members by the majority; cf. Glawe and Wagner 2021b, for an extensive discussion).

While institutional quality is certainly an important factor for learning club membership, our findings suggest that learning spillover effects among EU neighbor countries are even more important for accomplishing the transition toward a higher learning trajectory (they “trump” institutions and institutional spillovers once included simultaneously in a regression model). Educational reforms should focus especially on school *quality* and should not be restricted to individual countries since the strong spatial spillover effects call for joint efforts. Against this background, two recently adopted initiatives of the European Commission appear to be a step in the right direction. In particular, the *Vision of the European Education Area by 2025* (EEA) “proposes new initiatives, more investment and stronger cooperation of Member States to help all Europeans, of all ages, benefit from the EU’s rich education and training offer” (cf. European Commission 2020). The EEA is underpinned by six dimensions, including (educational) quality, digital transition, as well as higher education. It is complemented by the *Digital Education Action Plan (2021–2027)* with a focus on the enhancement of digital competences. While these two initiatives sound promising, it remains to be seen whether they prove successful in improving educational quality in the EU.

Taking a closer look at the reform measures implemented by the educationally best performing EU countries (such as Finland) reveals that educational reforms targeted at upgrading teaching quality combined with an empowerment of the teaching



profession appear to be promising strategies to improve the learning capacity and learning outcome of students. This can be, among others, achieved through the introduction of teacher certification programs and in general a stronger research orientation of teacher degree programs. If this foundation is established, greater local autonomy can also prove conducive to students' knowledge creation if being paired with adequate accountability mechanisms, clear general guidelines, and a favorable institutional environment. Several countries that show strong catching-up tendencies (such as Portugal and Estonia) with respect to students' learning performance have successfully implemented reforms that follow the "Finnish method."

Finally, it is noteworthy that spatial spillovers might not only be restricted to within the EU but also involve EU candidate countries. Also, these spillovers may have temporal adjustment lags that limit the spatial diffusion of learning in the short run. Such analyses, however, are beyond the scope of our current paper but certainly constitute interesting topics for future research.

Appendix A: Dynamics of Per-capita Income

See Fig. 6.

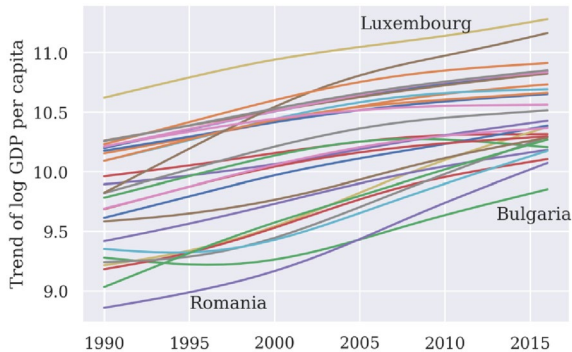


Fig. 6 Dynamics of per-capita income in EU countries. Notes: Consistent with the methodology of Phillips and Sul (2007a), the long-run trends of GDP per capita are computed using the HP filter with a smoothing parameter of 400. An interactive version of this figure is online available at: <https://embed.deepnote.com/dfc1465f-493d-4a8b-b11c-8412ba77aed7/791b85ea-394f-4043-871c-97a0fe09ab86/4aac90f0-7dc6-4d39-929a-2ed579ec5aa7?height=601>



Appendix B: Convergence Clubs in Learning

See Fig. 7.

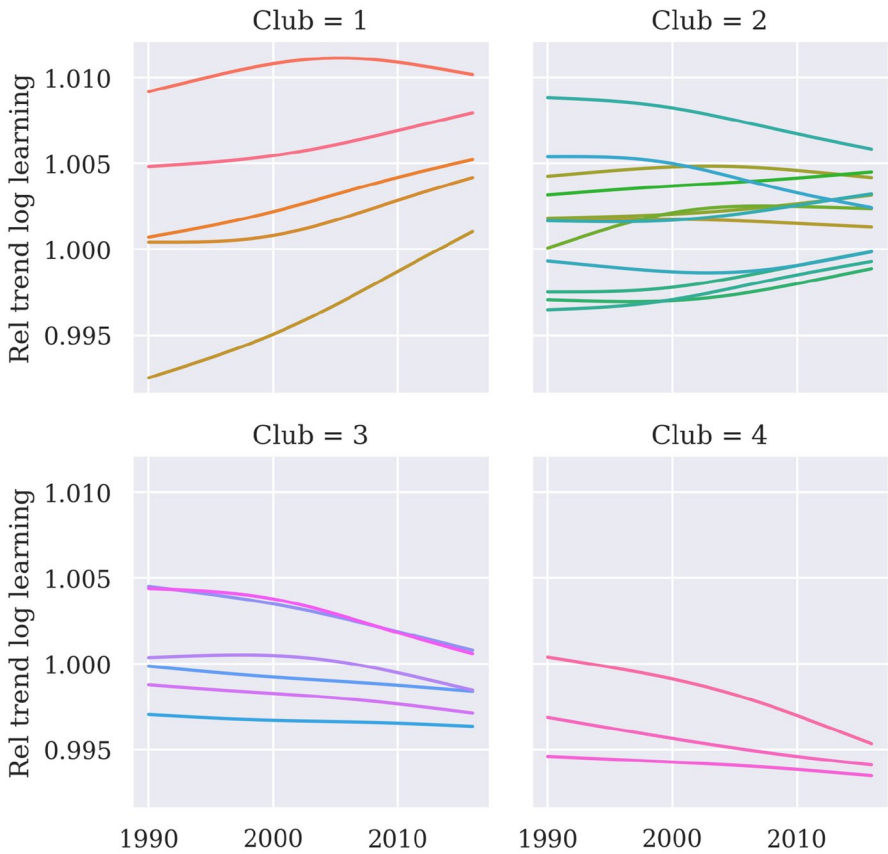


Fig. 7 Convergence clubs in learning. Notes: An interactive version of this figure is online available at: <https://embed.deepnote.com/dfe1465f-493d-4a8b-b11c-8412ba77aed7/6f6d59a2-f068-4848-a39c-8c7021e7a7ad/00015-6b46c060-195d-4202-b08d-4a245d49f6d9?height=600.96875>

Appendix C: Merge Test of the Convergence Clubs in Learning

See Table 8.

Table 8 Merge test of the convergence clubs

	Club 1+2	Club 2+3	Club 3+4
Coeff	-0.12	-0.24	-0.28
<i>T</i> -stat	-2.12	-5.88	-6.27

The null hypothesis of convergence is rejected with the *T*-stat is smaller than -1.65



Appendix D: Determinants of Learning Clubs, Individual WGI

See Tables 9, 10, 11, 12, 13 and 14.

Table 9 Ordered probit, voice and accountability (VA)

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Institutions (VA)	0.2998** (0.1268)	0.1343 (0.1401)	-0.2428* (0.1322)	-0.1914 (0.1227)
<i>Panel B</i>				
Spatial lag of institutions (VA)	0.4490*** (0.1595)	0.2275 (0.2377)	-0.4033* (0.2142)	-0.2732* (0.1552)
<i>Panel C</i>				
Institutions (VA)	0.2126* (0.1202)	0.1244 (0.1261)	-0.2189 (0.1349)	-0.1180 (0.0815)
Spatial lag of institutions (VA)	0.3623** (0.1612)	0.2121 (0.2259)	-0.3732* (0.2173)	-0.2012 (0.1342)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis

Table 10 Ordered probit, political stability and violence avoidance (PV)

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Institutions (PV)	0.1784 (0.1109)	0.0695 (0.0768)	-0.1255 (0.0992)	-0.1224 (0.0781)
<i>Panel B</i>				
Spatial lag of institutions (PV)	0.3494* (0.1877)	0.1450 (0.1959)	-0.2647 (0.2046)	-0.2298 (0.1727)
<i>Panel C</i>				
Institutions (PV)	0.1279 (0.1221)	0.0559 (0.0655)	-0.1011 (0.1036)	-0.0827 (0.0711)
Spatial lag of institutions (PV)	0.2831 (0.2007)	0.1237 (0.1822)	-0.2238 (0.2016)	-0.1831 (0.1772)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis



Table 11 Ordered probit, government effectiveness (GE)

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Institutions (GE)	0.1761** (0.0826)	0.0753 (0.0940)	-0.1370 (0.0926)	-0.1144 (0.0788)
<i>Panel B</i>				
Spatial lag of institutions (GE)	0.2510*** (0.0873)	0.1246 (0.1465)	-0.2239* (0.1281)	-0.1517* (0.0921)
<i>Panel C</i>				
Institutions (GE)	0.0621 (0.1129)	0.0314 (0.0683)	-0.0565 (0.1076)	-0.0370 (0.0694)
Spatial lag of institutions (GE)	0.2085* (0.1144)	0.1055 (0.1307)	-0.1896 (0.1324)	-0.1244 (0.0926)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis

Table 12 Ordered probit, regulatory quality (RL)

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Institutions (RL)	0.2589*** (0.0956)	0.1232 (0.1309)	-0.2214** (0.1041)	-0.1607 (0.1011)
<i>Panel B</i>				
Spatial lag of institutions (RL)	0.4035*** (0.1452)	0.2550 (0.2313)	-0.4398** (0.1899)	-0.2187* (0.1284)
<i>Panel C</i>				
Institutions (RL)	0.1337 (0.0828)	0.0939 (0.1079)	-0.1602 (0.1209)	-0.0674 (0.0531)
Spatial lag of institutions (RL)	0.3324** (0.1398)	0.2334 (0.2099)	-0.3982** (0.1788)	-0.1676 (0.1020)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis



Table 13 Ordered probit, rule of law (RL)

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Institutions (RL)	0.1649** (0.0746)	0.0705 (0.0804)	-0.1278 (0.0825)	-0.1076 (0.0681)
<i>Panel B</i>				
Spatial lag of institutions (RL)	0.2538*** (0.0928)	0.1207 (0.1420)	-0.2177* (0.1311)	-0.1568 (0.0955)
<i>Panel C</i>				
Institutions (RL)	0.1063 (0.0880)	0.0551 (0.0710)	-0.0989 (0.0915)	-0.0625 (0.0558)
Spatial lag of institutions (RL)	0.2003* (0.1054)	0.1038 (0.1294)	-0.1864 (0.1314)	-0.1178 (0.0893)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis

Table 14 Ordered probit, control of corruption (CC)

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Institutions (CC)	0.1509** (0.0659)	0.0693 (0.0770)	-0.1248* (0.0694)	-0.0954 (0.0616)
<i>Panel B</i>				
Spatial lag of institutions (CC)	0.2020*** (0.0725)	0.0988 (0.1132)	-0.1774* (0.0997)	-0.1233* (0.0744)
<i>Panel C</i>				
Institutions (CC)	0.0917 (0.0764)	0.0488 (0.0595)	-0.0868 (0.0745)	-0.0537 (0.0492)
Spatial lag of institutions (CC)	0.1396* (0.0804)	0.0743 (0.0932)	-0.1322 (0.0998)	-0.0817 (0.0617)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis



Appendix E: Determinants of Learning Clubs, Ordered Probit Model

See Tables 15, 16 and 17.

Table 15 Ordered Probit, institutions and learning

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Spatial lag of learning	8.1759** (4.0376)	5.6005 (4.4475)	-9.5600** (4.6087)	-4.2164 (2.8811)
Institutions	0.0407* (0.0220)	0.0279 (0.0276)	-0.0476 (0.0290)	-0.0210 (0.0177)
<i>Panel B</i>				
Spatial lag of learning	8.2655 (5.3921)	5.6602 (3.9102)	-9.6617** (4.8765)	-4.2640 (3.1988)
Institutions	0.0410* (0.0233)	0.0281 (0.0245)	-0.0480* (0.0256)	-0.0212 (0.0173)
Spatial lag of institutions	-0.0013 (0.0347)	-0.0009 (0.0232)	0.0015 (0.0402)	0.0007 (0.0178)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis

Table 16 Ordered Probit, trust and learning

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Spatial lag of learning	9.2654** (4.1907)	5.3234 (4.4162)	-9.1719* (4.8474)	-5.4169 (3.4821)
Trust	0.0198 (0.0560)	0.0114 (0.0346)	-0.0196 (0.0566)	-0.0116 (0.0338)
<i>Panel B</i>				
Spatial lag of learning	7.4320* (4.2879)	4.3337 (4.3133)	-7.4001 (5.3439)	-4.3656 (3.0857)
Trust	0.0035 (0.0578)	0.0020 (0.0341)	-0.0035 (0.0579)	-0.0020 (0.0339)
Spatial lag of trust	0.1522 (0.2136)	0.0887 (0.1196)	-0.1515 (0.1880)	-0.0894 (0.1380)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis



Table 17 Ordered Probit, all variables

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
Spatial lag of learning	8.4525** (4.1951)	5.8403 (4.6112)	-9.9368** (4.7350)	-4.3560 (2.9189)
Institutions	0.0451* (0.0239)	0.0312 (0.0307)	-0.0530* (0.0313)	-0.0232 (0.0195)
Trust	-0.0275 (0.0565)	-0.0190 (0.0415)	0.0323 (0.0660)	0.0142 (0.0308)
Observations	26	26	26	26

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis

Appendix F: Robustness Check: Controlling for the “Return to Education”

See Table 18.

Table 18 Ordered Probit, return to education

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A</i>				
Returns to education	0.0062 (0.0314)	0.0017 (0.0080)	-0.0035 (0.0168)	-0.0045 (0.0225)
<i>Panel B</i>				
Spatial lag of learning	12.8129** (5.5184)	6.9599 (6.5457)	-11.8869* (7.1512)	-7.8860* (4.1067)
Initial learning	1.6544 (2.4171)	0.8987 (1.5452)	-1.5348 (2.2599)	-1.0182 (1.6241)
Returns to education	0.0035 (0.0267)	0.0019 (0.0137)	-0.0033 (0.0238)	-0.0022 (0.0165)
<i>Panel C</i>				
Spatial lag of learning	11.5583** (5.6202)	6.8580 (6.6353)	-12.1011* (6.9455)	-6.3151* (3.6553)
Institutions	0.0381* (0.0224)	0.0226 (0.0256)	-0.0399 (0.0253)	-0.0208 (0.0182)
Returns to education	0.0043 (0.0264)	0.0026 (0.0150)	-0.0045 (0.0266)	-0.0024 (0.0147)
<i>Panel D</i>				
Spatial lag of learning	13.2655** (5.6122)	7.2310 (6.8259)	-12.3545* (7.1475)	-8.1420* (4.3933)
Trust	0.0518 (0.0656)	0.0282 (0.0419)	-0.0482 (0.0605)	-0.0318 (0.0438)
Returns to education	0.0059 (0.0288)	0.0032 (0.0147)	-0.0055 (0.0254)	-0.0036 (0.0179)
Observations	25	25	25	25

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis



Appendix G: Robustness Check: Controlling for School Characteristics and Related Variables

See Table 19.

Table 19 Ordered Probit, return to education

	(1)	(2)	(3)	(4)
	Club 1	Club 2	Club 3	Club 4
<i>Panel A1</i>				
Class size	0.0149 (0.0239)	0.0015 (0.0060)	-0.0082 (0.0132)	-0.0081 (0.0152)
<i>Panel A2</i>				
Spatial lag of learning	12.2343** (5.3117)	2.2757 (4.9371)	-8.9242 (6.2032)	-5.5858* (3.2553)
Initial learning	0.9474 (3.8851)	0.1762 (0.7575)	-0.6911 (2.7717)	-0.4326 (1.8065)
Class size	0.0083 (0.0238)	0.0015 (0.0059)	-0.0060 (0.0173)	-0.0038 (0.0119)
Observations	20	20	20	20
<i>Panel B1</i>				
Duration compulsory educ.	0.0106 (0.0318)	0.0037 (0.0126)	-0.0068 (0.0213)	-0.0076 (0.0228)
<i>Panel B2</i>				
Spatial lag of learning	9.5743** (4.5810)	5.7745 (4.4664)	-9.8267* (5.4839)	-5.5221 (3.4587)
Initial learning	1.6265 (2.3001)	0.9810 (1.7372)	-1.6694 (2.5353)	-0.9381 (1.5028)
Duration compulsory educ.	-0.0166 (0.0306)	-0.0100 (0.0179)	0.0171 (0.0317)	0.0096 (0.0167)
Observations	26	26	26	26
<i>Panel C1</i>				
Gov. education expend.	0.5802* (0.3090)	0.2237 (0.2490)	-0.4237* (0.2432)	-0.3803* (0.2303)
<i>Panel C2</i>				
Spatial lag of learning	7.2572* (3.7314)	3.9781 (3.7212)	-7.0962* (4.1861)	-4.1391 (2.9099)
Initial learning	0.6468 (2.2231)	0.3545 (1.3429)	-0.6324 (2.2337)	-0.3689 (1.3251)
Gov. education expend.	0.3722 (0.2829)	0.2040 (0.1709)	-0.3639 (0.2302)	-0.2123 (0.1884)
Observations	25	25	25	25

(*), (**), (***) denote significance at the 10, 5, and 1% level. Robust s.e. are in parenthesis. 'Gov. education expend.' stands for 'the log of the government education expenditure share in GDP' and 'educ.' Stands for education



Acknowledgements We would like to thank the editor and two anonymous referees for their constructive and insightful comments.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acemoglu, D., F.A. Gallego, and J.A. Robinson. 2014. Institutions, human capital, and development. *Annual Review of Economics* 6(1): 875–912.
- Acemoglu, D., S. Johnson, and J.A. Robinson. 2001. The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91(5): 1369–1401.
- Acemoglu, D., S. Johnson, and J.A. Robinson. 2005. Institutions as a fundamental cause of long-run growth. *Handbook of Economic Growth* 1: 385–472.
- Agénor, P.-R. 2017. Caught in the middle? The economics of middle-income traps. *Journal of Economic Surveys* 31(3): 771–791.
- Aghion, P., and P. Howitt. 1992. A model of growth through creative destruction. *Econometrica* 60: 323–351.
- Andere, E. 2015. Are teachers crucial for academic achievement? Finland educational success in a comparative perspective. *Education Policy Analysis Archives/Archivos Analíticos De Políticas Educativas* 23: 1–24.
- Barro, R.J., and J.W. Lee. 2013. A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics* 104: 184–198.
- Barro, R.J., and X. Sala-i-Martin. 1992. Convergence. *Journal of Political Economy* 100(2): 223–251.
- Barro, R.J., X. Sala-i-Martin, O.J. Blanchard, and R.E. Hall. 1991. Convergence across states and regions. *Brookings Papers on Economic Activity* 1991(1): 107–182.
- Bartkowska, M., and A. Riedl. 2012. Regional convergence clubs in Europe: Identification and conditioning factors. *Economic Modelling* 29(1): 22–31.
- Basile, R., R. Capello, and A. Caragliu. 2012. Technological interdependence and regional growth in Europe: Proximity and synergy in knowledge spillovers. *Papers in Regional Science* 91(4): 697–722.
- Baumol, W.J. 1986. Productivity growth, convergence, and welfare: What the long-run data show. *American Economic Review* 76(5): 1072–1085.
- Benhabib, J., and M. Spiegel. 2005. Human capital and technology diffusion. In *Handbook of Economic Growth*, ed. P. Aghion and S. Durlauf, 935–966. Amsterdam: Elsevier.
- Beyaert, A., J. García-Solanes, and L. Lopez-Gomez. 2019. Do institutions of the Euro Area converge? *Economic Systems* 43(3–4): 100720.
- Bietenbeck, J. 2014. Teaching practices and cognitive skill. *Labour Economics* 30: 143–153.
- Bjørnskov, C. 2009. Social trust and the growth of schooling. *Economics of Education Review* 28(2): 249–257.
- Buchberger, F., and I. Buchberger. 2004. Problem-solving capacity of a teacher education system as a condition of success? An analysis of the “Finnish Case.” In *Education policy analysis in a comparative perspective*, ed. F. Buchberger and S. Berghammer, 222–237. Trauner: Linz.
- Byrne, K., and A. Plekhanov. 2021. Education reforms and adult skills: Evidence from Estonia. *Economics of Education Review* 82: 102106.



- Caragliu, A., and P. Nijkamp. 2016. Space and knowledge spillovers in European regions: The impact of different forms of proximity on spatial knowledge diffusion. *Journal of Economic Geography* 16(3): 749–774.
- Cohen, D. 1996. Tests of the “convergence hypothesis”: Some further results. *Journal of Economic Growth* 1(3): 351–361.
- Coleman, J.S. 1988. Social capital in the creation of human capital. *American Journal of Sociology* 94: S95–S120.
- Crato, N. 2020. Curriculum and educational reforms in Portugal: An analysis on why and how students’ knowledge and skills improved. In *Audacious education purposes*, ed. F.M. Reimers. Cham: Springer. https://doi.org/10.1007/978-3-030-41882-3_8.
- Darling-Hammond, L. 2009. *The Flat World and Education: How America’s Commitment to Equity Will Determine Our Future*. New York: Teachers College Press.
- Darling-Hammond, L. 2010. What we can learn from Finland’s successful school reform. Stanford Center for Opportunity Policy in Education. <https://edpolicy.stanford.edu/library/publications/543>
- Davoli, M. and Entorf, H. 2018. The PISA shock, socioeconomic inequality, and school reforms in Germany. IZA Policy Paper No. 140.
- De la Fuente, A., and R. Doménech. 2006. Human capital in growth regressions: How much difference does data quality make? *Journal of the European Economic Association* 4(1): 1–36.
- Dearmon, J., and K. Grier. 2009. Trust and development. *Journal of Economic Behavior & Organization* 71(2): 210–220.
- Dearmon, J., and R. Grier. 2011. Trust and the accumulation of physical and human capital. *European Journal of Political Economy* 27(3): 507–519.
- Dias, J., and E. Tebaldi. 2012. Institutions, human capital, and growth: The institutional mechanism. *Structural Change and Economic Dynamics* 23(3): 300–312.
- Dollar, D., and A. Kraay. 2003. Institutions, trade and growth. *Journal of Monetary Economics* 22: 1147–1175.
- Durlauf, S.N., and P.A. Johnson. 1995. Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics* 10(4): 365–384.
- Ertur, C., and W. Koch. 2007. Growth, technological interdependence and spatial externalities: Theory and evidence. *Journal of Applied Econometrics* 22(6): 1033–1062.
- European Commission. 2020. Achieving a European Education Area by 2025 and resetting education and training for the digital age. https://ec.europa.eu/commission/presscorner/detail/e%20n/ip_20_1743.
- Filmer, D., H. Rogers, N. Angrist, and S. Sabarwal. 2020. Learning-adjusted years of schooling (LAYS): Defining a new macro measure of education. *Economics of Education Review* 77: 101971.
- Fritsche, U., and V. Kuzin. 2011. Analysing convergence in Europe using the nonlinear single factor model. *Empirical Economics* 41(2): 343–369.
- Galor, O. 1996. Convergence? Inferences from theoretical models. *The Economic Journal* 106(437): 1056–1069.
- Glawe, L., and H. Wagner. 2016. The middle-income trap: Definitions, theories and countries concerned—A literature survey. *Comparative Economic Studies* 58(4): 507–538.
- Glawe, L., and H. Wagner. 2020. The middle-income trap 2.0: The increasing role of human capital in the age of automation and implications for developing Asia. *Asian Economic Papers* 19(3): 40–58.
- Glawe, L., and H. Wagner. 2021a. Convergence, divergence, or multiple steady states? New evidence on the institutional development within the European Union. *Journal of Comparative Economics* 49(3): 860–884.
- Glawe, L., and H. Wagner. 2021b. Divergence tendencies in the European integration process: A danger for the sustainability of the E(M)U? *Journal of Risk and Financial Management* 14(3): 104.
- Glawe, L., and H. Wagner. 2022. Is schooling the same as learning? - The impact of the learning-adjusted years of schooling on growth in a dynamic panel data framework. *World Development* 151: 105773.
- Guiso, L., P. Sapienza, and L. Zingales. 2006. Does culture affect economic outcomes? *Journal of Economic Perspectives* 20(2): 23–48.
- Hanushek, E.A., M. Piopiunik, and S. Wiederhold. 2019. Do smarter teachers make smarter students?: International evidence on teacher cognitive skills and student performance. *Education next* 2019: 57–64.
- Hanushek, E.A., and L. Woessmann. 2008. The role of cognitive skills in economic development. *Journal of Economic Literature* 46(3): 607–668.
- Hanushek, E.A., S. Link, and L. Woessmann. 2013. Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics* 104: 212–232.



- Islam, N. 2003. What have we learnt from the convergence debate? *Journal of Economic Surveys* 17(3): 309–362.
- Jakku-Sihvonen, R. 2002. Differences between school-specific results in terms of equality in education. In *Evaluation of the equal opportunities in the Finnish comprehensive school 1998–2001*, ed. R. Jakku-Sihvonen and J. Kuusela, 9–26. Finland: National Board of Education.
- Jakku-Sihvonen, R., and H. Niemi. 2006. Introduction to the Finnish education system and teachers' work. In *Research-based teacher education in Finland: Reflections by Finnish teacher educators*, ed. R. Jakku-Sihvonen and H. Niemi, 7–13. Turku: Finnish Educational Research Association.
- Jakubowski, M. 2021. Poland: Polish education reforms and evidence from international assessments. In *Improving a country's education*, ed. N. Crato. Cham: Springer.
- Karlsson, C., and U. Gråsjö. 2019. Knowledge flows, knowledge externalities, and regional economic development. In *Handbook of regional science*, ed. M. Fischer and P. Nijkamp, 413–437. Berlin: Springer.
- Kaufmann, D., Kraay, A., and Mastruzzi, M. 2010. The worldwide governance indicators: Methodology and analytical issues. World Bank Policy Research Working Paper 5430, 5430.
- Kotthoff, H.-G., and E. Terhart. 2013. 'New' solutions to 'old' problems? Recent reforms in teacher education in Germany. *Revista Española De Educación Comparada* 22: 73–92.
- Knack, S., and P. Keefer. 1995. Institutions and economic performance: Cross-country tests using alternative institutional measures. *Economics & Politics* 7(3): 207–227.
- Kumpulainen, K., and T. Lankinen. 2012. Striving for educational equity and excellence: Evaluation and assessment in Finnish basic education. In *Miracle of education: The principles and practices of teaching and learning in Finnish schools*, ed. H. Toom, Niemi and A. Kallioniemi, 68–81. Rotterdam: Sense Publishers.
- La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R.W. Vishny. 1997. Legal determinants of external finance. *The Journal of Finance* 52(3): 1131–1150.
- Laukkanen, R. 2006. Finnish strategy for high-level education for all. Paper presented at the conference Educational Systems and the Challenge of Improving Results, University of Lausanne, Lausanne, Switzerland, 15–16 September 2006.
- Laukkanen, R. 2008. Finnish strategy for high-level education for all. In *Governance and performance of education systems*, ed. N.C. Soguel and P. Jaccard, 305–332. Netherlands: Springer.
- Lee, S.W. 2018. Pulling back the curtain: Revealing the cumulative importance of high performing, highly qualified teachers on students' educational outcome. *Educational Evaluation and Policy Analysis* 40(3): 359–381.
- Lees, M. 2016. *Estonian education system 1990–2016: Reforms and their impact*. Estonia: Ministry of Education of Estonia.
- Lim, S.S., R.L. Updike, A.S. Kaldjian, R.M. Barber, K. Cowling, H. York, J. Friedman, R. Xu, J.L. Whisnant, H.J. Taylor, et al. 2018. Measuring human capital: A systematic analysis of 195 countries and territories, 1990–2016. *The Lancet* 392(10154): 1217–1234.
- Lucas, R.E. 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22(1): 3–42.
- Mankiw, N.G., D. Romer, and D.N. Weil. 1992. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics* 107(2): 407–437.
- Ministry of Education. 2008. OECD PISA survey: Finland's success in the PISA survey. Finland.
- Monfort, M., J.C. Cuestas, and J. Ordóñez. 2013. Real convergence in Europe: A cluster analysis. *Economic Modelling* 33: 689–694.
- Montenegro, C. E. and Patrinos, H. A. 2014. Comparable estimates of returns to schooling around the world. World Bank Policy Research Working Paper 7020.
- Nelson, R.R., and E.D. Phelps. 1966. Investment in humans, technological diffusion, and economic growth. *American Economic Review* 56(1/2): 69–75.
- OECD. 2001. Reviews of national policies for education: Estonia 2001. *OECD Centre for Co-Operation with Non-Members*. <https://doi.org/10.1787/9789264189966-en>.
- OECD. 2015a. *Improving Schools in Sweden: An OECD Perspective*. Paris: OECD.
- OECD. 2015b. *Germany, in Education Policy Outlook 2015: Making Reforms Happen*. Paris, France: OECD Publishing. Retrieved from. <https://doi.org/10.1787/9789264225442-24-en>.
- OECD. 2016. Germany. Country Note—Results from PISA 2015. Paris, France: OECD Publishing. Organisation for Economic Co-Operation and Development (2018). *Catching Up? Country Studies on Intergenerational Mobility and Children of Immigrants*. OECD, Paris.



- O'Neill, D. 1995. Education and income growth: Implications for cross-country inequality. *Journal of Political Economy* 103(6): 1289–1301.
- Phillips, P.C., and D. Sul. 2007a. Transition modeling and econometric convergence tests. *Econometrica* 75(6): 1771–1855.
- Phillips, P.C., and D. Sul. 2007b. Some empirics on economic growth under heterogeneous technology. *Journal of Macroeconomics* 29(3): 455–469.
- Phillips, P.C., and D. Sul. 2009. Economic transition and growth. *Journal of Applied Econometrics* 24(7): 1153–1185.
- Pritchett, L. 2013. *The rebirth of education: Schooling ain't learning*. CGD Books.
- Quah, D. 1996. Regional convergence clusters across Europe. *European Economic Review* 40(3–5): 951–958.
- Ringarp, J. 2016. PISA lends legitimacy: A study of education policy changes in Germany and Sweden after 2000. *European Educational Research Journal* 15(4): 447–461.
- Rodrik, D., A. Subramanian, and F. Trebbi. 2004. Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth* 9(2): 131–165.
- Romer, P.M. 1990. Endogenous technological change. *Journal of Political Economy* 98(5, Part 2): S71–S102.
- Rothman, B. 2017. After the shock: The German education system in 2017. The National Center on Education and the Economy (NCEE). <https://ncee.org/quick-read/after-the-shock-the-german-education-system-in-2017/>.
- Roser, M. and Ortiz-Ospina, E. 2016. Global education. *Our World in Data*. <https://ourworldindata.org/global-education>.
- Sab, R. and Smith, S. C. 2001. Human capital convergence: International evidence. IMF Working Paper 32, Washington D.C.: International Monetary Fund.
- Sahlberg, P. 2006. Education reform for raising economic competitiveness. *Journal of Educational Change* 7(4): 259–287.
- Sala-i Martin, X.X. 1996. The classical approach to convergence analysis. *The Economic Journal* 106(437): 1019–1036.
- Shultz, G. P. and Hanushek, E. A. 2012. Education is the key to a healthy economy. *Wall Street Journal*, 1.
- Skolverket [NAE]. 2004a. *Skolverkets lägesbedömning 2004a*. Stockholm: Skolverket.
- Skolverket [NAE]. 2004b. *Handlingsplan för en rättssäker och likvärdig betygssättning*. 2004b-02-23. Stockholm: Skolverket.
- Temple, J. 2002. Growth effects of education and social capital in the OECD countries. *Historical Social Research/Historische Sozialforschung* 27(4): 5–46.
- Terhart, E. 2008. Die Lehrerbildung. In *Das Bildungswesen in der Bundesrepublik Deutschland: Strukturen und Entwicklungen im Überblick*, ed. K. Cortina, L. Baumert, A. Leschinsky, K.U. Mayer, and L. Trommer, 745–772. Reinbek: Rowohlt.
- Topel, R. 1999. Labor markets and economic growth. In *Handbook of labor economics*, ed. O. Ashenfelter and D. Card, 2943–2984. Amsterdam: Elsevier.
- Von Lyncker, K., and R. Thoennessen. 2017. Regional club convergence in the EU: Evidence from a panel data analysis. *Empirical Economics* 52(2): 525–553.
- Väljjarvi, J., Kupari, P., Linnakylä, P., Reinikainen, P., Sulkunene, S., Törnroos, J., and Arffman, I. 2007. The Finnish success in PISA—and some reasons behind it: PISA 2003. Jyväskylä, Finland.
- Wiśniewski, J., and M. Zahorska. 2020. Reforming education in Poland. In *Audacious education purposes*, ed. F. Reimers. Cham: Springer.

