Letters in Spatial and Resource Sciences
https://doi.org/10.1007/s12076-022-00310-y

ORIGINAL PAPER

Policy spillover effects on student achievement: evidence from PISA

Clara-Christina E. Gerstner[1] · Emmanuel S. Tsyawo[2]

Received: 1 February 2022 / Accepted: 7 June 2022

Abstract
National education reforms do not occur in isolation. Countries look towards each other to identify ways that improve the quality of their education systems. When evaluating the effect of an education policy, it is worth considering both local effects of the policy and its spillover effects on other countries. Ignoring spillover effects between countries can lead to biased estimates of policy effects and suboptimal decision making. This paper examines spillover effects of one widespread education policy, school autonomy, on student achievement using three waves of data from the Programme for International Student Assessment (PISA). The spatial autoregressive model is applied to capture both spillover and local effects of school autonomy. Overall, school autonomy raises student achievement in Reading, Mathematics, and Science. We confirm the existence of positive and statistically significant average spillover effects; thus, estimates based on linear regression underestimate the impact of school autonomy. Our findings indicate that there is spatial dependence in student achievement across countries linked to the geographic proximity between countries. Possible extensions of this work are discussed.

Keywords PISA · Student achievement · School autonomy · Spillover effect · Spatial autoregressive model

JEL Classification C21 · I20

We thank Michael Rovine, Thanh Lu, Keisha Solomon, Eric Kadio, and Ashley Mcfarlane for valuable comments.

✉ Emmanuel S. Tsyawo
emmanuel.tsyawo@um6p.ma

Clara-Christina E. Gerstner
clarag@upenn.edu

1 Graduate School of Education, University of Pennsylvania, 3700 Walnut St, Philadelphia, PA 19104, USA

2 AIRESS and FGSES, Université Mohammed VI Polytechnique, Rabat, Morocco

Published online: 18 July 2022

1 Introduction

National education reforms do not occur in isolation. Countries look towards each other to identify best practices that can improve the quality of their education systems. For instance, many European countries granted greater autonomy to schools around the same time in the 1990s due to common geographical, political, or organisational aims and conditions (Coghlan and Desurmont 2007). Policy decisions and outcomes in one country can affect decisions and outcomes in close countries. Besides geographic proximity, closeness between countries can be defined in terms of political, cultural, and economic similarities. Such interdependencies between countries are not always taken into account when examining the impact of policies. Yet, ignoring spatial interdependence in the evaluation of education policies results in an inadequate understanding of a policy's effectiveness (Obinger et al. 2013). Spatial analytic approaches allow one to overcome this issue by estimating local effects of a policy and its spillover effects on other countries.[1] Although remarkable progress has been realised in the estimation of spillover effects in micro-level studies, specifically in the peer effects literature (see, e.g., Manski (1993), Bramoullé et al. (2009), and Lin (2010), little attention has been devoted to the study of spillover effects at the country level.[2]

Spatial thinking in the analysis of education policy matters now more than ever as globalisation enhances policy borrowing between countries (Larsen and Beech 2014; Gulson and Symes 2007).[3] Governments study and implement policies that proved successful in other countries, react to external pressures to adopt new policies, or strategically respond to policies adopted in neighbouring countries. Obinger et al. (2013) identifies four mechanisms of policy diffusion—learning, emulation, competition, and coercion. In the context of education reforms, learning and competition appear to be the most relevant mechanisms. Learning occurs when countries adopt "lessons learned" by other countries in order to lessen the uncertainty and implementation costs associated with adopting a new policy. For example, France switched from administering standardised student assessments at the beginning of the academic year to the end of the year—a practice which is more common in neighbouring countries (Baird 2016).[4] Learning is tightly linked to competition; governments consider other countries' policy decisions and outcomes to increase their comparative advantage (Obinger et al. 2013). Different mechanisms of policy diffusion are not easily distinguished in empirical studies. It is challenging (if not infeasible) to reliably identify which causal mechanisms drive policy transfers (i.e., spillover effects) using observational data—see, e.g.,

[1] LeSage and Pace (2010) refers to local and spillover effects as direct and indirect effects—see also Fischer et al. (2009), LeSage and Pace (2010), and LeSage and Pace (2013).

[2] Teltemann and Windzio (2018) takes a first step in this direction by testing and controlling for spatial proximity between countries in an analysis of student performance. The authors do not, unlike this paper, quantify spillover effects.

[3] The process by which countries influence each others' policy decisions is defined as policy borrowing, transfer, or diffusion—see, e.g., Meseguer and Gilardi (2009).

[4] Due to push-back from parents and teachers, the 2007 reform was revoked in 2013.

Obinger et al. (2013), Sacerdote (2011), and Bramoullé et al. (2019). Hence, this paper does not study specific transfer mechanisms, but focuses rather on the overarching conditions (e.g., contextual factors) that foster education policy transfers.

International organisations such as the Organisation for Economic Co-operation and Development (OECD) foster learning and competition between countries by providing policy recommendations and metrics to assess the success of policies. In particular, the OECD provides a platform for competition to flourish through large-scale assessments such as the Programme for International Student Assessment (PISA hereafter). Various researchers have argued and provided evidence that PISA influences national policy-making processes (Baird 2016; Forestier and Crossley 2015; Grek 2009; Parcerisa et al. 2020). Most policy transfer studies draw their findings from single case studies, e.g, examples of one country borrowing policies from another country, expert interviews with education policy makers, or document analyses of international collaborations. For example, these qualitative works suggest that PISA is driving educational policy borrowing through increasing competition and social pressure on national policy makers when confronted with their countries' performance in international rankings (Sellar and Thompson 2017). Countries such as Norway, Germany, and Spain have experienced a "PISA shock" following the first rounds of PISA results which sparked fierce debates concerning educational change and motivated reforms (Breakspear 2012). As a result, for instance, German policy makers and scholars travelled to "PISA wonderlands" such as Finland and Sweden to identify reforms that might benefit them (Ringarp and Rothland 2010). Similarly, experts from Denmark, Norway, and Sweden showed increased interest in Finish policies after the publication of early PISA results; they tried to identify differences in Scandinavian systems and learn from the Finish PISA success story (Breakspear 2012).

Policy transfer theories highlight that international interdependence between countries is a key driver in the dissemination and adoption of policies (Parcerisa et al. 2020; Phillips and Ochs 2003; Obinger et al. 2013). Countries often look towards "reference societies", i.e., other countries that serve as models worth imitating and learning from (Waldow 2017). As the examples above show, such reference societies are typically high-performing PISA countries such as Finland which are studied very closely. However, policy borrowing processes also depend on other contextual factors such as geographical, linguistic and cultural similarities between countries (Phillips and Ochs 2003). Although countries such as Hong Kong and Singapore are ranked among the top-performing PISA countries, European policy makers and scholars have argued that student populations in these countries are too dissimilar to successfully transfer policies (Gill and Benton 2013; Gray and Galton 2011). Countries may be more likely to learn and borrow from countries that are considered peers and share more similarities. Moreover, PISA policy recommendations are often used to justify and validate existing policy proposals rather than adopt new and vastly different policies from other countries (Cantley 2019). Thus, for example, it is unlikely that the UK would completely switch to a Hong Kong-style education system, but rather policy makers selectively choose policies that are in line with the existing education system (Forestier and Crossley 2015).

Finally, researchers have argued that specific conditional factors which drive policy diffusion are testable (Holzinger and Knill 2005; Obinger et al. 2013). Conditional factors are defined as forms of proximity between countries that define the intensity and nature of interdependencies. Geographic distance is presumably the most widely used conditional factor. Closeness between countries can increase, e.g., the exchange of information and labour. Other important conditional factors include common language, shared cultural heritage, and economic relations between countries, e.g., OECD membership.[5] Spatial econometric models can be used to test different forms of proximity between countries and identify them as potential vehicles of policy spillovers. In such models, the strength and nature of bilateral relationships (e.g., geographic and economic) between countries are defined via the spatial matrix.

Increasing school autonomy is one of several policy reforms promoted by the OECD which is theorised to improve student achievement—see, e.g., Maslowski et al. (2007), Wößmann (2007), and Hanushek et al. (2013). School autonomy measures the degree of decentralisation in decision making and control by local actors. Schools with higher autonomy have the freedom to make independent decisions about, e.g., their budget, hiring processes, and curriculum. The main argument for increasing autonomy is that it puts schools in a better position to meet students' demands, e.g., buying textbooks and developing course content. Previous studies suggest that higher school autonomy in personnel management (Maslowski et al. 2007), textbook choice, and hiring decisions (Fuchs and Wößmann 2008) can improve efficiency of public schools and, thereby, positively impact student achievement especially in high income countries (Hanushek et al. 2013). School autonomy policies often go hand in hand with school accountability (Coghlan and Desurmont 2007; Wößmann 2007), and higher parental involvement (Ammermueller 2013). Examples of countries increasing school autonomy include the United Kingdom which responded to its low 2009 PISA results by expanding the "free schools" programme, awarding more autonomy and control to schools as this strategy proved effective in other countries (Baird 2016). Similarly, France introduced reforms in response to a widening PISA achievement gap by allowing schools to manage budgets more autonomously and personalise educational resources. However, the network effect of such policy changes through the interdependence between countries (i.e., the spillover dimension) remains unexplored.

This paper contributes to the literature by examining the spillover effect of school autonomy on student achievement across countries. Teltemann and Windzio (2018) is perhaps the closest to this paper. Unlike Teltemann and Windzio (2018), which treats spatial interdependencies as nuisance parameters, this paper focusses on quantifying spillover effects which we interpret as evidence that policy borrowing between countries and its impact on student achievement exist. Like Hanushek and Wößmann (2006), and Hanushek et al. (2013), we construct a panel using data sourced from PISA. The data comprise 56 OECD and non-OECD countries and

[5] In the online appendix, we explore other conditional factors such as GDP per capita, OECD membership, and income categories.

span three waves, namely 2009, 2012, and 2015. We estimate spillover effects of school autonomy on student achievement in Reading, Mathematics, and Science. In addition, we study different subdomains of autonomy, e.g., autonomy in choosing textbooks vs. autonomy in creating assessments. We also examine the effectiveness of autonomy policies by sub-samples, e.g., European and OECD member countries.

We confirm the existence of positive average spillover effects based on geographic proximity between countries. Based on our preferred specification, a 10 percentage point increase in school autonomy increases reading achievement by 25.9% of a standard deviation. More than 40% of the total effect can be attributed to spillover effects. In examining four subdomains of school autonomy, we find the largest effects for autonomy in designing assessments and the smallest for choosing course content. Our main findings do not hold for personnel autonomy, but we find smaller positive effects for budget autonomy. Thus, autonomy in designing the academic content delivered in schools appears to be a more important driver of student test performance than autonomy in selecting teaching personnel or allocating budgets. Our robustness analyses indicate that our findings are robust to plausible variations in specifications.

The outline of the rest of the paper is as follows. Section 2 provides a description of the dataset and the construction of the spatial matrix. In Sect. 3, we present the spatial autoregressive model and parameters of interest such as local and spillover effects. We discuss empirical findings in Sect. 4, and conclude in Sect. 5. Details of the econometric model and additional empirical results are available in an online appendix.

2 Data and spatial matrix

The OECD has conducted PISA every three years since 2000. PISA aims to monitor trends in student performance across both OECD and non-OECD countries. To ensure comparability, PISA targets 15-year-old students in each participating country regardless of institutional affiliation or grade level. The PISA sampling procedure consists of a two-stage cluster-sample design to ensure that samples are representative of each country's target population (OECD. PISA 2016).

PISA performance tests are computer-administered or paper and pencil tests that last up to two hours. Test items include both multiple-choice and open-ended questions which are constructed to test a range of relevant skills and competencies in Mathematics, Reading, and Science. In addition to test scores, PISA also provides various background variables through questionnaires administered to students, teachers, principals, and sometimes parents. In this study, we use student performance measures and school information from three waves of PISA. We limit the analysis to 56 countries which are observed for all three waves.

Table 1 provides summary statistics of all variables (pooled across all three waves) included in our model specifications.[6] The outcome variables include PISA

[6] Country-level summary statistics are provided in the online appendix.

Table 1 Summary statistics

	Mean	SD
PISA reading score	471.411	44.004
PISA mathematics score	469.061	52.401
PISA science score	475.552	47.720
Autonomy—Academic Content	0.739	0.226
Autonomy in assessments	0.772	0.208
Autonomy in textbooks	0.828	0.244
Autonomy in course content	0.688	0.260
Autonomy in course choice	0.667	0.277
Autonomy—Personnel	0.434	0.253
Autonomy—Budget	0.632	0.224
External monitoring	0.716	0.201
Private school share	0.238	0.223
Shortage of instruct. materials		
Strongly	0.079	
Not at all	0.379	
GDP per capita (1000 $)	32.862	25.972
OECD membership	0.625	
World Bank Country and Lending Groups		
High income country	0.714	
Upper-middle income country	0.250	
Lower-middle income country	0.036	

Mean mean weighted by sampling probabilities, *SD* standard deviation for continuous variables only

reading, mathematics, and science scores. Autonomy and its four subdomains are the policy variables of interest. Alongside country and year dummies, observed school-level and country-level characteristics are included as controls in order to help distinguish policy spillovers from spatial patterns such as correlated shocks—see, e.g., König et al. (2019, 483) and 2010. Student observations with missing background measures are dropped.[7]

Measures of school autonomy are obtained based on the answers of principals on the school questionnaire survey. In line with previous studies, e.g., Hanushek et al. (2013), and Teltemann and Windzio (2018), we first code subdomain measures indicating full autonomy at the school level when the principal, the school's board, department heads, or teachers are key decision makers in the school (coded as one and zero otherwise). Four subdomains are averaged to construct a single variable for School Autonomy (also called Academic Content Autonomy). The four subdomains include (i) autonomy in creating assessments, (ii) autonomy in choosing textbooks, (iii) autonomy in deciding on course content, and (iv) autonomy in choosing

[7] For variables included in our analyses, this constitutes less than 5% of student observations.

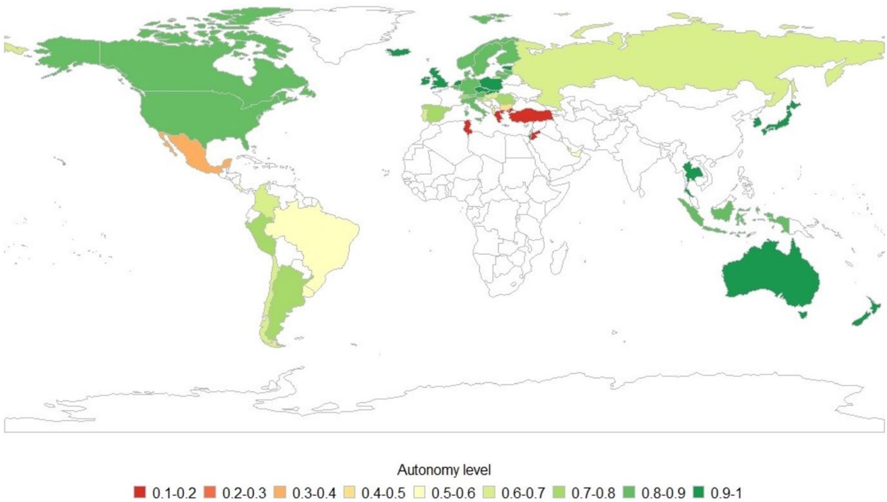

Fig. 1 School autonomy in 2015

courses. Personnel and Budget Autonomy measures are constructed based on six other subdomains, e.g., responsibility for hiring and firing teachers, and budget allocation—see Hanushek et al. (2013) for further discussion.

Figure 1 displays school autonomy levels of countries in the 2015 sub-sample. We observe that autonomy levels are the highest in Australia, Japan, parts of Europe (like the Netherlands and the Czech Republic) and Great Britain. Autonomy levels are also above 80% in the US, Canada, and Indonesia. The lowest autonomy levels are observed in Greece, Tunisia, Turkey, and Mexico (all below 40%). Autonomy levels do not change markedly across the three waves.

In addition to school autonomy, we include an accountability measure (external monitoring of school performance), private school status, and shortage of instructional materials as school-level variables in our models—see also West and Woessmann (2010), and Hanushek et al. (2013). All measures are obtained from PISA school background questionnaires. For private school status, we create a binary variable for school type (coded as one for private and zero for public). We impute missing values for private school status and run a robustness check, which shows that excluding the variable does not alter our findings (see specification (1) Table 5 below).[8] The accountability measure is coded as one for schools that publicly share test information and zero for schools that do not share information. Two measures indicating the existence or absence of a shortage of instructional materials are added. Each school-level variable is aggregated at the country level using school-specific weights. OECD membership, Gross Domestic Product (GDP), and the World Bank

[8] Private school status values for Sweden and Israel were missing in 2015. These two missing values were imputed using the respective country averages of 2009 and 2012.

Country and Lending Group Categories are added as country-level controls. These economic measures are included to account for countries' economic conditions and the quality of skills and institutions.

Aggregation of data to the country level is done because this paper focusses on inter-country analyses of spillover effects. Moreover, aggregation to the country level eliminates bias in estimates arising from within-country selection, selection into schools by students, and general equilibrium effects (Hanushek et al. 2013, pp. 213, 219). Also, spatial connectivity used in the construction of the spatial matrices is observed at the country level. Student-level analyses of inter-country spillover effects become challenging because capturing spatial connectivity between students within-country relative to students between-country is prone to ambiguity. To buttress this point further, student-level analyses in the peer effects literature are already quite advanced e.g., Schneeweis and Winter-Ebmer (2007), Bramoullé et al. (2009), and Lin (2010). For this reason, the marginal contribution of this paper in such a direction may not be weighty. Finally, institutional factors such as autonomy or accountability measures often display little intra-country variation as they are defined within national legal frameworks (Wößmann 2007; Coghlan and Desurmont 2007).[9]

We construct spatial matrices using information on the geographic distance between capital cities. We extract data on longitudes and latitudes from the GeoDist database (Mayer and Soledad 2011) and compute geographic distances between capital cities. A contiguity matrix that specifies $w_{ij} = 1$ for country-pairs with shared borders (and zero otherwise) is not considered because it isolates some countries. The row in $\boldsymbol{w}$ corresponding to an isolated country comprises 0/0, i.e., is not defined, when $\boldsymbol{w}$ is row-normalised. Examples of isolated countries in the data include Tunisia, New Zealand, Australia, and Iceland. To circumvent this problem, Teltemann and Windzio 2018 uses a distance threshold of 2500 km between capitals to construct a spatial matrix of 0/1 binaries. The choice of a cut-off implies that the spatial matrix is a function of the cut-off and by consequence, the parameters are also a function of the chosen cut-off. We opt for a non-binary geographic spatial matrix in order to avoid sensitivity of results to threshold choice. The (i, j)'th entry of the geographic spatial matrix is computed using the inverse of the geographic distance between the capital cities of countries i and j and the spatial matrix is row normalised.

3 Methodology

Spatial dependence arises due to interactions between units or unobserved characteristics in space. This paper uses the spatial autoregressive (SAR) model. The SAR model is a spatial extension of linear regression; it addresses violations of the independence assumption by taking account of spatial correlations. Spillover effects

[9] For example, countries either have nationwide exit exams or they do not.

are typically estimated using spatial econometric techniques (e.g., LeSage and Pace 2010, and Fischer et al. 2009).

3.1 The SAR model

The SAR is given by

$$y_{it} = \rho \sum_{j=1}^{N} w_{ij} y_{jt} + \gamma_0 + \boldsymbol{x}_{it}\boldsymbol{\gamma} + \eta_i + \delta_t + \varepsilon_{it} \tag{1}$$

where y_{it} is the outcome, $\boldsymbol{x}_{it}$ is a $1 \times k$ vector of covariates, η_i is the country i fixed effect, δ_t is the year t fixed effect, and $\boldsymbol{\theta} = (\rho, \gamma_0, \boldsymbol{\gamma}', \eta_1, \dots, \eta_{N-1}, \delta_1, \dots, \delta_{T-1})'$ constitute the vector of parameters.[10] ε_{it} is assumed independently distributed across i with $\mathbb{E}(\varepsilon_{it}) = 0$ for each $t \in \{1, \dots, T\}$. ε_{it} and $\varepsilon_{it'}$, are allowed to be arbitrarily correlated and heteroskedastic for all $i = 1, \dots, N$ and $t, t' = 1, \dots, T$. By convention, $w_{ij} \geq 0$ is the (i, j) 'th element of the $N \times N$ spatial matrix $\boldsymbol{w}_N$, $w_{ii} = 0$, and $\sum_{j=1}^{N} w_{ij} = 1$ due to row-normalisation. The spatial matrix defines the nature of interdependence between countries, e.g., geographical proximity. Stacking across $n = NT$ country-year observations, (1) in matrix notation is

$$\mathbf{y}_n = \rho \boldsymbol{w}_n \mathbf{y}_n + \mathbf{x}_n \boldsymbol{\beta} + \boldsymbol{\varepsilon}_n \tag{2}$$

where $\boldsymbol{\beta} = (\gamma_0, \boldsymbol{\gamma}', \eta_1, \dots, \eta_{N-1}, \delta_1, \dots, \delta_{T-1})'$, $\mathbf{x}_n$ absorbs the constant term, $\boldsymbol{x}_{it}$, country dummies, and year dummies. $\boldsymbol{w}_n$ is an $n \times n$ block-diagonal matrix $\mathbf{I}_T \otimes \boldsymbol{w}_N$. Following the spatial econometrics literature (e.g., Kelejian and Prucha 1998, Lee 2007, and Anselin et al. 2008), the spatial matrix $\boldsymbol{w}_n$ is assumed known. GMM is used for estimation, and the set of instruments comprises linearly independent columns of $[\mathbf{x}_n, \boldsymbol{w}_n \mathbf{x}_n, \boldsymbol{w}_n^2 \mathbf{x}_n, \dots]$—see, e.g., Kelejian and Prucha (1998), and Bramoullé et al. (2009).

One encounters a number of identification issues in estimating (1)—see, e.g., König et al. (2019, sect. VI.B) for a discussion. Relevant ones to this paper include the simultaneity of outcomes y_{it} and y_{jt}, and correlated shocks.[11] As standard practice in the spatial econometrics literature, we solve the simultaneity problem by instrumenting $w_{ij} y_{jt}$ with $(w_{ij} x_{1,jt}, \dots, w_{ij} x_{k,jt})$. Following the literature, e.g., König, Liu, and Zenou (2019, p. 483) and Bramoullé et al. (2019, sect. 3.4), we include fixed effects in order to alleviate the concern of correlated shocks. Time-varying correlated shocks that are not adequately captured by a one- or two-way fixed effect as in (1), e.g., one with an interactive fixed effect structure $\eta_i \times \delta_t$ are, however, a potential concern that is not accounted for in this paper. The well-known reflection problem of Manski (1993) is naturally addressed by an intransitive spatial matrix $\boldsymbol{w}_n$ (Bramoullé et al. 2009; Lin 2010; Bramoullé et al. 2019).[12] The geographic spatial

[10] To avoid perfect collinearity with the constant term, η_N and δ_T are dropped.

[11] Correlated shocks arise when factors, e.g., a "PISA shock" or global economic shocks, cause countries to react similarly.

[12] $\boldsymbol{w}_n$ is intransitive when a neighbour's neighbour is not a neighbour.

matrix as used in this paper satisfies the criterion of intransitivity because the entries are continuous and vary across country-pairs.

3.2 Partial effects

Econometric models containing spatially lagged outcomes add a richness to the model by taking account of different forms of cross-sectional dependence—see, e.g., LeSage and Pace (2013), and Anselin et al. (2008). They do, however, require a more careful interpretation of parameters unlike linear regression models whose coefficients have a straightforward interpretation due to the assumed linearity and independently sampled observations. The interpretation of $(\rho, \boldsymbol{\gamma}')'$ is structural—see e.g., Bramoullé et al. (2009). A positive (negative) ρ suggests a country's outcome is increasing (decreasing) in neighbours' outcomes. An element in $\boldsymbol{\gamma}$ gives the partial effect of the corresponding covariate on the outcome after controlling for spatial correlation.[13]

From (2), $(\mathbf{I}_n - \rho \boldsymbol{w}_n)\mathbf{y}_n = \mathbf{x}_n \boldsymbol{\beta} + \boldsymbol{\varepsilon}_n$ leads to the reduced-form

$$\mathbf{y}_n = (\mathbf{I}_n - \rho \boldsymbol{w}_n)^{-1}\mathbf{x}_n \boldsymbol{\beta} + (\mathbf{I}_n - \rho \boldsymbol{w}_n)^{-1}\boldsymbol{\varepsilon}_n = \sum_{j=1}^{k} \boldsymbol{S}_n(\rho, \beta_j)\mathbf{x}_{.j,n} + \boldsymbol{V}_n(\rho)\boldsymbol{\varepsilon}_n$$

where $\boldsymbol{S}_n(\rho, \beta_j) = (\mathbf{I}_n - \rho \boldsymbol{w}_n)^{-1}\beta_j = \mathbf{I}_T \otimes (\mathbf{I}_N - \rho \boldsymbol{w}_N)^{-1}\beta_j$ is an $n \times n$ block diagonal matrix, $\mathbf{x}_{.j,n}$ is the j'th column of $\mathbf{x}_n$, and $\boldsymbol{V}_n(\rho) = (\mathbf{I}_n - \rho \boldsymbol{w}_n)^{-1}$. From the foregoing, a partial effect of $x_{\imath j,t}$ on y_{it} has the form $\frac{\partial y_{it}}{x_{\imath j,t}} = \boldsymbol{S}_n(\rho, \beta_j)_{i\imath}$, $\imath \in \{1, \dots, n\}$ for all $t = 1, \dots, T$, and $\boldsymbol{S}_n(\rho, \beta_j)_{i\imath}$ denotes the $(i, \imath)$'th element of $(\mathbf{I}_N - \rho \boldsymbol{w}_N)^{-1}\beta_j$. Local effects are on the diagonal of the $n \times n$ matrix $\boldsymbol{S}_n(\rho, \beta_j)$, and spillover effects are off-diagonal.

The series expansion $\boldsymbol{S}_n(\rho, \beta_j) = (I_n - \rho \boldsymbol{w}_n)^{-1}\beta_j = (\mathbf{I}_n + \rho \boldsymbol{w}_n + \rho^2 \boldsymbol{w}_n^2 + \dots)\beta_j$ decomposes partial effects into orders $0, 1, 2, \dots$. This shows β_j as a zero-order effect from observations themselves. A simple interpretation of β_j ignores the impact of immediate neighbours (first-order), neighbours' neighbours (second-order), and so on. When $\rho \neq 0$, β_j is not a valid local effect because of feedback effects—see, e.g., Lesage and Pace (2009, sect. 2.7.4).[14]

Partial effects from SAR models exhibit much heterogeneity, and one may be interested in summary measures. Lesage and Pace (2009) proposes summary measures, namely, the average local effect $LE_j(\rho, \beta_j) = \frac{1}{n}tr(\boldsymbol{S}_n(\rho, \beta_j))$, average total effect $TE_j(\rho, \beta_j) = \frac{1}{n}\mathbf{1}'_n \boldsymbol{S}_n(\rho, \beta_j)\mathbf{1}_n$, and average spillover effect $SpE_j(\rho, \beta_j) = TE_j(\rho, \beta_j) - LE_j(\rho, \beta_j)$. The average total effect, when $\boldsymbol{w}_n$ is row-normalised, has the simple expression $TE_j(\rho, \beta_j) = \frac{1}{n}\mathbf{1}'_n \boldsymbol{S}_n(\rho, \beta_j)\mathbf{1}_n = \beta_j/(1-\rho)$—see, e.g., Lesage and Pace (2009, sect. 2.7.1). For inference on partial effects, we follow Lesage and Pace (2009, sect. 2.7.3). The approach involves simulating the empirical

[13] This is the type of interpretation found in Teltemann and Windzio (2018), and Gaku and Tsyawo (2021) as the authors consider spatially-lagged outcome as a control.

[14] A feedback effect occurs when an impact from a country passes through its neighbour back to the country of origin.

Table 2 OLS results by Subject

	Reading (1)	Mathematics (2)	Science (3)
Autonomy	74.033***	95.152***	85.513***
	(18.323)	(19.867)	(16.951)
Private school share	– 14.992	– 22.181	– 20.865
	(17.393)	(26.557)	(20.714)
External monitoring	– 14.481	– 29.910	– 23.606
	(14.397)	(18.222)	(15.802)
No shortage of educ. materials	31.390	43.750	48.058
	(26.962)	(35.400)	(29.237)
Large shortage of educ. materials	– 124.707**	– 140.512**	– 122.863**
	(53.730)	(60.957)	(52.903)
GDP per capita	0.161	0.263*	0.147
	(0.118)	(0.153)	(0.131)
OECD membership	21.185***	16.861*	21.863***
	(7.128)	(9.013)	(7.696)
R^2	0.62	0.63	0.64

Total number of country-year observations is 168. Robust standard errors adjusted for clustering at the country level are in parentheses. All specifications control for year fixed effects and income categories

Significance levels: ***1%, **5%, *10%

distribution of the partial effects from the (asymptotically) normal distribution of parameter estimates—see also Fischer et al. (2009). Partial effects may not be symmetrically distributed even if ρ and β_j are normally distributed—see discussion in LeSage and Pace (2013, p. 1545). For this reason, we construct confidence intervals using percentiles of the simulated empirical distribution of partial effects.

4 Empirical results

This section presents our findings comparing results from the Ordinary Least Squares (OLS) and Spatial Autoregressive (SAR) models to determine the magnitude and relevance of spillover effects on student achievement. We study the local and spillover effects of school autonomy on reading, mathematics, and science achievement. We also provide results for different subdomains and types of autonomy such as teacher autonomy and budget autonomy. Finally, we conduct robustness analyses to verify the sensitivity of our results.

Table 3 SAR Results by Subject

	Reading	Mathematics	Science
	(1)	(2)	(3)
Panel A	Coefficients		
Autonomy	67.362***	72.432***	68.999**
	(13.971)	(14.159)	(12.928)
ρ	0.409**	0.685***	0.628***
	(0.176)	(0.162)	(0.156)
Panel B	Partial effects		
Local	68.280	76.293	71.822
	[41.3;96.7]	[49.6;113]	[47;101.7]
Spillover	45.697	153.346	113.837
	[4.7;192.9]	[40.9;1229.7]	[32;675.6]
Total	113.977	229.639	185.658
	[59.7;270.6]	[104.9;1323]	[92.4;759.3]
Panel C	Goodness of fit		
J-stat (df)	6.27 (8)	6.90 (8)	7.05 (8)
p-value	0.62	0.55	0.53
R^2	0.61	0.61	0.63

Notes: Total number of country-year observations is 168. Robust standard errors adjusted for clustering at the country level are in parentheses. The 95% confidence intervals for partial effects provided in brackets are constructed from 10,000 simulations—see Lesage and Pace (2009), and Fischer et al. (2009). All specifications control for year fixed effects, income category, private school share, external monitoring, shortage of instruction materials, and GDP per capita. *Significance levels*: *** 1%, ** 5%, * 10%

4.1 Ordinary least squares results

Table 2 shows the direct effect of school autonomy on reading, mathematics, and science achievement after controlling for school- and country-level covariates as well as time-fixed effects. The results suggest that a 10 percentage point increase in school autonomy would increase reading achievement by 16.8%, mathematics achievement by 18.2%, and science achievement by 17.9% of a standard deviation.[15] According to Wößmann (2016), a year of learning approximately translates into one-third of a standard deviation increase in PISA test scores. This rule of thumb suggests that a 10% increase in school autonomy increases reading test scores by half a year's worth of learning. Other estimates in Table 2 suggest that countries with better equipped schools (e.g., Japan and Singapore) have higher test scores in Reading, Mathematics and Science than countries where many schools report shortages

[15] The standard deviations of PISA test scores are given in Table 1.

Table 4 SAR Results by Subdomains for Reading

	School				Personnel	Budget
	Assessments	Textbooks	Content	Course choice		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	Coefficients					
Autonomy	70.211***	50.859***	37.611***	52.585***	17.863	48.996***
	(14.959)	(12.135)	(11.369)	(11.360)	(12.735)	(13.313)
ρ	0.374*	0.512***	0.483***	0.431***	0.502***	0.374**
	(0.204)	(0.174)	(0.165)	(0.151)	(0.166)	(0.180)
Panel B	Partial effects					
Local	70.991	52.051	38.374	53.397	18.263	49.541
	[42.5;101.3]	[28.6;77.6]	[16.3;61.3]	[31.3;76.4]	[-6.8;43.9]	[24;76]
Spillover	41.167	52.089	34.357	39.058	17.64	28.75
	[-1.9;215.1]	[10;256.1]	[6.5;135.7]	[7.7;137.6]	[-9.3;80.1]	[1.1;114.8]
Total	112.158	104.14	72.731	92.454	35.903	78.29
	[58.8;288.9]	[49.4;315.6]	[30.6;182.2]	[47.9;202.8]	[-16.4;113.4]	[37.2;172.6]
Panel C	Goodness of fit					
J-stat (df)	6.84 (8)	6.74 (8)	5.78 (8)	5.28 (8)	7.48 (8)	7.03 (8)
p-value	0.55	0.57	0.67	0.73	0.49	0.53
R^2	0.60	0.57	0.57	0.62	0.52	0.58

Notes: Dependent variable: PISA reading score. Total number of country-year observations is 168. Robust standard errors adjusted for clustering at the country level are in parentheses. The 95% confidence intervals provided in brackets are constructed from 10,000 simulations of partial effects. All specifications control for year fixed effects, income category, private school share, external monitoring, shortage of instruction materials, and GDP per capita. *Significance levels*: *** 1%, ** 5%, * 10%

in educational materials (e.g., Indonesia and Turkey). On average, OECD member countries perform better on PISA by about 20 points.

4.2 Local and spillover effects of autonomy on test scores

Table 3 presents the SAR results by subject area. Panel A reports estimates of the coefficient on Autonomy and the spatial coefficient ρ, Panel B reports partial effects viz. local, spillover, and total effects of autonomy with 95% confidence intervals in brackets, and Panel C provides the over-identifying restrictions test (J-statistic, its degrees of freedom, and p-value) and the R^2 coefficient. All specifications control for year fixed effects, school- and country-level covariates. The set of instruments used comprises the linearly independent columns of $[\mathbf{x}_n, \boldsymbol{w}_n\mathbf{x}_n]$.

Coefficients on Autonomy are large and positive across specifications, although slightly smaller than coefficients in Table 2. The direct effect of school autonomy on student achievement ranges from 67.36 points for Reading to 72.43 points for

Mathematics. Further, all three spatial coefficients (ρ) are significant at the 5% level and range from 0.41 for Reading to 0.69 for Mathematics.

Partial effects are computed based on the coefficient for Autonomy and the spatial coefficient ρ—see formulae in Sect. 3.2. The total effect for each specification is obtained by adding the local and spillover effects. The 95% confidence intervals provided in brackets indicate that all spillover and total effects are statistically significant at the 5% level. Large upper bounds in the asymmetric confidence intervals of local, spillover, and total effects are indicative of right-skewed distributions of partial effects.[16] The total effect in column (1) indicates that a 10 percentage point increase in school autonomy increases reading achievement by 25.9% of a standard deviation, mathematics achievement by 43.8% of a standard deviation, and science achievement by 38.9% of a standard deviation.[17] Spillover effects are large and consequential across specifications. About 40% of the total effect of school autonomy on reading test scores can be attributed to spillover effects. For Mathematics and Science, almost 2/3 of the total effect can be attributed to spillovers. Although the spillover effects for Mathematics and Science are large in magnitude, their confidence intervals also suggest some imprecision in their estimation. Relative to the local effects, the spillover effects especially for Mathematics and Science which are quite large in Table 4 are also less precisely estimated. For instance, both confidence intervals do not exclude spillover effects as low as 40 which are excluded by the respective confidence intervals of the local effects. One observes that the total effect of Autonomy on reading achievement using the SAR model is about 1.5 times the OLS estimate, and more than twice as large for Mathematics and Science. Hence, estimates based on linear regression underestimate the impact of school autonomy by ignoring education policy transfer processes and spillover effects of policies between countries such as those described by Breakspear (2012) and Parcerisa et al. (2020).

4.3 Local and spillover effects of autonomy by subdomains

There are different types of autonomy assessed in PISA. Our main results are based on school autonomy (also called academic-content autonomy) which is made up of four subdomains—designing assessments, choosing textbooks, course content, and course offerings. In addition, PISA provides data on budget autonomy (i.e., schools can decide how to allocate their budgets) as well as personnel autonomy (i.e., schools can decide whether to hire or fire a teacher).

Table 4 provides estimates and partial effects for school autonomy subdomains, personnel autonomy, and budget autonomy.[18] Autonomy coefficients are positive

[16] Partial effects are highly non-linear functions of the asymptotically normally distributed parameter estimates—see Sect. 3.2.

[17] E.g., for Reading, this value is computed as $\frac{113.977/10}{44.004} \times 100\% \approx 25\%$ where the total effect is 113.977 and the standard deviation of reading scores is 44.004 (see Table 1).

[18] The results in Table 4 is based on Reading. See the online appendix for results by subdomain on Mathematics and Science.

Table 5 Robustness Analyses for Reading

	Drop Missing (1)	OECD only (2)	Europe only (3)	Country-fixed effects (4)
Panel A	Coefficients			
Autonomy	69.352***	41.944***	40.878**	58.743***
	(14.488)	(10.910)	(14.056)	(11.905)
ρ	0.579***	0.012	0.626	0.190***
	(0.166)	(0.533)	(0.392)	(0.045)
Panel B	Partial effects			
Local	71.633	41.944	42.697	58.894
	[44.1;103.1]	[22.2;65.2]	[18.2;87.5]	[36;82.3]
Spillover	93.285	0.494	66.52	13.629
	[23.6;501.7]	[-25.5;231.1]	[-7.8;1199.7]	[5.5;25.6]
Total	164.918	42.439	109.217	72.523
	[83.3;585.2]	[19.7;278.5]	[29.8;1276]	[43.8;104.7]
Panel C	Goodness of fit			
J-stat (df)	5.60 (8)	11.42 (7)	10.28 (7)	354.29 (8)
p-value	0.69	0.12	0.17	0.00
N (Countries)	56	35	31	56
R^2	0.62	0.58	0.45	0.96

Notes: Dependent variable: PISA reading score. Robust standard errors adjusted for clustering at the country level are in parentheses. The 95% confidence intervals provided in brackets are constructed from 10,000 simulations of partial effects. Standard errors in specification (4) are only heteroskedasticity-robust because the cluster-robust covariance matrix is not invertible. All specifications control for year fixed effects, income categories, private school share, external monitoring, shortage of instruction materials, and GDP per capita. *Significance levels*: *** 1%, ** 5%, * 10%

and significant save personnel autonomy (specification 5), though overall smaller than the school autonomy coefficient in Table 3. Spatial coefficients are positive and significant across specifications and range from 0.37 for budget and assessment autonomy to 0.51 for textbook autonomy.

The table indicates that all subdomains of autonomy are important in explaining the overall effect of school autonomy on reading achievement. Partial effects are shown in Panel B and indicate that spillovers are large and consequential. Spillover effects are significant at the 5% level for specifications 2–4 and at the 10% level for specification 1. The total effects range from 72.73 points for autonomy in choosing course content to 112.16 points for autonomy in designing assessments. These effects are smaller than the total effect reported for school autonomy but are still of meaningful magnitude.

The last two specifications in the table provide estimates for personnel and budget autonomy. We find no effect for personnel autonomy in column (5) and a positive but smaller effect for budget autonomy. This suggests that autonomy in designing the academic content delivered in schools is a more important driver of student test performance than autonomy in selecting teaching personnel or allocating budgets.

4.4 Robustness analyses

To verify the sensitivity of our results to plausible variations in specifications, we provide robustness analyses in Table 5 for Reading.[19] In column (1), Israel and Sweden are excluded from the analysis because both countries have missing values which were imputed and specification (1) of Table 3 is re-run. There are some changes in the estimates for autonomy and ρ, with the spillover still being large and consequential. Also, the specification is not rejected by the data. This confirms that the imputation of private school shares does not qualitatively alter the results.

Specifications (2) and (3) consider sub-samples (OECD and European countries) of the dataset. Both specifications have smaller and less significant estimates for autonomy relative to Table 3. The spatial coefficient for the OECD subsample, where several member countries are geographically dispersed, is close to zero and the spillover effect is inconsequential. For the European sub-sample where countries are in close geographical proximity, the spatial coefficient is large but non-significant and the 95% confidence interval of the spillover effect does not exclude zero.[20]

Column (4) considers the addition of country fixed effects in the model. Other researchers working with PISA panel data have used both time and country fixed effects to reduce bias from time-invariant unobserved country characteristics—e.g., Hanushek et al. (2013). This model does not seem to be supported by the data as suggested by the J-test. The coefficients and partial effects are nonetheless meaningful and not at considerable variance with our main conclusions in Table 3 for example.

5 Conclusion

Works that examine the impact of education policy at the country-level largely ignore spillover effects that may arise when countries influence each other's education policy decisions and outcomes. Spillover effects are, however, important from a welfare perspective when evaluating the overall impact of an education policy across countries. This paper contributes to the literature by examining spillover effects of an education policy, viz. school autonomy, on student achievement using international student assessment data.

[19] Robustness analyses for Mathematics and Science give similar results and are thus not provided in the main text—see the online appendix.

[20] This is likely attributable to the smaller sample size.

Geographic proximity is tied to other conditional factors such as economic ties between countries, common language and cultural similarities. Our findings confirm spatial dependence in student achievement across countries linked to geographic proximity between countries. This result supports Obinger et al. (2013)'s argument that the exchange of information and competition between neighbouring countries produce spillover effects. We find positive and significant average local and spillover effects of school autonomy on student reading, mathematics, and science performance. Our preferred specification indicates that a 10 percentage point increase in school autonomy increases reading achievement by 25.9% of a standard deviation. About 40 percent of the total policy effect can be attributed to spillover effects.

Our analysis of subdomains of school autonomy confirms that all four subdomains contribute to School Autonomy's overall effect on reading achievement. We find that autonomy in designing assessments produces the largest local and spillover effects, and autonomy in choosing course content produces the smallest effect on reading achievement. Subdomain-specific spillover effects are significant for textbooks, course content, and course choice. Our main findings do not hold for personnel autonomy, but we find smaller positive effects for budget autonomy. Hanushek et al. (2013) also find stronger effects associated with school autonomy compared to personnel and budget autonomy. Moreover, Teltemann and Windzio (2018) argue that dimensions of autonomy directly related to learning processes, such as choosing course content or textbooks, are more predictive of student achievement.

Spatial econometric models allow researchers to determine the existence of spillover effects in the data. They do not identify specific causal mechanisms that drive spillover effects. Consequently, we cannot use these models to determine which causal mechanisms (e.g., policy borrowing or international standardisation) explain policy spillovers. Other researchers using more qualitative approaches such as document analyses have identified and linked various PISA education reforms to specific policy diffusion mechanisms—see, e.g., Parcerisa et al. (2020). Our approach, however, enables us to identify relevant forms of proximity that foster policy spillovers between countries, such as geographic proximity. Aggregation of data to the country level masks a number of features that are not explored in this paper. We are unable to study spillover effects at the subnational or regional level in this paper as not all countries in the PISA study provide region-specific identifiers to study intra-country mechanisms. With more waves of PISA data available, it may be possible in the future to construct larger panels for individual countries and examine such regional effects. Moreover, cross-country research often uncovers only broad patterns of association but precludes a detailed exploration of local policy decisions (Wößmann 2007). Exploring inter- and intra-country spillovers of education policy using student- and school-level data is another interesting extension worth exploring in future work.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12076-022-00310-y.

Funding The authors did not receive any funding for the execution of this project.

Availability of data and material Data used in this project are publicly available at https://www.oecd.org/pisa/data/

Code availability All estimation was conducted using the open source software R. All codes are available from the authors.

Declarations

Conflict of interest The authors have no conflicts of interest or competing interests.

References

Ammermueller, A.: Institutional features of schooling systems and educational inequality: cross-country evidence from PIRLS and PISA. German Econ. Rev. **14**(2), 190–213 (2013)
Anselin, L., Gallo, J.L., Jayet, H.: Spatial Panel Econometrics. The Econometrics of Panel Data. Springer, pp. 625–660 (2008)
Baird, J.-A.: et al. On the supranational spell of PISA in policy. Educ. Res. **58.2**, 121–138 (2016)
Bramoullé, Y., Djebbari, H., Fortin, B.: Identification of peer effects through social networks. J. Econometr. **150**(1), 41–55 (2009)
Bramoullé, Y., Djebbari, H., Fortin, B.: Peer Effects in Networks: A Survey" (2019)
Breakspear, S.: The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance. OECD (2012)
Cantley, I.: PISA and policy-borrowing: a philosophical perspective on their interplay in mathematics education. Educ. Philos. Theory **51**(12), 1200–1215 (2019)
Coghlan, M., Desurmont, A.: School Autonomy in Europe Policies and Measures. Brussels (2007)
Fischer, M.M., Bartkowska, M., Riedl, A., Sardadvar, S., Kunnert, A.: The impact of human capital on regional labor productivity in Europe. Lett. Spatial Resour. Sci, **2.2-3**, 97–108 (2009)
Forestier, K., Crossley, M.: International education policy transfer-borrowing both ways: the Hong Kong and England experience. Compare J. Comp. Int. Educ. **45.5**, 664–685 (2015)
Fuchs, T., Wößmann, L.: What accounts for international differences in student performance? A re-examination using PISA data. The Economics of Education and Training. Springer, pp. 209–240 (2008)
Gaku, S., Tsyawo, Emmanuel, S.: Neighbourhood effects and the incidence of child labour. Lett. Spatial Resour. Sci. **14.3**, pp. 247–259 (2021)
Gill, T., Benton, T.: Investigating the Relationship Between Aspects of Countries' Assessment Systems and Achievement on the Programme for International Student As sessment (PISA) Tests. Cambridge Assessment, Cambridge (2013)
Gray, J., Galton, M., Colleen, M.: Wellbeing and the Young Adolescent. Cambridge Scholars Publishing, The supportive school (2011)
Grek, S.: Governing by numbers: the PISA 'effect' in Europe. J. Educ. Policy **24**(1), 23–37 (2009)
Gulson, K.N., Symes, C.: Spatial Theories of Education: Policy and Geography Matters, vol. 9. Routledge Research in Education. Routledge, New York (2007)
Hanushek, E.A., Wößmann, L.: Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. Econ. J. **116.510**, C63–C76 (2006)
Hanushek, E.A., Link, S., Wößmann, L.: Does school autonomy make sense everywhere? Panel estimates from PISA. J. Dev. Econ. **104**, 212–232 (2013)
Holzinger, K., Knill, C.: Causes and conditions of cross-national policy convergence. J. Eur. Publ. Policy **12**(5), 775–796 (2005)
Kelejian, H.H., Prucha, I.R.: A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. J. Real Estate Finance Econ. **17.1**, 99–121 (1998)
König, M.D., Liu, X., Zenou, Y.: R &D networks: theory, empirics, and policy implications. Rev. Econ. Stat. **101.3**, 476–491 (2019)

Larsen, M.A., Beech, J.: Spatial theorizing in comparative and international education research. Comp. Educ. Rev. **58**(2), 191–214 (2014)
Lee, L.-F.: GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. J. Econometr. **137**(2), 489–514 (2007)
LeSage, J.P., Pace, R.K.: Spatial Econometric Models. Handbook of Applied Spatial Analysis. Springer, pp. 355–376 (2010)
LeSage, J.P., Pace, R.K.: Interpreting Spatial Econometric Models. Handbook of Regional Science. Springer, pp. 1535–1552 (2013)
LeSage, J.P., Pace, R.K.: Introduction to Spatial Econometrics. CRC Press (2009)
Lin, X.: Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. J. Law Econ. **28**(4), 825–860 (2010)
Manski, C.F.: Identification of endogenous social effects: the reflection problem. Rev. Econ. Stud. **60**(3), 531–542 (1993)
Maslowski, R., Scheerens, J., Luyten, H.: The effect of school autonomy and school internal decentralization on students' reading literacy. Sch. Eff. Sch. Improv. **18**(3), 303–334 (2007)
Mayer, T., Soledad, Z.: Notes on CEPII's Distances Measures: The GeoDist Database. SSRN Electr. J. (2011)
Meseguer, C., Gilardi, F.: What is new in the study of policy diffusion? Rev. Int. Polit. Econ. **16**(3), 527–543 (2009)
Obinger, H., Schmitt, C., Starke, P.: Policy diffusion and policy transfer in comparative welfare state research. Social Policy Admin. **47**(1), 111–129 (2013)
OECD.: PISA 2015 Results (vol. I): Excellence and Equity in Education. OECD Publishing (2016)
Parcerisa, L., Clara, F., Antoni, V.: Understanding the PISA influence on national education policies: a focus on policy transfer mechanisms. International perspectives on school settings, education policy and digital strategies. A transatlantic discourse in education research, pp. 185–198 (2020)
Phillips, D., Ochs, K.: Processes of policy borrowing in education: some explanatory and analytical devices. Comp. Educ. **39**(4), 451–461 (2003)
Plümper, T., Neumayer, E.: Model specification in the analysis of spatial dependence. Eur. J. Polit. Res. **49**(3), 418–442 (2010)
Ringarp, J., Rothland, M.: Is the grass always greener? The effect of the PISA results on education debates in Sweden and Germany. Eur. Educ. Res. J. **9**(3), 422–430 (2010)
Sacerdote, B.: Peer effects in education: How might they work, how big are they and how much do we know thus far? Handbook of the Economics of Education. Vol. 3. Elsevier, pp. 249–277 (2011)
Schneeweis, N., Winter-Ebmer, R.: Peer effects in Austrian schools. Emp. Econ. **32**(2–3), 387–409 (2007)
Sellar, S., Thompson, G., David, R.: Taking the Measure of PISA and International Testing. Brush Education, The Global Education Race (2017)
Teltemann, J., Windzio, M.: The impact of marketisation and spatial proximity on reading performance: international results from PISA 2012. Compare J. Comp. Int. Educ. **49.5**, 777–794 (2018)
Waldow, F.: Projecting images of the 'good' and the 'bad school': Top scorers in educational large-scale assessments as reference societies. Compare J. Comp. Int. Educ. **47.5**, pp. 647–664 (2017)
West, M.R., Woessmann, L.: 'Every Catholic Child in a Catholic School': historical resistance to state schooling, contemporary private competition and student achievement across countries. Econ. J. **120.546**, F229–F255 (2010)
Wößmann, L.: International evidence on school competition, autonomy, and accountability: a review. Peabody J. Educ. **82**(2–3), 473–497 (2007)
Wößmann, L.: The importance of school systems: evidence from international differences in student achievement. J. Econ. Perspect. **30**(3), 3–32 (2016)