POVERTY, INEQUALITIES AND EDUCATIONAL OUTCOMES: A TERRITORIAL ANALYSIS IN THE AGE OF TRANSITIONS[1]

Simona Cafieri, Gianmarco Borrata, Manuela Barba, Paola Bianco

Abstract. This paper provides a multidimensional spatial analysis of educational inequality in Italy by integrating regional data (2019–2023) with K-means clustering and Geographically Weighted Regression. The results show that structural poverty, digital exclusion, youth vulnerability, and parental educational background produce markedly uneven literacy and numeracy outcomes, with the magnitude and direction of these effects varying across space. By examining how demographic, digital, and ecological transitions interact with existing vulnerabilities, the study identifies emerging territorial pressures that intensify educational disparities. Overall, the findings offer updated evidence to support more targeted, place-based strategies aimed at reducing persistent and evolving forms of educational disadvantage.

1. Introduction

In Italy, territorial disparities remain among the most persistent forms of social inequality. Education—one of the main channels of upward mobility—is strongly shaped by structural and territorial imbalances, as clearly revealed during the COVID-19 pandemic. The transition to distance learning exposed how unequal access to infrastructure and resources constrained educational opportunities in disadvantaged areas, further widening existing divides *(Borgonovi and Ferrara, 2023).*

While prior research has documented regional educational inequalities, limited attention has been devoted to how structural vulnerabilities and transition-related processes jointly produce spatially heterogeneous effects. This study addresses this gap by analysing how demographic, digital, and ecological transitions interact with multidimensional forms of disadvantage to shape literacy and numeracy outcomes *(Cantalini et al., 2025).*

The analysis is conducted at the regional level, the most appropriate scale for integrating harmonised socioeconomic, digital, and educational indicators and for capturing the cumulative effects associated with long-term structural transitions.

[1] This study represents the result of a collaborative effort between the authors.

While the abstract emphasises the need to move beyond the traditional North–South divide, this should not be interpreted as an attempt to analyse sub-regional patterns such as centre–periphery or urban–rural gradients, which cannot be adequately captured with available data. Rather, the paper advances beyond the binary macro-regional classification by adopting a multidimensional and spatially sensitive regional perspective that uncovers differentiated mechanisms within and across regions.

Although territorial inequalities in education represent a well-established field of research in Italy, existing analyses have predominantly relied on macro-regional comparisons or global regression models that implicitly assume spatial stationarity. Such approaches tend to overlook how the determinants of educational disadvantage vary across space and how structural vulnerabilities interact with transition-related pressures in differentiated ways. This study advances the literature by integrating a multidimensional framework of territorial vulnerability with spatially explicit modelling at the regional scale. By combining cluster analysis with Geographically Weighted Regression, the paper reveals mechanisms that remain hidden in conventional models and shows that socioeconomic, digital and demographic factors do not exert uniform effects across Italian regions. In doing so, it offers updated evidence on the geography of educational disadvantage in the post-pandemic and transition-oriented context and provides a more nuanced understanding of how cumulative vulnerabilities shape regional outcomes. This perspective contributes to bridging the gap between descriptive territorial disparities and the need for analytically grounded, place-sensitive policy insights.

2. Background

The understanding of poverty has evolved from a narrow focus on income deprivation to a multidimensional perspective. Contemporary frameworks emphasise the interplay of cognitive, social, and cultural dimensions in shaping educational trajectories *(Spicker et al., 2008).* In this view, educational poverty refers to limited access to quality learning environments, essential skills, and opportunities for cognitive development. Educational disadvantage is not solely rooted in individual or family conditions but is also shaped by contextual and territorial factors *(OECD, 201*2). In Italy, persistent spatial inequalities have produced regional clusters of deprivation where economic, social, and institutional vulnerabilities intersect and reinforce one another *(Ballarino, 2009).*

Despite extensive research on regional disparities in Italy, existing studies rarely adopt a multidimensional and transition-oriented perspective, nor do they examine how the digital, demographic, and ecological transitions translate into spatially differentiated educational outcomes. This leaves a gap in understanding the mechanisms through which structural territorial vulnerabilities shape learning inequalities—a gap that the present study addresses by integrating these transitions into a spatial analytical framework.

Recent literature highlights the need to interpret educational inequalities in light of three major transitions currently reshaping territorial dynamics: digital, demographic, and ecological *(Parsons et al., 2024; Rosário and Disa, 2022).*

• Digital transition: limited broadband coverage, low device availability, and weak digital skills constrain access to technology-enhanced learning and exacerbate existing gaps.

• Demographic transition: population ageing and shrinking youth cohorts reduce school network density and continuity of provision *(Muttarak and Lutz, 2014)*, particularly in shrinking regions.

• Ecological transition: exposure to environmental risks—such as heatwaves or hydrogeological instability—affects the resilience and quality of school infrastructure and disrupts learning conditions.

In this perspective, each transition is expected to generate territorially differentiated impacts: digital constraints amplify learning gaps most acutely in low-connectivity regions; demographic decline heightens educational vulnerability in ageing and shrinking areas where school provision becomes fragile; and ecological pressures disproportionately affect territories exposed to environmental risks that compromise the continuity and quality of learning conditions. Taken together, these transitions do not simply coexist but operate as structural amplifiers of territorial inequality, reinforcing context-specific mechanisms that shape geographically differentiated educational outcomes.

By integrating the three structural transitions into a spatially sensitive analytical framework and combining cluster analysis, OLS, and GWR, this study moves beyond traditional accounts of the North–South divide and provides a novel multidimensional reading of educational inequality in Italy. This approach not only identifies territorial disparities but explains the mechanisms through which digital, demographic, and ecological vulnerabilities translate into regionally differentiated educational outcomes—an aspect largely overlooked in previous research.

3. Data and methodology

The empirical analysis integrates data from a broad range of official statistical sources for the period 2019–2023. We combined EU-SILC *(Istat, 2022–2023),* LFS *(Istat, 2023),* INVALSI and PNRR information to capture socioeconomic, educational and policy dimensions across Italian regions. The regional scale allows the integration of harmonised indicators and provides sufficient territorial variation to analyse spatial heterogeneity through GWR, even though it does not permit an explicit investigation of urban–rural or centre–periphery gradients.

The methodological strategy comprises two steps. First, we conduct a descriptive territorial analysis and apply K-means clustering to classify regions according to structural determinants, using INVALSI proficiency distributions and indicators of socioeconomic and digital vulnerability. The cluster typologies serve as a structural lens to contextualise the subsequent regression and GWR results, helping to interpret how different configurations of vulnerability shape spatially varying relationships between socioeconomic factors and educational outcomes.

In the second step, we estimate the relationship between educational outcomes and socioeconomic conditions using multiple linear regression models. Since these models do not account for spatial dependence, we adopt GWR , following recent contributions *(Sacco and Falzetti, 2021).* GWR allows the strength and direction of associations to vary across space, highlighting mechanisms that remain hidden in global models.

This combined exploratory–explanatory approach enables both the identification of structural regional profiles and the analysis of spatially heterogeneous relationships between vulnerabilities and educational outcomes. “Structural determinants” refer here to persistent socioeconomic and educational characteristics—such as youth vulnerability, digital skills, household income and parental education—selected in line with established literature on territorial inequalities.

3.1. Cluster Analysis of Regional Education Performance

To classify regional heterogeneity, we applied K-means clustering, using the regional distributions of student performance in INVALSI literacy and numeracy assessments for 2019 and 2023, grouped into five ordered proficiency levels across primary and secondary school grades. The elbow method indicated a three-cluster solution, supported by empirical interpretability and consistency with well-documented territorial disparities.

The clustering relies on the following INVALSI indicators (regional level) for years: 2019, 2023:

1) Literacy proficiency distributions (5 ordered levels: Grades 2 and 5 primary; Grades 3 and 5 secondary);
2) Numeracy proficiency distributions (5 ordered levels: Grades 2 and 5 primary; Grades 3 and 5 secondary).

The clustering captures structural dimensions that the literature consistently associates with territorial inequalities, including socioeconomic fragility, digital capital, household resources and parental education. These dimensions shape cumulative learning opportunities and help identify stable regional profiles of disadvantage. Rather than explaining educational outcomes directly, the cluster typology provides a theoretically grounded framework for interpreting the subsequent regression and GWR analyses. In particular, it offers a structural lens through which to understand how different configurations of vulnerability relate to the spatially varying relationships revealed by the GWR.

3.2. Cluster Composition and Interpretation

The cluster analysis identifies three distinct regional profiles, reflecting well-established territorial divides in Italy (Figures 1, and 2).

The first group includes regions characterised by stronger socioeconomic and educational endowments; the second captures intermediate and internally heterogeneous contexts; and the third comprises structurally disadvantaged regions with persistently lower proficiency levels. These profiles provide the structural backdrop for interpreting the regression and GWR results, highlighting how different combinations of vulnerabilities underpin the spatially varying mechanisms identified in the subsequent analysis.

Figure 1 – *Cluster of Italian regions for 2019: inadequate numerical and literacy competence.*

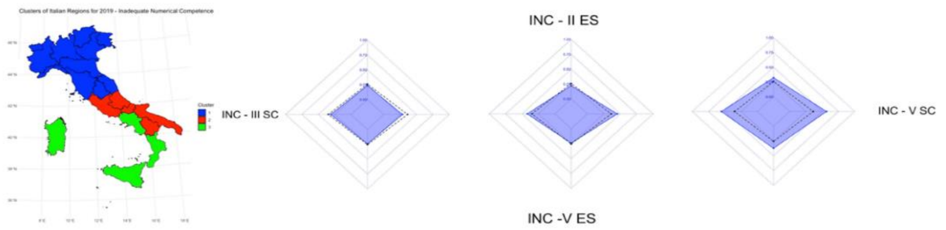
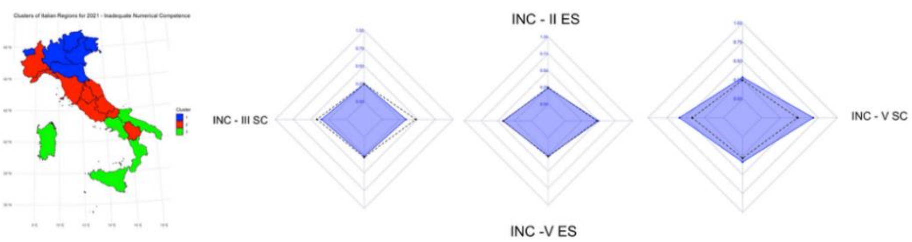


Figure 2 – *Cluster of Italian regions for 2021: inadequate numerical and literacy competence.*



3.3. Cluster Results to Territorial Patterns of Educational Disadvantage

The three clusters identified in Section 3.2 reveal distinct territorial patterns of educational disadvantage, distinguishing regions marked by persistent structural vulnerabilities from those characterised by more favourable socioeconomic and educational conditions. These profiles reflect long-standing territorial divides and illustrate how cumulative disadvantages shape regional differences in literacy and numeracy outcomes.

By linking the cluster results to the regional socioeconomic landscape, it becomes possible to interpret the spatial variation observed in the regression and GWR analyses. The clusters highlight how different territorial endowments condition the strength and direction of key determinants, thereby providing a coherent framework for understanding the spatially heterogeneous mechanisms identified through GWR.

4. Determinants of education deprivation: Geographically Weighted Regression

To examine how structural vulnerabilities relate to literacy and numeracy outcomes across space, we complement the cluster analysis with a set of Geographically Weighted Regression (GWR) models. Table 1 summarises the socioeconomic indicators used in the empirical analysis. The four dependent variables (Y1–Y4) reflect inadequate literacy and numeracy competences at the end of primary and upper secondary school.

The explanatory variables capture key dimensions of territorial vulnerability: digital proficiency (X1); youth disadvantage, including NEET rates (X2) and early school leaving (X3); household economic resources (X4); and school service provision (X5). Additional indicators (X6–X9) represent parental educational composition.

GWR models are estimated for each dependent variable to assess how the strength and direction of these relationships vary across regions.

GWR is used to capture spatial heterogeneity in the relationship between literacy and numeracy outcomes and their socioeconomic determinants, allowing coefficients to vary across regions *(Brunsdon et al., 1998; Sacco et al., 2021).* This approach complements the global regression by identifying territorial differences in the magnitude and direction of key associations.The spatial analyses have been performed using the R package spdep[2] e GWmodel[3].

[2] https://cran.r-project.org/web/packages/spdep/index.html
[3] https://cran.r-project.org/web/packages/GWmodel/index.html

Table 1 – *Socioeconomic Indicators used in the analysis.*

Variable	Indicator	Definition
Y1	Low literacy – Grade 5 (Primary)	Percentage of students in 5th grade of primary school with inadequate literacy skills
Y2	Low literacy – Grade 5 (Upper Secondary)	Percentage of students in final year of upper secondary school with inadequate literacy skills
Y3	Low numeracy – Grade 5 (Primary)	Percentage of students in 5th grade of primary school with inadequate numeracy skills
Y4	Low numeracy – Grade 5 (Upper Secondary)	Percentage of students in final year of upper secondary school with inadequate numeracy skills
X1	Digital skills	Share of individuals with at least basic digital skills
X2	NEET	Percentage of youth aged 15–24 not in employment, education or training
X3	Early leavers from education and training	Percentage of 18–24 year-olds with at most lower secondary education who are not in further education or training
X4	Average household income	Mean income per household
X5	Schools offering basic services	Percentage of schools providing basic infrastructure (internet, labs, accessibility, etc.)
X6	Husbands: upper secondary – Wives: upper secondary	Share of couples where both spouses have upper secondary education
X7	Husbands: tertiary – Wives: upper secondary	Share of couples where husbands have tertiary and wives upper secondary education
X8	Husbands: tertiary – Wives: tertiary	Share of couples where both spouses have tertiary education
X9	Husbands: upper secondary – Wives: tertiary	Share of couples where husbands have upper secondary and wives tertiary education

4.1. Results of Geographically Weighted Regression

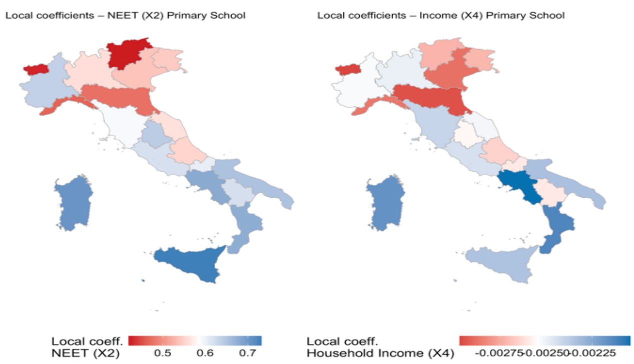
Table 2 presents the GWR estimates for inadequate literacy competence in Grade 5 of primary school. To support the interpretation of spatial heterogeneity, we include maps illustrating the geographical variation of selected coefficients. These visualisations show how the effects of specific determinants differ across regions, highlighting territorial patterns concealed in global models.

Household income (X4) is the only predictor consistently significant across all regions, showing a uniformly negative association with inadequate literacy outcomes. Early school leaving (X3) and the provision of basic school services (X5) also display negative effects, though their significance varies territorially (65% and 55% of regions, respectively). Digital skills (X1) and NEET rates (X2) exert more localised influences, emerging as significant only in specific areas, while parental education variables (X6–X9) show generally weak associations.

Table 2 – *Results of the geographically weighted regression (GWR) model for Inadequate Literacy Competence - Grade 5 (Primary School, 2021).*

	Coefficient Range						
	Min	1st Q	Mean	Median	3rd Q	Max	% Sig. Coef.
Intercept	-24.72	-22.31	-20.29	-19.47	-18.19	-16.98	85
X1	0.72	0.79	0.84	0.85	0.91	1.01	45
X2	0.42	0.55	0.60	0.61	0.66	0.74	30
X3	-0.34	-0.29	-0.26	-0.26	-0.21	-0.15	65
X4	-0.003	-0.003	-0.003	-0.003	-0.002	-0.002	100
X5	-0.52	-0.46	-0.45	-0.44	-0.40	-0.34	55
X6	-0.38	-0.30	-0.28	-0.27	-0.21	-0.18	20
X7	-0.15	-0.08	0.02	0.01	0.09	0.19	10
X8	-1.80	-1.72	-1.70	-1.68	-1.64	-1.60	0
X9	0.78	0.85	0.90	0.92	0.95	1.02	25
R^2	0.85	0.87	0.88	0.88	0.89	0.91	
	R^2 *Global* = 0.88				R^2 *OLS* = 0.81		

Figure 3 – *Local GWR coefficients for NEET (X2) and Household Income (X4) – Literacy Grade 5 (Primary School, 2021).*



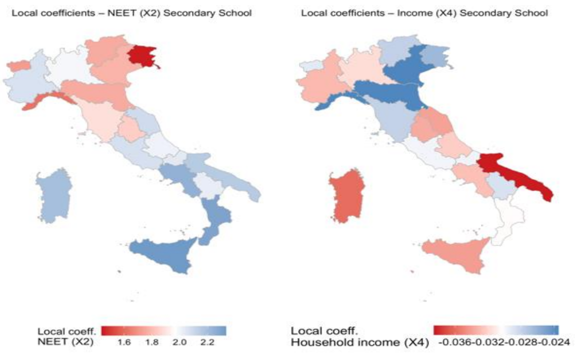
To assess robustness, we performed diagnostic checks: the GWR model outperformed the global OLS specification (lower AICc), Moran's I on residuals confirmed the absence of remaining spatial autocorrelation, and alternative bandwidth selection procedures yielded comparable coefficient surfaces, supporting the stability of the spatial patterns identified.

Table 3 presents the results of the GWR model estimating the determinants of inadequate literacy competence among students in the Grade 5 of secondary school.

Table 3 – *Results of the geographically weighted regression (GWR) model for Inadequate Literacy Competence - Grade 5 (Secondary School, 2021).*

	Coefficient Range						
	Min	1st Q	Mean	Median	3rd Q	Max	% Sig. Coef.
Intercept	65.2	70.8	74.5	74.1	78	82.6	80
X1	-1.85	-1.52	-1.39	-1.37	-1.21	-0.95	90
X2	1.42	1.70	1.94	1.95	2.15	2.36	100
X3	1.12	1.39	1.59	1.58	1.73	2.01	65
X4	-0.04	-0.036	-0.031	-0.030	-0.027	-0.022	100
X5	-1.54	-1.34	-1.22	-1.21	-1.12	-0.98	70
X6	-1.12	-0.95	-0.87	-0.86	-0.75	-0.63	35
X7	-2.35	-2.05	-1.88	-1.87	-1.69	-1.45	40
X8	-1.95	-1.66	-1.52	-1.51	-1.39	-1.20	50
X9	-1.10	-0.87	-0.74	-0.75	-0.62	-0.45	25
R^2	0.88	0.90	0.92	0.92	0.93	0.95	
	$R^2\ Global\ =\ 0.93$					$R^2\ OLS\ =\ 0.92$	

Figure 4 – *Local GWR coefficients for NEET (X2) and Household Income (X4) – Literacy Grade 5 (Secondary School, 2021).*



Compared to the global OLS estimates, the GWR results reveal marked spatial variation in several relationships. Predictors that appear weak in the global specification, such as digital skills and school service provision, display pronounced effects in some regions, whereas household income emerges as a consistently influential factor. These patterns show that global coefficients mask relevant territorial differences, underscoring the value of adopting a spatially explicit modelling strategy.

5. Education Policy and Public Investment

The discussion Public investment in education, including PNRR resources, has the potential to mitigate the territorial vulnerabilities highlighted by the empirical analysis. However, the current allocation of funds does not consistently reflect the spatial patterns identified by the GWR model. Regions where inadequate literacy outcomes are more strongly driven by digital constraints, socioeconomic disadvantage, or demographic decline do not always correspond to those receiving proportionally greater investment in digital infrastructure, school network consolidation, or ecological adaptation.

The findings suggest that PNRR measures could be more effective if aligned with the locally varying determinants of educational disadvantage. A more spatially targeted approach—sensitive to the mechanisms revealed by cluster analysis and GWR—would enhance the capacity of public investment to reduce structural inequalities rather than address needs in a uniform manner.

6. Discussions and conclusions

This study contributes to the analysis of territorial inequalities in Italy by integrating a multidimensional framework of vulnerability with a geographically weighted modelling strategy. By linking structural poverty, digital exclusion and youth disadvantage to spatially varying educational outcomes, it provides new evidence on mechanisms that remain invisible in conventional regional or macro-regional analyses. This spatial lens shows that inequalities are driven by locally differentiated processes rather than by uniform national patterns.

The empirical results confirm that the determinants of educational disadvantage vary substantially across space. These findings advance existing research by demonstrating pronounced spatial non-stationarity—an aspect largely overlooked in previous studies—and by clarifying how multidimensional vulnerabilities interact with demographic, digital and ecological transitions at the regional level.

The cluster analysis strengthens this interpretation by identifying distinct territorial profiles. Regions with strong socio-institutional endowments exhibit more resilient educational outcomes, whereas structurally fragile territories reveal compounded disadvantages that mirror the spatial patterns captured by the GWR. Rather than serving as explanatory variables, the clusters provide a structural lens through which the localised regression results can be interpreted, clarifying how different configurations of vulnerabilities shape spatially differentiated educational outcomes.

The policy implications follow directly from these spatially heterogeneous mechanisms. Regions in Cluster 1, where vulnerabilities are limited, may prioritise innovation and the consolidation of effective practices. Cluster 2 regions require targeted interventions to address internal fragmentation and mixed disadvantage profiles. Cluster 3 regions—marked by persistent deprivation and low performance—need long-term, integrated strategies combining income support, community-based programmes and investments in educational infrastructure. The GWR results underscore that policies designed without consideration of geographical variation risk misallocating resources, as identical measures may have markedly different effects depending on local conditions.

Overall, the study demonstrates that the geography of educational disadvantage is shaped by non-stationary mechanisms that can only be detected through spatially explicit models. Despite the constraints of regional data, the analysis provides updated insights into how structural vulnerabilities interact with demographic, digital and ecological transitions, offering a timely contribution to the post-pandemic and transition-oriented policy agenda.

Finally, the analysis points to several avenues for future research. Structural causal models and directed acyclic graphs (DAGs) could strengthen variable selection and reduce risks of collider bias. Additional factors—such as teacher quality, governance structures or school-level practices—could not be included due to data limitations but may play a substantive role in shaping territorial outcomes. Extending the framework to micro-data or panel designs would allow for a more precise identification of dynamic mechanisms. Evaluating the territorial effects of PNRR-funded interventions and examining how demographic, digital and ecological transitions reshape educational opportunities will be essential for informing future place-sensitive policies.

References

BALLARINO G., BERNARDI F., REQUENA M., SCHADEE H. 2009. Persistent inequalities? Expansion of education and class inequality in Italy and Spain. *European Sociological Review*, Vol. 25, No. 1, pp.123-138.

BORGONOVI F., FERRARA A. 2023. COVID-19 and inequalities in Educational achievement in Italy. *Research in social stratification and mobility,* Vol. 83:100760.

BRUNSDON C., FOTHERINGHAM S., CHARLTON M. 1998. Geographically weighted regression. *Journal of the Royal Statistical Society*, pp. 431-443.

CANTALINI S., PANICHELLA N., PIETROLUCCI A., TRIVENTI M. 2025 Regional disparities, geographical marginality, and educational pathways: A study on upper secondary education in Italy. *Social Inclusion*, Vol. 13.

INVALSI 2024. *Rilevazioni Nazionali degli Apprendimenti*. Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e Formazione.

ISTAT 2022-2023 EU-SILC – Statistics on Income and Living Conditions. Roma: Istituto Nazionale di Statistica.

ISTAT 2023. Labour Force Survey – LFS. Roma: Istituto Nazionale di Statistica.

ITALIADOMANI 2025. *PNRR Open Data Portal*. https://www.italiadomani.gov.it

MEF 2023. Ripartizione territoriale delle risorse del PNRR.

MUTTARAK R., LUTZ W. 2014. Is education a key to reducing vulnerability to natural disasters and hence climate change? *Ecology and society*, Vol. 19, No. 1.

OECD 2012. *Equity and quality in education: Supporting disadvantaged students.* Paris: Organisation for Economic Co-operation and Development.

OPENCOESIONE 2025. *Portale Open Data – Politiche di Coesione*. https://opencoesione.gov.it

ROSARIO A. T., DISA J. C. 2022. Sustainability and the digital transition: A literature review. *Sustainability,* Vol. 14, No. 7, pp. 4072.

SACCO C., FALZETTI P. 2021. Spatial variations of school-level determinants of reading achievement. *Large-scale Assessments in Education*, Vol. 9, No.1, No.12.

PARSONS E. S., JOWELL A., VEIDIS E., BARRY M., ISRANI S. T. 2024. Climate change and inequality. *Pediatric research*, pp. 1-8.

SPICKER P., LEGUIZAMON S. Á., GORDON D. (Eds.). 2008. *Poverty: an international glossary.* London: Bloomsbury Publishing.

Simona CAFIERI, Istat, simona.cafieri@istat.it
Gianmarco BORRATA, Università Federico II, gianmarco.borrata@unina.it
Manuela BARBA, Istat, manuela.barba@istat.it
Paola BIANCO, Istat, paola.bianco@istat.it