CELEBRATING 50

Public Sector Economics

2/2026

Vol. 50, No. 2 | pp. 167-328
June 2026 | Zagreb

ISSN: 2459-8860
https://doi.org/10.3326/pse.50.2

Institute of Public Finance

CELEBRATING 50

Public Sector Economics
2/2026

Vol. 50, No. 2 | pp. 167-328 | June 2026 | Zagreb

TABLE OF CONTENTS

Convergence of European healthcare systems: a cluster analysis

PAOLO LIBERATI, Ph.D.*

Article**
JEL: H11, H51, I1
https://doi.org/10.3326/pse.50.2.1

* I acknowledge financial support from PRIN 2022 "Reconciling efficiency and equity in the Italian healthcare system after the pandemic crisis", financed by the Italian Ministry of Universities and Research (Project code 2022NKLHHT; CUP Master H53D23002530006; CUP F53D23003130006). The author is grateful to two anonymous referees who have contributed to the quality of the final version of the paper.

** Received: May 21, 2025
Accepted: February 16, 2026

Paolo LIBERATI
University Roma Tre, Department of Economics, Via Silvio D'Amico 77, 00145, Rome, Italy
Research Centre in Economics and Public Finance (CEFIP), Via Silvio D'Amico 77, 00145, Rome, Italy
e-mail: paolo.liberati@uniroma3.it
ORCiD: 0000-0002-6317-9701

PUBLIC SECTOR ECONOMICS
50 (2) 167-213 (2026)

Abstract

This paper examines whether European healthcare systems can be meaningfully grouped into distinct clusters based on institutional characteristics and health outcomes. To address this question, we conduct a cluster analysis of 26 European countries over the period 2001-2021. The findings provide limited empirical support for the conventional classification into Nordic, Continental, and Liberal models within the healthcare domain, both in terms of institutional features and health outcomes. Rather, most Western European countries cluster together, while a distinct and robust cluster of Eastern European countries clearly emerges.

Keywords: health, welfare state, cluster, Europe, Esping-Andersen

1 INTRODUCTION

In recent decades, European healthcare systems have faced increasing pressures from ageing populations, rising healthcare costs, and growing demands for equitable and high-quality services. These challenges have been accompanied by broader transformations in welfare states, including shifts in public spending priorities, marketisation of services, and evolving roles of the state, family, and private actors in providing care.[1] As a result, the organisation and performance of healthcare systems have become not only a matter of national policy, but also a key dimension of comparative welfare research, reflecting how different countries balance efficiency, equity, and social protection.

Accordingly, understanding whether healthcare systems continue to exhibit distinct national patterns or whether they exhibit performance similarities is crucial for both policymakers and academics. If similarities occur, traditional welfare state typologies may be less informative in explaining differences in health system performance. Conversely, persistent clustering along established welfare regime lines would suggest that historical institutional logics continue to shape healthcare provision and outcomes.

Against this backdrop, European healthcare systems can be examined in terms of both their institutional characteristics and health outcomes to assess whether distinct clusters emerge or whether performance similarities increase over time. A natural point of reference for this discussion – although not originally developed for the health sector – is Esping-Andersen's (1990) influential classification of welfare states. Esping-Andersen distinguishes three welfare regimes – liberal, corporatist, and social democratic – based on the extent to which social rights enable de-commodification, understood as individuals' capacity to access essential goods and services independently of market forces. Liberal regimes rely predominantly on markets, with limited public provision, modest benefits, and strict eligibility criteria. Corporatist regimes link social entitlements closely to occupational

[1] Hermann (2010); Taylor-Gooby, Leruth and Chung (2017), especially chapter 1; Pestieau and Lefebvre (2018), especially chapter 11.

status, while social democratic regimes are grounded in universalism and solidarity, providing comparatively generous benefits as a citizenship right. In this pioneering framework, European countries were grouped into distinct and internally coherent models, each characterised by a specific logic of welfare provision.[2]

Building on Esping-Andersen's approach, a vast literature has since emerged seeking to confirm, revise, or challenge the original "three worlds" typology in light of new empirical evidence and theoretical developments.[3] Despite these efforts, the debate remains far from settled. As Powell, Yörük and Bargu (2019:68) note, empirical support for the three-world classification is mixed at best, and few countries fully conform to the ideal-typical characteristics identified by Esping-Andersen. Rather than interpreting Esping-Andersen's typology as a rigid representation of empirical reality, it may therefore be more productive to conceptualise it as a set of ideal types differentiated by policy goals and instruments (Palier, 2010:22). This perspective allows for the possibility that different components of the welfare state – such as health, education, and pensions – may evolve asynchronously with respect to institutional structures, rights, benefits, financing, and organizational arrangements.

Thus, although Esping-Andersen's typology was developed primarily with reference to labour-market regulation and cash benefits, it is theoretically plausible that regime-specific logics of state-market-family relations also shape the organisation, financing, and performance of in-kind services such as healthcare. Indeed, one influential line of critique argues that analysing specific welfare domains separately may better capture cross-national variation, since different sectors do not necessarily follow the same trajectory. From this perspective, policy-specific analyses may offer greater explanatory power than global welfare typologies (Kasza, 2002:284). This is an argument originally advanced by Alber (1995:132) and reiterated by Bertin, Carrino and Pantalone (2021). However, with the notable exceptions of Kangas (1994), who examined health insurance in 1950 and 1985, and Bambra (2005a), the literature has devoted surprisingly limited attention to healthcare services. This omission is striking, given that healthcare is typically delivered in kind, often provides universal coverage, and is largely independent of labour market participation.[4] Health spending also constitutes a major share of public social expenditure – typically second only to pensions – and remains a key item in national budgets. In 2023, average health expenditure across OECD countries was approximately 9.3 per cent of GDP; in Europe, the figure was just below 8 per cent, with most spending allocated to hospital services (OECD, 2025a). Despite widespread recognition of healthcare as a core welfare function, it has often been marginal in comparative welfare analyses – still "at the corner of their eye rather than in the centre of their vision" (Moran, 2000:136).

[2] For a discussion see Bambra (2005b); Bergqvist, Yngwe and Lundberg (2013); Fosse (2011).
[3] See, for example, Arts and Gelissen (2002); Ferragina and Seeleib-Kaiser (2011); Danforth (2014); Emmenegger et al. (2015); Saint-Arnaud and Bernard (2003); Taylor Brown and Ben Brik (2024).
[4] See Rothgang (2021); for a recent review, Powell, Yörük and Bargu (2019).

PUBLIC SECTOR
ECONOMICS
50 (2) 167-213 (2026)

When applied to healthcare systems, the available empirical evidence provides only limited support for Esping-Andersen's original clustering. Most studies agree that reconciling national healthcare arrangements with the three-world typology is problematic, particularly because welfare systems in many European countries have undergone substantial redefinition in recent decades. As a result, a complete overlap between healthcare typologies and the Esping-Andersen model is rarely observed (Bertin, Carrino and Pantalone, 2021). This suggests that health regimes are more hybrid than standard classifications imply,[5] although elements of the typology remain aligned with specific national health policies.[6] Recent contributions further question the continued relevance of prototypical cases. Sowula et al. (2023:11) argue that Germany and Sweden may no longer represent conservative and social democratic regimes, respectively. Similarly, Szebehely and Meagher (2017:304) document processes of de-universalisation across Nordic countries, including rising for-profit provision in publicly funded care, increased reliance on family care, and growing out-of-pocket payments – developments that challenge the principle of universalism.

Debates on a possible convergence of healthcare models reflect similar ambiguities. Giaimo and Manow (1999), analysing reforms in Britain, Germany, and the United States, found no evidence of convergence toward a common reform trajectory or a generalised shift toward privatisation. By contrast, Rothgang et al. (2010) identify convergence across OECD countries in public-private mixes, as measured by both beta and sigma convergence,[7] resulting in increasingly hybrid healthcare systems. Swank (2002:230) similarly notes significant retrenchment and efficiency-oriented reforms across pensions, income support, healthcare, and social services.

Other studies have examined specific healthcare domains. In the field of long-term care, evidence from Europe points to a gradual retrenchment of universalism. Ranci and Pavolini (2015:282) show that, since the financial crisis, universal access has often been curtailed through increased private-sector involvement, giving rise to forms of restricted universalism. While formal entitlements have largely remained intact, effective access has become increasingly dependent on constraints in provision and quality. Importantly, these developments cut across political orientations and are better explained by broader fiscal pressures than by ideological shifts (Barbier, 2012:391). In line with this interpretation, Tine et al. (2022:204) document growing inequalities in Denmark's long-term care provision driven primarily by resource constraints. Primary care, another key healthcare domain, has followed a different but related trajectory. Over the past two

[5] See also Reibling, Ariaans and Wendt (2019).

[6] See, for example, Fosse (2011) on health promotion policies. Similar challenges also arise with earlier classifications, such as Wilensky and Lebeaux (1965) residual-institutional distinction, later adopted by Pinker (1971) and Titmuss (1974). See Higgins (1981); Pierson (2006); Carreira Da Silva (2017).

[7] Beta-convergence occurs when units with initially lower levels of a given indicator grow faster over time than those with higher levels; sigma-convergence refers to a reduction in dispersion across units.

decades, it has been reshaped by systemic reforms, technological innovation, and, more recently, post-pandemic restructuring.[8] By the end of the 2000s, comparative studies identified relatively strong primary care systems – particularly in terms of organizational regimes, accessibility, continuity, and coordination – in countries such as Belgium, Denmark, Estonia, Finland, the Netherlands, Portugal, Slovenia, Spain, and the United Kingdom (Kringos et al., 2013). At the same time, former Eastern bloc countries made substantial progress by strengthening the role of general practitioners and introducing gatekeeping mechanisms aimed at moving away from hospital-centred models. During the 2010s, a range of organisational and professional innovations – including nurse prescribing, patient choice of family doctor, integrated health and social care, electronic health records, and telemedicine – became increasingly widespread. The COVID-19 pandemic subsequently underscored the central role of primary care in system resilience, prompting European WHO member states in 2021 to endorse a comprehensive strategy to further strengthen primary health care.[9]

In light of these transformations, traditional classifications of health systems – based on Bismarck, Beveridge, and Esping-Andersen – appear increasingly inadequate. Bismarckian and Beveridgean models primarily distinguish systems according to financing mechanisms, while Esping-Andersen's framework, as noted above, emphasizes degrees of de-commodification; yet, these dimensions may coexist. As a result, real-world health systems have become increasingly hybrid, limiting the explanatory power of these typologies for understanding contemporary patterns of access, inequality, and system performance.[10]

Taken together, this literature does not imply that welfare state classifications have become irrelevant, but rather that their applicability – particularly to healthcare – cannot be taken for granted. The aim of this paper is therefore to assess whether, and to what extent, observable patterns in healthcare systems and outcomes across Europe can be identified. By systematically analysing health-related indicators over a long-time horizon, the paper evaluates whether established welfare regime distinctions still correspond to meaningful differences in health system performance, or whether new empirical clusters have emerged that cut across traditional classifications.

The analysis adopts a long-term perspective, covering the period from 2001 to 2021, to capture cumulative policy effects and avoid short-term fluctuations – particularly those associated with the 2008 financial crisis. The study includes 26 European countries, enabling comparisons among states operating within broadly similar socio-economic and regulatory contexts, yet still exhibiting substantial diversity in organization and access rules, financing mechanisms, and health

[8] The European Practice Assessment (EPA), introduced in the early 2000s, established a common framework for evaluating primary care performance across countries.
[9] See Kringos et al. (2015); Gilardi, Füglister and Luyet (2009); Dubas-Jakóbczyk et al. (2020); Polin et al. (2021).
[10] On Bismarck-type models, see Palier (2010).

PUBLIC SECTOR ECONOMICS
50 (2) 167-213 (2026)

outcomes. This diversity provides a rich empirical basis for assessing whether national health systems conform to Esping-Andersen's typology or have developed hybrid configurations in response to internal and external pressures, including EU-level recommendations promoting fiscal consolidation and efficiency following the crisis (Alsasua, Bilbao-Ubillos and Olaskoaga, 2007; Bilbao-Ubillos, 2023; Vaughan-Whitehead, 2017).

Methodologically, the paper applies cluster analysis to a set of indicators capturing multiple dimensions of healthcare, with the aim of identifying emerging empirical patterns, while acknowledging the inherently fuzzy boundaries of welfare regimes (Shalev, 2007; Taylor Brown and Ben Brik, 2024). Findings indicate that healthcare clustering in Europe does not closely align with traditional welfare state classifications. Instead, there is evidence that Western European countries tend to cluster together, primarily reflecting performance similarities in health outcomes rather than a true convergence in institutional structures. At the same time, a distinct group of Eastern European countries emerges, with this pattern remaining robust across alternative cluster specifications.

2 METHODS

2.1 THE CLUSTER ANALYSIS

The available literature on the classification of health regimes has made extensive use of cluster analysis. One fundamental reason is that understanding different welfare state regimes often requires examining multiple dimensions that cannot easily be reduced to a single, cross-nationally comparable indicator. When multidimensionality is an issue, clustering countries according to their internal cohesion (homogeneity), while also clarifying the nature of their external separation, adds value to the categorization of welfare states (Obinger and Wagschal, 2001; Kautto, 2002; Powell and Barrientos, 2004; Jensen, 2008; Wendt, 2009; 2014; Minas, Jacobson and McMullan, 2014; Reibling, Ariaans and Wendt, 2019). Additionally, Gough (2001:169) describes cluster analysis as "robust, meaningful and simple," strongly recommending it for the study of welfare regimes. In other contexts, Witt et al. (2018:21) also argue that cluster analysis can be used to investigate the complex and interrelated dimensions of nations, serving as "a foundational tool for sense-making and conceptualization of the object under investigation" or, as Ermakoff (2019) suggests, as a way of letting the data speak for themselves.

In the context of cross-sectional time-series data, a key methodological consideration is whether to perform a cluster analysis separately for each year or, alternatively, to conduct a single cluster analysis on pooled data. In what follows, the analysis initially proceeds by estimating a yearly-cluster solution. This strategy enables the examination of potential movements of countries across clusters over time. However, it implicitly treats each year as an independent dataset, thereby precluding the identification of trajectory effects and limiting the analysis to the observation of repeated relative positions. Such a limitation is particularly consequential when the objective is to assess the presence of absolute convergence in health outcomes across countries. For this reason, a subsequent step involves the implementation of a cluster analysis on pooled data, which allows for a more direct investigation of whether countries display increasing similarity in the joint distribution of structural and performance-related indicators over the period considered. In this context, the pooled-data clustering is intended to capture overall patterns of similarity across countries while retaining the time dimension at the observation level. Although this approach does not explicitly model transition dynamics or provide formal convergence tests, it allows for a comparative assessment of how country profiles relate to one another over time.[11]

Moreover, to prevent variables measured in different units from disproportionately influencing the dissimilarity matrix, all variables are standardised using a z-scoring procedure. Specifically, each variable is transformed by subtracting its mean and dividing by its standard deviation.[12] It is important to note that the choice between year-by-year and pooled clustering entails different z-scoring strategies. When clustering is performed separately for each year, standardisation is based on the mean and standard deviation computed within that specific year. Under this approach, even substantial changes in a variable's scale within a given year – for instance, a doubling of its values – leave its z-scores unchanged and therefore do not affect the resulting cluster solution. Moreover, because standardisation is applied across countries at each time point, mean-level changes over time cannot be detected, as all annual means are normalised to zero. By contrast, when clustering is conducted on pooled data, a sudden shift in a variable in a single year (such as a doubling of its values) alters the overall mean and standard deviation, thereby affecting the corresponding z-scores and potentially modifying the resulting clusters. Although year-by-year analysis can still provide valuable inter-temporal information, it is important to consider that the distribution of multiple indicators typically evolves unevenly across years. As a consequence, year-specific standardisation may modify the relative contribution of each variable to the clustering solution, potentially introducing artefactual temporal variation and complicating the interpretation of long-term patterns.[13] Since the two approaches

[11] As recently argued by Lu (2024), when applied to longitudinal data, clustering seeks to group units of analysis (individuals, regions, countries, etc.) according to the patterns and trends characterizing their repeated measurements over time. Hierarchical clustering represents one methodological option capable of accommodating temporal trajectories, even in the presence of sparse or irregular observations (Zhou, Zhang and Tu, 2023). Applications of longitudinal clustering span multiple domains, including criminology, sociology, medicine, and ecology. Recent examples include the identification of subgroups exhibiting distinct cigarette-smoking patterns for the purpose of predicting health outcomes (Lee et al., 2016), as well as the characterization of adolescent substance-use trajectories and their association with leisure experience (Weybright et al., 2016). For a comparative assessment of methods for clustering longitudinal data, see Den Teuling, Pauws and van den Heuvel (2021).

[12] After z-scoring, the variance of each variable equals 1, which can be interpreted as a special case of weighting based on the reciprocal of the sample standard deviation. This implies that, under the z-score transformation, the influence (or weight) of a variable decreases as its variability increases.

[13] To assess the temporal properties of the series across countries, correlograms were computed for each national time series. The analysis revealed significant autocorrelation in all cases, indicating that observations are not temporally independent but are influenced by their own past values. This evidence supports the use of a pooled-data framework for the subsequent cluster analysis, as it enables the identification of cross-country similarities in underlying dynamic patterns while appropriately accounting for the serial dependence inherent in each series and, potentially, cross-correlation across countries.

PUBLIC SECTOR ECONOMICS
50 (2) 167-213 (2026)

can be viewed as complementary rather than mutually exclusive, the cluster analysis is performed using both methods.

With regard to the clustering method, the analysis is conducted using hierarchical agglomerative clustering, which involves a series of successive agglomerations of *n* countries into groups, beginning with *n* single-member clusters. A limitation of this method is that, once two countries are merged at a given agglomeration step, they cannot be separated in subsequent steps (Kaufman and Rousseeuw, 1990).

Regarding the dissimilarity measure, the analysis employs the squared Euclidean distance in combination with Ward's method. This hierarchical clustering approach merges clusters according to the criterion of minimizing total within-cluster variance (i.e., the error sum of squares). Ward's method is particularly effective in reducing within-cluster heterogeneity, often yielding clusters that are more compact and better separated.[14]

As a final point, a decision must be made regarding the definition of the optimal number of clusters. To this end, the most common and reliable stopping rules are those proposed by Duda and Hart (1973) and Calinski and Harabasz (1974), based, respectively, on a pseudo-T-squared statistic and a pseudo-F statistic.[15] A sensitivity analysis will additionally be conducted using the gap statistic.

2.2 DATA AND VARIABLES

The data used in this analysis are drawn from several sources: OECD Health Statistics; the European System of Integrated Social Protection Statistics (ESSPROS) database; the Health for All (HFA) database of the World Health Organization (WHO); the Social Insurance Entitlement Dataset (SIED); and the Comparative Political Data Set (CPDS). Data are collected for 26 European countries over the period 2001-2021; however, not all countries are observed in every year. Notable exceptions include Iceland (observed in 2005, 2010, 2015, and 2020), Estonia (observed from 2007 to 2021), Luxembourg (observed in 2010, 2015, and 2020), and Norway (observed from 2008 to 2021). Consequently, the resulting dataset constitutes an unbalanced panel.

The variables used in this paper capture various dimensions of the health sector. To address the observation that health spending alone is insufficient to meaningfully distinguish health regimes – due to the general tendency of countries to allocate similar shares of GDP to health – a set of health indicators is introduced to reflect the different aspects of health provision, as also suggested by the OECD classification

---

[14] Given two pairs of clusters with equally distant centroids, Ward's linkage will preferentially merge the smaller pair, aiming to minimize the total within-cluster inertia (or error sum of squares) rather than the direct distance between clusters. This implies that Ward's method is, in principle, less sensitive to noise and outliers, and it tends to produce compact, approximately spherical clusters of similar size and variance (see Everitt et al., 2011, Chapter 4; Kaufman and Rousseeuw, 1990).

[15] See also Milligan and Cooper (1985).

of health data (Esping-Andersen, 1990:19; Castles, 2002:616; 2007; Jensen, 2008:159).

Furthermore, the variables were selected to cover the period from 2001 to 2021 for 26 European countries without significant gaps. Occasional missing values were addressed by extrapolating the observed series. Limitations in the available data prevented the use of other potentially relevant indicators.[16] The analysis initially employs variables that allow all 26 countries to be included; subsequent analyses consider variables that may necessitate the exclusion of one or more countries. The variables used are described in table 1 and cover various dimensions of health characteristics that, in principle, allow for the differentiation of health regimes according to multiple aspects: (a) health expenditures; (b) health status; (c) health quality; (d) health care utilisation; (e) health insurance; (f) long-term care; (g) health resources; and (h) health protection (Kuitto, 2011:350).

To begin with, health expenditures reflect the share of public resources allocated to health, although they do not fully capture the intensity or quality of care across countries (Huber and Stephens, 2001). Health expenditures can, in fact, embody the labour-intensive nature of medical services and the differential dynamics of their prices relative to the overall price index, as well as the potentially rising costs of technological innovations, particularly in specific medical areas (Sorenson, Drummond and Kahn, 2013:228). They may also reflect the increased demand for therapies resulting from the success of medical treatments and the growing number of elderly chronically ill patients, generating a kind of expenditure-growth cycle (Zweifel, Steinmann and Eugster, 2005; Breyer and Lorenz, 2020). Furthermore, as noted by Rothgang et al. (2010), tracking public health expenditures over time may provide insights into certain forms of "indirect" privatisation of the health system (Bambra, 2005a). Overall, as suggested by Freeman (1999), public funding of health care is often considered a key indicator of governmental involvement in health provision.[17]

[16] For example, the Global Health Database provides very limited data even for European countries on key variables related to primary care and health quality indicators.
[17] See also Burau and Blank (2006:65).

Table 1
Variables and source of data

OECD category	Description	Source	Variable	Unit of measure
Health expenditures	Health spending	OECD HS	health	Percentage of GDP
	Cash benefits non means-tested	ESSPROS	V1A	Share of total health benefits
	In kind benefits non means-tested	ESSPROS	V1B	Share of total health benefits
	Cash benefits means-tested	ESSPROS	V1C	Share of total health benefits
	In kind benefits means-tested	ESSPROS	V1D	Share of total health benefits
	Households out-of-pocket payments	OECD HS	outpocket	Percentage of GDP
	Private sector expenditures	WHO	H568	Percentage of GDP
	Public pharmaceutical expenditures	WHO	H580	Percentage of total spending
Health status	Life expectancy at birth	OECD HS	lifexpbirth	Age
	Death rate	OECD HS	deathrate	Per 1,000 inhabitants
	Population aged 65+	OECD HS	oldpop	Percentage of total population
Health quality	Congestive heart failure hospital admissions	OECD HS	congestive	Per 100,000 inhabitants
	Asthma and COPD hospital admissions	OECD HS	asthma	Per 100,000 inhabitants
	Diabetes hospital admissions	OECD HS	diabete	Per 100,000 inhabitants
Health care utilisation	Inpatient discharges	OECD HS	disch	Per 100,000 inhabitants
	Average length of stay (all hospitals)	WHO	H540	Number of days
Health insurance	Government compulsory health insurance	OECD HS	govins	Percentage of the population
	Voluntary health insurance	OECD HS	volins	Percentage of the population
Long-term care	Beds in residential long-term care facilities	OECD HS	longtermbeds	Per 1,000 population aged 65+
Health resources	Total health employment	OECD HS	totemp	Per 1,000 inhabitants
	Number of beds in publicly owned hospitals	OECD HS	beds	Per 1,000 inhabitants

OECD category	Description	Source	Variable	Unit of measure
Health protection	Sickness, gross 26-week replacement rate	SIED	srtsw26s	Percentage of wage
	Weeks during which sickness benefit is payable	SIED	sduratio	Number of weeks
	Waiting days before sickness payment	SIED	swaiting	Number of days
	Coverage ratio as proportion of labour force	SIED	scovratl	Percentage of workers
Institutional and political variables	Cabinet composition	CPDS	gov-party	Index from 1 to 5
	Voter turnout in elections	CPDS	vturn	Percentage of voters
	Fractionalization of party system	CPDS	rae-ele	Index from 0 to 1
	Index of federalism	CPDS	fed	Index from 0 to 2
	Executive-Legislative relations	CPDS	pres	Index from 0 to 3

Note: OECD HS = OECD Health Statistics, ESSPROS = European System of Integrated Social Protection Statistics, WHO = World Health Organization, SIED = Social Insurance Entitlement Dataset, CPDS = Comparative Political Data Set.

Still on the side of monetary resources, a set of additional variables is considered:[18] means-tested and non-means-tested cash benefits; means-tested and non-means-tested in-kind benefits; household out-of-pocket payments as a percentage of GDP; private-sector expenditures as a percentage of GDP; and public pharmaceutical expenditures as a percentage of total health spending.[19] In particular, including cash transfers in our analysis helps identify the terms on which individuals can make claims on public resources and the type of solidarity fostered by systems of public support. The way cash transfers are organised also reveals the priorities that underpin governmental action (Daly, 1997). Moreover, observing cash and in-kind transfers over the last twenty years may capture the effects of reforms that have potentially restricted access to benefits by tightening eligibility criteria and regulations; reforms aimed at increasing the use of means-testing and replacing universal payments; or, finally, reforms promoting the expansion of private-sector involvement through the contracting out of services.[20]

In order to understand the dynamics of healthcare systems, it is also necessary to complement the previous information with data on the characteristics of health protection, so as to capture the relevance of the institutional architecture – particularly access requirements and the duration of sickness benefits – as a proxy for the decommodification of the labour market. To this purpose, four variables are included in this group: the gross 26-week replacement rate for sickness provisions; the number of weeks during which sickness benefits are payable; the number of "waiting days" at the beginning of a sickness spell when no benefits are paid; and the coverage ratio of sickness benefits as a proportion of the labour force. These variables make it possible to ascertain whether certain elements of health systems converge across countries, thereby contributing to a broader convergence of other system components (Barros, 2007; Leiter and Theurl, 2012).

With regard to health resources, two main indicators are used: the total health workforce and the number of beds in publicly owned hospitals. In particular, total employment can be interpreted as a complement to the level of health spending. As argued by Wendt and Kohl (2010), there may be only a weak correlation between the financial resources invested in a nation's health system and the level of health employment, which suggests that total employment should be taken into account when assessing health performance.

Concerning health care utilisation, reference is made to inpatient discharges and to the average length of stay in hospital. To some extent, these two variables can provide information about the "indirect privatisation" of the health service,

---

[18] See Sowula et al. (2023) for Sweden and Germany.

[19] Means-tested and non-means-tested cash and in-kind benefits are included under the category of health expenditure, as they represent either a direct monetary transfer or a direct form of public spending. To some extent, they also constitute an element of health coverage. However, both in-kind and cash components of the welfare state are classified within the same category, as part of a broader definition of collective resources. See, for example, Lundberg et al. (2015).

[20] See Polin et al. (2021) for a description of health reforms in 2018-2019.

achieved through shorter lengths of stay and a more rapid turnover. However, as shown in several studies, this does not necessarily imply lower prices or higher quality. Braithwaite, Travaglia and Corbett (2011), reviewing a large number of articles, found weak and at times conflicting evidence, while Tiemann, Schreyogg and Busse (2012) also showed that private hospital ownership is not necessarily associated with higher efficiency in the case of Germany.

Concerning health quality, the lack of available data for the entire period and for all countries substantially restricts the possible choices. One option – with gaps – is to use the rate of hospital admissions for congestive heart failure as a proxy for the quality of treatments. Asthma and chronic obstructive pulmonary disease (COPD) hospital admissions, as well as diabetes hospital admissions, are also included in some steps of the analysis, although this entails losing observations for certain countries.

With regard to health status, three indirect indicators of health outcomes have been selected: the share of the population aged over 65, life expectancy at birth, and the death rate. On the one hand, these indicators may imperfectly represent the general quality of "health outcomes"; on the other hand, they may help investigate the presence of performance similarities in health outcomes, given the correlation commonly observed between health care resources and health conditions. The choice is debatable also because the ranking of health performance proposed by the World Health Organization (WHO) in 2000 was criticised on academic grounds[21] and was effectively rejected at the political level by countries that had received an unfavourable ranking. Nevertheless, these variables are included in order to capture those cases, if any, in which comparable amounts of resources may give rise to non-comparable health outcomes.

The dimension of health insurance is also considered, including the share of the population covered by compulsory government health insurance and the share covered by voluntary health insurance. Furthermore, the dimension of long-term care is approximated by the number of beds in residential long-term care facilities.

Finally, a set of institutional and political variables is included to serve as proxies for missing data on welfare institutions. Specifically, the following variables are introduced: the composition of the cabinet, particularly whether it leans to the right or left; voter turnout in elections, as a proxy for individuals' participation in the political process; an index of party system fractionalization, which may correlate with the number of veto points affecting welfare provision; an index of federalism, to account for the degree of decentralization in the delivery of health services; and the nature of executive-legislative relations – particularly the prevalence of either parliamentary or presidential systems – which may also influence the overall structure of the welfare state.

[21] See, for example, Musgrove (2003).

It is important to note that the variables discussed above capture structural, process, and outcome dimensions. Accordingly, any observed convergence would reflect the combined effect of these indicators, without allowing for a clear identification of their individual contributions. However, the empirical analysis attempts to isolate the extent to which convergence can be attributed to structural rather than to performance-related factors.

Of particular interest in the attempt to classify health regimes is the possibility that the same country may belong to different clusters over time, a feature that allows for the consideration that health regimes – even within a given country – may have changed.[22] On the other hand, a general caveat in this analysis is that not all variables will be used simultaneously, as the inclusion of certain variables entails gaps either in years or in countries. However, in a first step, the focus will be on those variables that maximise the number of countries observed.[23]

3 RESULTS

3.1 STEP 1: REPLICATING EXISTING CLUSTERING

The analysis begins with a hierarchical clustering method based on the set of variables common to all European countries included. In a first step, the analysis seeks to maximise comparability by replicating the cluster analysis using the number of clusters proposed in similar studies, regardless of the specific information provided by the stopping rule. To this end, table 2 first describes how different countries have been grouped in the health sector by each available study. It can be noted that one study classifies countries into three clusters (Wendt, 2009), one study into five clusters (Reibling, Ariaans and Wendt, 2019), and one study into six clusters.[24] The remaining three studies identify four clusters (Jensen, 2008; Reibling, 2010; Wendt, 2014).

The same table also attempts to show how often the same countries are grouped together in different studies. Some persistence is observed for certain groups of countries: in particular, Austria, Belgium, France, Germany, and Luxembourg (all countries or a subset thereof) are often grouped in the same cluster; Finland, Spain, and Portugal form another frequently observed group; Denmark and the

[22] See, for example, De Simone, Gaeta and Ercolano (2012).

[23] A principal component analysis (PCA) was also performed to identify core characteristics of the health system. Although detailed results are not reported in the text, the analysis suggests that each variable has a significant loading (either higher or lower than ±0.3) on at least one component, and that each component can be meaningfully interpreted in terms of healthcare characteristics. If every variable has at least one significant loading on some component, it contributes to the variance (or correlation structure) of the data within the PCA subspace. This reduces the risk of including completely "noisy" or irrelevant variables. It also implies that clustering performance is less affected by the presence of less influential variables, as long as they are not entirely "silent" – i.e., they contribute with a significant loading somewhere (see, for example, Giuliani and Vici, 2024). Nevertheless, the presence of many "weak" variables – even if minimally contributing – can still introduce noise, increase model complexity, and reduce the stability or interpretability of the clusters. To avoid arbitrary variable selection, especially given the constraints posed by limited data availability, all variables were retained in the subsequent cluster analysis.

[24] Joumard, André and Nicq (2010). OECD (2025b) also provides a cluster analysis, identifying seven distinct clusters that include European countries. However, the analysis is not limited to European countries, encompassing other OECD member states.

PUBLIC SECTOR
ECONOMICS
50 (2) 167-213 (2026)

Netherlands; Italy and the United Kingdom; and Estonia, Hungary, Poland, and Slovakia are recurrently grouped together across studies. It is also worth noting that in Joumard, André and Nicq (2010), two clusters are very small, as one includes only Belgium and France, and another includes only Iceland and Sweden. Furthermore, in Jensen (2008) there is a one-country cluster (Ireland) out of a total of four, while in the cases of Wendt (2009; 2014), some countries are not classified in any cluster (Greece, the Netherlands, Norway, and Switzerland). As it stands, table 2 leaves a considerable degree of uncertainty regarding the possibility of a stable classification of health regimes.

In the second stage, a hierarchical cluster analysis on pooled data is conducted to cover a range of 3 to 6 clusters, as in the previous studies, regardless of the optimal number of clusters identified by the stopping rule. For this exercise, average linkage is used for the sake of comparability, as it is the method most commonly adopted in the studies reported in table 2.

Tables A1 to A4 in appendix show the results. In all cases, and in contrast with the findings of other studies, a meaningful clustering of countries is hardly identifiable. Instead, two countries (Iceland and Ireland) consistently form their own separate clusters – an outcome that is insufficient to support claims of distinctive healthcare models, but that may suggest a broader convergence, at least when compared to the classifications emerging in earlier studies. Moreover, replicating the analysis while stopping at two clusters again reveals a lack of differentiation among countries, with the second cluster consisting solely of the Netherlands in five years (not reported in the table).

Table 2
Health systems: clusters of countries

Cluster	Jensen (2008)	Wendt (2009)	Reibling (2010)	Joumard et al. (2010)	Wendt (2014)	Reibling, Ariaans and Wendt (2019)
A	Spain	Spain	Italy	Hungary	Iceland	Finland
	Norway	Finland	Portugal	Ireland	Finland	Portugal
		Portugal	Finland	Italy	Spain	Sweden
				United Kingdom	Portugal	Norway
				Norway	Sweden	
				Poland		
B	Denmark	Denmark	Denmark	Denmark	Denmark	Denmark
	Netherlands	Ireland	Netherlands	Spain	Czechia	Netherlands
	Finland	Italy	Spain	Finland	Estonia	Spain
	Sweden	United Kingdom	Poland	Portugal	Hungary	United Kingdom
	United Kingdom				Poland	Italy
					Slovakia	
					Ireland	
					Italy	
					United Kingdom	
					Netherlands	
					Slovenia	

Cluster	Jensen (2008)	Wendt (2009)	Reibling (2010)	Joumard et al. (2010)	Wendt (2014)	Reibling, Ariaans and Wendt (2019)
C	Austria	Austria	Austria	Belgium	Austria	Austria
	Belgium	Belgium	Belgium	France	Belgium	Belgium
	France	France	France		France	France
	Germany	Germany	Sweden		Germany	Germany
	Italy	Luxembourg	Switzerland		Luxembourg	Luxembourg
						Czechia
						Iceland
						Ireland
D	Ireland	–	Czechia	Germany	Greece	Estonia
			Germany	Netherlands		Hungary
			Greece	Slovakia		Poland
				Switzerland		Slovakia
E	–	–	–	Austria	–	Switzerland
				Czechia		
				Greece		
				Luxembourg		
F	–	–	–	Iceland	–	–
				Sweden		
NC	–	Greece	–	–	Norway	–
		Netherlands			Switzerland	

Note: NC = Not assigned to any cluster.

3.2 STEP 2: A REPEATED CLUSTER ANALYSIS FOR EACH YEAR

In this section, a new cluster analysis is performed based on the information provided by the Duda-Hart stopping rule and using Ward's linkage method. To obtain a clearer picture of clustering dynamics over time, the analysis is conducted separately for each year, treating observations as independent across time. This assumption will be relaxed in the following section. A general caveat in this type of analysis is that the optimal number of clusters should not be expected to remain constant across years. As shown in table 3, the number of clusters ranges from a minimum of 2 in 2021 to a maximum of 6 in 2009.[25]

The evolution of healthcare system typologies across European countries reveals both persistent structures and significant temporal shifts. On the one hand, there is a relatively stable quasi-Nordic-Continental model, encompassing countries with both Bismarck-type and Beveridge-type institutional arrangements, such as Austria, Belgium, Denmark, Germany, the Netherlands, Norway, Spain, Sweden, and Switzerland. These nine countries exhibit very high co-clustering stability overall from 2001 to 2021. The core group – Austria, Belgium, Germany, and the Netherlands – is clustered together 100 percent of the time. Countries such as Spain, Sweden, Norway, and Switzerland also maintain strong alignment, usually above 90 percent, although some pairs involving Denmark and Switzerland show lower consistency (around 75-85 percent). In short, this group represents a stable Western/Northern European cluster, with minor fluctuations mostly associated with Denmark and Switzerland.

This pattern suggests strong continuity in their institutional and policy frameworks, as well as relative internal homogeneity. Notably, this occurs despite Spain and Sweden typically being classified as Beveridge-type welfare systems, in contrast to the Bismarck-type model followed by the other cluster members. This may indicate that similar healthcare outcomes can emerge from different institutional configurations, thereby calling into question the determinative power of welfare regimes in shaping health system performance. These findings resonate with earlier work highlighting the diminishing explanatory power of traditional institutional models, such as Beveridge and Bismarck, in accounting for differences in health system performance.[26]

[25] The optimal number of clusters is determined using both the pseudo-F and pseudo-T-squared statistics from the Duda-Hart rule. In general, higher pseudo-F values indicate greater separation between clusters, suggesting a more distinct clustering structure. Conversely, lower pseudo-T-squared values imply greater within-cluster homogeneity. Thus, the optimal number of clusters reflects a balance between cluster separation and internal cohesion.

[26] This result is consistent with a recent study by OECD (2025b), which – addressing three key policy areas in the health domain – shows that there is no indication that any one group of health systems would systematically outperform another. See also Wendt (2009); Böhm et al. (2013). As argued by Giaimo and Manow (1999) and Reibling (2010), despite their formal differences, systems often converge functionally in response to shared policy pressures.

Table 3

Cluster assignments by country and year (2001-2021)

Country	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Austria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Belgium	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Czechia	4	4	4	4	4	3	3	5	6	4	4	4	3	3	1	4	3	3	4	3	2
Denmark	1	1	1	1	1	1	1	2	2	3	1	1	1	1	3	1	1	1	1	1	1
Estonia	–	–	–	–	–	–	–	3	3	5	6	4	4	4	3	3	5	4	3	3	2
Finland	4	1	1	1	1	1	1	2	2	3	1	1	1	1	1	1	1	1	1	1	1
France	2	2	2	2	2	1	1	3	3	1	2	2	2	2	2	2	1	1	2	2	1
Germany	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Greece	2	2	2	1	1	3	1	3	4	1	2	2	2	2	1	1	1	1	1	1	2
Hungary	4	4	4	4	4	–	3	5	6	4	4	4	3	3	5	4	3	3	4	3	2
Iceland	–	–	–	–	1	–	–	1	–	–	3	–	–	–	–	1	–	–	–	–	–
Ireland	3	3	3	3	3	2	2	4	5	2	3	3	2	2	4	3	2	2	3	2	1
Italy	1	1	1	1	1	1	1	3	4	1	2	2	2	2	1	1	1	1	1	1	2
Latvia	4	4	4	4	4	3	3	5	6	4	4	4	3	3	5	4	3	3	4	3	2
Lithuania	4	4	4	4	4	3	3	5	6	4	4	4	3	3	5	4	3	3	4	3	2
Luxembourg	–	–	–	–	–	–	–	–	3	–	–	–	–	–	3	–	–	–	–	–	1
Netherlands	1	1	1	1	1	1	1	2	2	3	1	1	1	1	3	1	1	1	1	1	1
Norway	–	–	–	–	–	–	–	–	2	2	3	1	1	1	1	3	1	1	1	1	1
Poland	4	4	4	4	4	3	3	5	6	4	4	4	3	3	5	4	3	3	4	3	2
Portugal	2	2	2	2	2	1	1	3	3	1	2	2	2	2	2	2	1	1	2	2	1
Slovakia	4	4	4	4	4	3	3	5	6	4	4	4	3	3	5	4	3	3	4	3	2
Slovenia	2	2	2	2	2	1	1	3	3	1	2	2	2	2	2	2	1	1	2	2	1
Spain	1	1	1	1	1	1	1	3	4	1	2	2	2	2	1	1	1	1	1	1	1
Sweden	1	1	1	1	1	1	1	2	2	3	1	1	1	1	3	1	1	1	1	1	1
Switzerland	1	1	1	1	1	1	1	1	1	1	2	2	1	2	1	1	1	1	2	2	1
United Kingdom	2	2	2	1	2	1	1	3	4	1	2	2	2	2	1	1	1	1	1	1	2
Optimal number of clusters	4	4	4	4	4	3	3	5	6	4	4	4	3	3	5	4	3	3	4	3	2

On the other hand, a distinct Eastern European model emerges, including countries such as Czechia, Estonia, Hungary, Latvia, Lithuania, Poland, and Slovakia, all of which have welfare systems broadly inspired by the Bismarck tradition. These seven countries form an exceptionally cohesive cluster throughout the entire period from 2001 to 2021, with a co-clustering percentage above 94 percent for each pair of countries, and 15 pairs consistently falling in the same clusters across the whole period. After 1990, these countries underwent a transition from the common legacy of socialism to more Western-type economies. Their health systems, to some extent, may have experienced a hybridization process, which often involved a form of "shock therapy" by pursuing the privatisation of some welfare programmes and partly shifting the responsibility for their provision to social funds and private insurance markets.[27] As also suggested by Cerami (2010:249), from a broader perspective, after 1990 "the introduction of a welfare system based on professional diversity and private arrangements [...] corresponded to the functional necessity of occupational and market diversification, which stemmed from the excessively centralized and homogenized economic system in force during communism". Accordingly, in Eastern European countries, a new and distinctive welfare regime has emerged, in which traditional welfare models have been recombined – so much so that it recently also includes some non-Eastern European countries, such as Greece, Italy, and the United Kingdom. This unexpected assimilation may reflect the impact of structural reforms, economic constraints, or divergent responses to exogenous shocks, such as the COVID-19 pandemic.[28] In this regard, the results presented in table 3 lay the groundwork for assessing whether a broader convergence may occur in the future – potentially reflecting the many health reforms implemented in European countries in areas such as organizational regimes, hospital care, primary and ambulatory care, care coordination and specialised services, and healthcare purchasing and payment.[29]

In between, there is a mixed group of countries – Finland, France, Greece, Italy, Portugal, Slovenia, and the UK – which often cluster together despite featuring Bismarck-type, Beveridge-type, and hybrid healthcare systems. Their co-clustering, however, is less consistent than in the other groups. In particular, France, Portugal, Italy, Slovenia, and the UK form a relatively stable cluster, co-clustering more than 75 percent of the time across the 21-year period. Finland shows a weaker and more variable clustering pattern, aligning less consistently with the group, especially with Greece and the UK (co-clustering rates below 60 percent). Greece displays moderate co-clustering stability, with stronger ties to France and Italy (about 76-81 percent), but weaker connections to Portugal, Slovenia, and the UK. Overall, this group reflects a mix of southern and northern European countries with varying degrees of cohesion, with France, Portugal, Italy, Slovenia, and the UK forming the core cluster.

[27] Castles and Obinger (2008); Kuitto (2016). For a different perspective, see Filipovic and Dobrotic (2022:199).
[28] See Lamping and Rüb (2004) for the definition of a "recombinant welfare system" in the case of pensions.
[29] Polin et al. (2021); Dubas-Jakóbczyk et al. (2020). The argument was also anticipated by Ruggles and O'Higgins (1987).

In any case, it is worth noting that in 2021 – following the pandemic crisis – clustering became more homogeneous, as the previously mixed group tended to redistribute either into a broader Continental model or into the Eastern European cluster. This shift introduced hybrid elements into what had previously been a more internally consistent grouping, providing some evidence in support of the convergence hypothesis.

With regard to Ireland's peculiar position – often appearing as a one-member cluster – it can be noted that, in the dataset, the values of several variables lie at either the upper or lower extreme. In particular, this is the case for the death rate, the old-age population share, cash and in-kind means-tested benefits, waiting days of sickness before payment, public pharmaceutical expenditures, and the average length of stay in hospital. Once these variables are removed from the cluster analysis, Ireland no longer forms a one-member cluster (results not reported in the table). In addition to these dataset-specific features, there are other characteristics unique to Ireland. For example, during the period 2010-2015 (included in the analysis), it was observed that "approximately 40-60 per cent of the reduction in emergency inpatients' length of stay observed during this period may have been a result of bed supply reductions experienced in those years" (Walsh et al., 2022:506). Moreover, it has been noted that long waits for public hospital services are a feature of the Irish healthcare system, with limited evidence that waits for private hospital services are shorter (Whyte, Connolly and Wren, 2020; Brick and Connolly, 2021), a characteristic that distinguishes Ireland from other countries, where waiting times in the private sector tend to be shorter. Not to be underestimated with respect to its impact on the entire system and the overuse of hospital care is the fact that Ireland is the only European country without universal coverage of primary health care (The Lancet Regional Health – Europe, 2025). Finally, in 2019 "Ireland […] had the third lowest public hospital bed density in the EU, standing at over 40 per cent below the EU average […]. In 2021, outpatient care accounted for 40 per cent of total out-of-pocket spending – a share twice the EU average, which partly stems from the fact that the majority of the Irish population access GP services on a private basis" (McGeady et al., 2024:162). With regard to the public-private mix, it should also be noted that Ireland has long had the highest share of duplicative private insurance in Europe (about 45 per cent of the population), that is, private coverage for healthcare services already available in the public sector (OECD, 2018:174).

All these elements make Ireland a case with peculiar characteristics compared to other countries. Since 2017, Ireland has been undertaking health system reform through the Sláintecare initiative, which aims to create a more universal and publicly funded healthcare system with equal access for all. The reform is still ongoing, and the system is currently in a transitional phase. Due to this transitional state, combined with persistent challenges such as limited resources and a funding model that mixes public funding with private insurance, Ireland does not fit neatly into conventional healthcare system categories. This makes it stand out as a distinct case in health system comparisons. Moreover, this transitional phase – together with the impact of the pandemic – may explain why Ireland no longer appears as a separate cluster in 2020 and 2021 (see table 3).[30]

3.3 STEP 3: A CLUSTER ANALYSIS WITH POOLED DATA

The year-to-year analysis highlights both episodes of differentiation in the periods 2001-2005 and 2008-2009, as well as similarities in more recent years. A complementary way to analyse the temporal dynamics is to conduct a cluster analysis on pooled data, which removes the assumption that each year represents an independent dataset, in order to assess whether the observed similarities in recent years are relatively stronger than the differentiation seen at the beginning of the century. For this purpose, the analysis initially uses the variables that maximise the number of countries and observations. Applying Ward's method again, the analysis identifies three as the optimal number of clusters. However, it is worth noting that the combination of the index calculated as the ratio of the within-cluster sum of squares after and before the split, and the pseudo-T-squared test (both outputs of the Duda-Hart rule), could also support an optimal number of clusters equal to two.

Table 4

Cluster composition

Cluster 1	Cluster 2	Cluster 3
Austria	France	Czechia
Belgium	Ireland	Estonia
Denmark	Portugal	Hungary
Finland	Slovenia	Latvia
Germany		Lithuania
Greece		Poland
Iceland		Slovakia
Italy		
Luxembourg		
Netherlands		
Norway		
Spain		
Sweden		
Switzerland		
United Kingdom		

Note: Optimal number of clusters: 3 (Duda-Hart rule, Ward linkage).

[30] Sláintecare is a cross-party, ten-year plan launched in Ireland in 2017 to create a universal, single-tier health and social care system. The plan aims to provide equitable access to care based on need rather than ability to pay, focusing on population health, integrated primary and community care, and reforms to entitlements, funding, and implementation. Key goals include shifting care to the community, expanding workforce capacity, and increasing public investment to deliver integrated, person-centred services for all Irish residents. See Burke et al. (2018); Thomas et al. (2021).

Results are reported in table 4, where it is evident that an Eastern European model emerges in Cluster 3. Cluster 1, by contrast, represents a blending of countries traditionally associated with all major welfare state classifications: the Nordic model (Denmark, Finland, Iceland, Norway, Sweden), the Continental model (Austria, Belgium, Germany, Luxembourg, the Netherlands, Switzerland, and the United Kingdom), and the Southern European model (Greece, Italy, and Spain). Cluster 2, meanwhile, consists of a small and less clearly defined group of countries (France, Ireland, Portugal, and Slovenia).

The analysis suggests two key points. First, a certain degree of differentiation can still be observed, although it does not align neatly with traditional welfare state models. Second, there appears to be similarities among Western European countries, as only France, Ireland, and Portugal do not fall within Cluster 1. To further investigate this – and considering that the Duda-Hart rule could also support two as the optimal number of clusters, as well as the fact that the Gap Statistic, computed for a range of 1 to 10 clusters with 300 iterations, also points to two clusters – the analysis is replicated using pooled data. In this latter case (not reported in the table), Cluster 2 disappears, with all countries falling into Cluster 1. This outcome supports the hypothesis of strong convergence across all European countries, with the only alternative recognisable cluster being that of the Eastern European countries.

The clustering presented in table 4 also reveals some interesting areas of overlap with the typology of "varieties of capitalism" proposed by Hall and Soskice (2001). Specifically, the cluster of "coordinated market economies" identified by Witt et al. (2018) – which includes Austria, Belgium, Denmark, Finland, Germany, the Netherlands, Norway, Sweden, and Switzerland – falls entirely within Cluster 1 of table 4. This cluster also includes some countries categorized as "European peripheral economies" (Greece, Italy, and Spain), a group which, in Witt et al. (2018:23), also includes several Eastern European nations. Similarly, Cluster 1 in table 4 overlaps with the "social democratic market economies" identified by Movahed (2023). In both cases, however, the clustering in table 4 is less sharply defined than the groupings reported in Witt et al. (2018) and Mohaved (2023).[31]

In order to verify the robustness of the previous results, the cluster analysis was replicated using a robust scaling procedure instead of a z-scoring technique, to reduce the potential influence of outliers on the clustering results. The robust scaling provides a new variable given by $x^{'} = \frac{x - median(x)}{IQR(x)}$, where $IQR(x) = p_{75}(x) - p_{25}(x)$ is the interquartile difference. In this case, the analysis identifies four as the optimal number of clusters. However, the differences from those reported in table 4 are negligible: in particular, Czechia moves from Cluster 3 to Cluster 1, and Ireland moves from Cluster 2 to form a single-member cluster. This provides reasonable confidence that the outcome does not depend critically on modelling choices. Furthermore, the same

[31] Note that Witt et al. (2018) use average linkage clustering, while Movahed (2023) employs a k-means method.

analysis was replicated with an alternative distance metric, namely the Manhattan (absolute) distance, in place of the squared Euclidean distance. While the overall cluster structure remained largely consistent, a few countries were reassigned under the alternative metric: Greece, Italy, Spain, and the UK moved to Cluster 2. This observation suggests that Clusters 1 and 2 may not be sharply separated and that certain countries lie near the boundaries between clusters. Such sensitivity to the choice of distance metric without altering the fundamental clustering outcome highlights a degree of overlap and underscores the importance of interpreting membership in Clusters 1 and 2 as indicative rather than definitive. By contrast, Cluster 3 remains unchanged, confirming the distinct profile of this group of countries in terms of health system performance.

All these results are based on the set of variables that maximizes the number of countries and observations in the dataset. Further analyses were conducted by progressively including additional variables, whose inclusion resulted in the loss of one or more countries. Several steps are considered:

a) Including long-term care beds, which results in the exclusion of Portugal. In this case, Clusters 1 and 2 from table 4 are slightly reconfigured, yielding a clearer Nordic-Continental model in Cluster 1, while Greece, Iceland, Italy, and Spain move to Cluster 2. The Eastern European model remains stable.
b) Adding congestive heart failure hospitalizations, which leads to the exclusion of Portugal, Greece, and Latvia. In this scenario, Austria, Germany, Spain, Switzerland, and the UK move to Cluster 2, producing a mixed and relatively undefined model. The Eastern European cluster remains stable.
c) Including the number of beds in publicly owned hospitals, which results in the exclusion of Denmark, Portugal, Greece, and Latvia. Compared to the clusters in table 4, only Switzerland moves to Cluster 2, while the Eastern European cluster persists.
d) Adding hospital admissions for diabetes, asthma, and chronic obstructive pulmonary disease results in the exclusion of Denmark, Portugal, Greece, Latvia, and Slovakia, and reveals a much stronger convergence in clustering. While the Eastern European cluster still persists, all other countries – with the exception of Ireland, which forms a single-member cluster – fall into Cluster 1.
e) The same pattern as in d) holds when out-of-pocket health expenditures are included, which leads to the loss of observations for Germany and Iceland.

In order to obtain more detailed information on whether performance similarities are driven by outcomes or by institutional characteristics, the cluster analysis was also replicated using the variables that best capture the structure of health systems. Specifically, the analysis focused on financing, coverage, and resources (i.e., health expenditures, health resources, and health protection variables in table 1). In this case (not reported in the table), using three clusters, a Nordic model – including Austria, Belgium, Denmark, Germany, the Netherlands, Norway, Sweden, and Switzerland – is clearly identified. A second cluster is composed of France, Ireland, Portugal, and Slovenia – as it was in Cluster 2 of table 4 – enlarged to include Greece and the United Kingdom, while the third cluster becomes a hybrid due to the presence of Italy, Luxembourg, and Spain. This implies that when focusing on institutional characteristics, the convergence of health models is weaker than the similarities observed for health outcomes. To further investigate this point, the Gap Statistic was recalculated on the new set of variables, yielding an optimal number of seven clusters. The results are reported in table 5, where Cluster 1 mostly identifies a Continental model, Cluster 2 a Nordic model, Cluster 3 (with the exception of Luxembourg) an Eastern European model, and Cluster 4 a subset of the Mediterranean model. The remaining three clusters, however, cannot be clearly classified according to standard definitions.

Table 5
Cluster composition

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Austria	Denmark	Czechia	Italy	France	Greece	Ireland
Belgium	Finland	Estonia	Spain	Portugal	UK	
Germany	Norway	Hungary		Slovenia		
Netherlands	Sweden	Latvia				
Switzerland		Lithuania				
		Luxembourg				
		Poland				
		Slovakia				

Note: Optimal number of clusters: 7 (Gap Statistics, Ward linkage).

In any case, what emerges from this sensitivity analysis is that similarities in outcomes do not necessarily entail convergence in health models. Put differently, this suggests that different institutional models may no longer produce distinct outcomes. As mentioned above, this finding supports the hypothesis that similar health outcomes can emerge from different institutional configurations, blurring the distinction between Beveridge and Bismarck models when moving from institutions to outcomes.

3.4 CHARACTERISTICS OF THE CLUSTERS

The analysis conducted so far shows that, if any clear partition exists, it is into two or three clusters. To comment on the specific characteristics of the larger clustering outcome, reference is made to the results in table 4, where a quasi-Nordic-Continental model is recognisable in Cluster 1, an Eastern European model is confirmed in Cluster 3, and a mixed group of countries belongs to Cluster 2. To meaningfully discuss the characteristics of these clusters, it is worth starting with the most persistent outcome – namely, the Eastern European model – and understanding the empirical evidence that explains why it forms a distinct cluster. Compared to the other two clusters, this group shows several distinctive indicators (see tables A5, A6, and A7 in appendix).

First, compared with Cluster 1, there is evidence of lower public health spending, with a marked upward trend observable only during the pandemic years. This lower spending level is associated with both a higher death rate and lower life expectancy at birth. Second, to some extent, a historically lower level of total health employment may have contributed to inadequate public health spending.[32]

Given that both the share of private health spending and the share of public pharmaceutical expenditures are also lower (H568 and H580), this group of countries can be clearly characterised by total health spending below the European average, which helps explain their distinct status in the cluster analysis. Furthermore, as additional indicators of divergence from the other two clusters, one can observe a slightly lower level of government compulsory health insurance (govins), suggesting relatively lower universal coverage; a shorter duration for which sickness benefits are payable (sduratio); and a longer average hospital stay (H540) – even when compared to the OECD average of 7.6 days. According to standard interpretation, longer hospital stays may reflect lower efficiency in bed management.

To some extent, compared with the other two clusters, there also appears to be greater difficulty in reducing the number of discharges (with the exception of 2020 and 2021), although this has no clear-cut interpretation. On the one hand, improved discharge rates can help free up hospital beds and staff time; on the other hand, premature discharges can worsen health outcomes and lead to costly readmissions. Regardless of interpretation, in 2019 – before the pandemic – the average number of discharges in Cluster 3 was about 185 per 1,000 inhabitants, compared to the OECD average of approximately 146.

Turning to Clusters 1 and 2, some differences can be observed. It is particularly useful to focus on Cluster 2 – a mixed cluster – and examine how it differs empirically from the more populous Cluster 1. In Cluster 2, there is a combination of variable public health expenditures, a mild upward trend in private health spending (H568), and a more pronounced downward trend in the share of public pharmaceutical expenditures (H580). Together, these three elements suggest a possible decline in overall healthcare expenditure, which – at least in some countries – supports the hypothesis that public health performances may be converging toward relatively lower standards.

It is also worth noting that, even during the pandemic, the average increase in health spending in Cluster 2 was lower than that observed in Cluster 1. This difference becomes more evident when considering a lower use of in-kind non-means-tested benefits (V1B), a lower level of discharges (disch) – an indirect efficiency indicator – a lower gross replacement rate for sickness benefits (srtsw26s), a significantly

[32] It is also worth noting that this shortage in Cluster 1 reflects a broader EU-wide issue. As reported by OECD (2024), the EU faces a health workforce deficit, with an estimated shortage of 1.2 million doctors, nurses, and midwives as of 2022. This shortage stems from multiple factors: demographic ageing affecting both patients and the health workforce, combined with difficult working conditions that contribute to staff burnout and retention challenges.

longer waiting period for payments (swaiting), a historically longer duration of payable sickness benefits (sduratio), and slightly lower coverage as a proportion of the labour force (scovratl).

It is also useful to examine the evolution of some key healthcare indicators across the three clusters. Public health spending has clearly grown in Cluster 1, helping maintain stability in private health expenditures, while private spending has increased in the other two clusters. Conversely, the share of public pharmaceutical expenditure in total health spending has decreased across all clusters. In Cluster 2, both the increase in private spending and the reduction in public pharmaceutical expenditures are particularly pronounced.

Furthermore, although not reported in the tables, two additional elements may support the idea of convergence: the number of beds in publicly owned hospitals (measured as days per person) has decreased across all clusters, while the increase in households' out-of-pocket payments is especially evident in Clusters 2 and 3.[33]

Regarding the robustness of these differences among clusters, all pairwise differences across the variables are statistically significant according to the Wilcoxon test, with one minor exception: the difference in health spending between Clusters 1 and 2, where the p-value exceeds the 5 percent significance level.

4 CONCLUSIONS

The main finding of this analysis is that the clustering of healthcare systems in Europe does not correspond to traditional welfare state classifications. The conventional distinctions between Nordic, Continental, and Liberal models receive little empirical support within the healthcare domain, whether in terms of system characteristics or health outcomes. Instead, most Western European countries cluster together, while a distinct and robust cluster of Eastern European countries emerges.

For Western Europe, this finding challenges the long-standing assumption that institutional legacies are the principal determinants of cross-national differences in healthcare performance (Pierson, 2004:17; Hacker, 1998; Taylor-Gooby, 1996; Wilsford, 1994). Rather, it lends support to the view of contemporary welfare states as "patchwork mixes of old and new policies and institutions" (Hemerijck, 2012:12) in which inherited arrangements coexist with newer policy instruments and governance logics.

Several factors may account for these patterns. First, the tendency of Western European countries to cluster together – despite heterogeneous institutional arrangements – may reflect their diminishing capacity to pursue autonomous healthcare policy paths.[34] As Ploug (1995) observed, policy debates have long revolved aro-

[33] Note that these indicators are not strictly comparable with the other data, as the number of countries for which data are available is smaller than in the general case.

[34] See also Rothgang et al. (2010:247).

und whether European states can maintain acceptable health standards while preserving universal access, or whether future trajectories point toward more restricted forms of universalism.

The absence of clear-cut distinctions across countries may also be linked to a broader shift toward market-compatible social policies, whereby welfare provision increasingly operates in accordance with market logics (Ferge, 1997). Since the early 1990s, this "rightward" shift has often been interpreted as a response to globalization and the lack of effective countervailing forces at the international level (Swank, 2002). In this context, Castles (1999) noted that whereas public healthcare expenditure had previously been driven primarily by political variables, cost-containment strategies have dominated since the 1980s, particularly across Europe during the 1990s.

Although differences in financing structures persist, convergence may be fostered by the introduction of market-based solutions in Beveridge-type systems, alongside efforts within Bismarckian systems to reduce inequalities in access to essential care. This has generated increasingly hybrid health systems and, ultimately, convergent health outcomes. However, the implications of these developments remain contested. Extensive welfare provision risks losing political legitimacy as an electoral platform, potentially encouraging more restricted forms of universalism and the adoption of minimum, rather than socially acceptable, standards.[35] Others caution that reductions in public intervention may facilitate the expansion of private insurance, thereby undermining the principle of universality and weakening collective protection against shared risks (Sunstein, 1993; Losada and Ares, 2021).

Finally, while healthcare reforms continue to be shaped by country-specific factors, they are increasingly influenced by processes of policy diffusion. As Panic (2003:74) observes, national stability and long-term progress often depend on policy developments in closely connected countries. This dynamic is particularly salient within an integrated European economic system, where states face similar fiscal and political constraints, even as a coherent "European social space" has yet to fully materialize (Alsasua, Bilbao-Ubillos and Olaskoaga, 2007:297). In this context, policy diffusion may represent an additional mechanism contributing to the clustering patterns identified in this analysis.

These findings carry important implications for both research and policy. The weak correspondence between healthcare clustering and traditional welfare typologies highlights the need to move beyond rigid classifications when analysing cross-national health outcomes. Future research could examine how hybrid health systems emerge, identifying which combinations of policies contribute to similarities in outcomes, and investigate the role of policy diffusion in shaping national reforms

[35] For a different perspective, see Goodin and Le Grand (1987).

within a shared European context. Researchers might also assess whether observed similarities in health outcomes mask persistent inequalities within populations.

For policymakers, the results suggest that institutional heritage alone does not determine performance. Recognizing emerging clusters of countries may help governments design adaptable reforms that balance universality, financial sustainability, and quality of care. In light of rising healthcare needs, indeed, the role of public intervention is likely to become increasingly complex, underscoring the importance of defining the appropriate level of public involvement needed to address these demands.

Disclosure statement

The author has no conflicts of interest to declare.

REFERENCES

1. Alber, J., 1995. A framework for the comparative study of social services. *Journal of European Social Policy*, 5(2), pp. 131-149. https://doi.org/10.1177/095892879500500204
2. Alsasua, J., Bilbao-Ubillos, J. and Olaskoaga, J., 2007. The EU integration process and the convergence of social protection benefits at national level. *International Journal of Social Welfare*, 16(4), pp. 297-306. https://doi.org/10.1111/j.1468-2397.2007.00483.x
3. Arts, W. and Gelissen, J., 2002. Three worlds of welfare capitalism or more? A state-of-the-art report. *Journal of European Social Policy*, 12(2), pp. 137-158. https://doi.org/10.1177/0952872002012002114
4. Bambra, C., 2005a. Worlds of Welfare and the Health Care Discrepancy. *Social Policy and Society*, 4(1), pp. 31-41. https://doi.org/10.1017/S1474746404002143
5. Bambra, C., 2005b. Health Status and the Worlds of Welfare. *Social Policy and Society*, 5(1), pp. 53-62. https://doi.org/10.1017/S1474746405002721
6. Barbier, J., 2012. Tracing the fate of EU "social policy": Changes in political discourse from the 'Lisbon Strategy' to Europe 2020. *International Labour Review*, 151(4), pp. 377-399. https://doi.org/10.1111/j.1564-913X.2012.00154.x
7. Barros, P. P., 2007. The slow and unnoticed changes in the funding mix. *Health Economics*, 16, pp. 437-440. https://doi.org/10.1002/hec.1241
8. Bergqvist, K., Yngwe, M. Å. and Lundberg, O., 2013. Understanding the role of welfare state characteristics for health and inequalities – An analytical review. *BMC Public Health*, 13, 1234, pp. 1-20. https://doi.org/10.1186/1471-2458-13-1234
9. Bertin, G., Carrino, L. and Pantalone, M., 2021. Do standard classifications still represent European welfare typologies? Novel evidence from studies on health and social care. *Social Science and Medicine*, 281, 114086. https://doi.org/10.1016/j.socscimed.2021.114086
10. Bilbao-Ubillos, J., 2023. The Social Dimension of the European Union: A Means to lock out Social Competition? *Social Indicators Research*, 165, pp. 267-281. https://doi.org/10.1007/s11205-022-03012-6
11. Böhm, K. [et al.], 2013. Five types of OECD healthcare systems: Empirical results of a deductive classification. *Health Policy*, 113(3), pp. 258-269. https://doi.org/10.1016/j.healthpol.2013.09.003
12. Braithwaite, J., Travaglia, J. F. and Corbett, A., 2011. Can Questions of the Privatization and Corporatization, and the Autonomy and Accountability of Public Hospitals, Ever be Resolved? *Health Care Analysis*, 19, pp. 133-153. https://doi.org/10.1007/s10728-010-0152-x
13. Breyer, F. and Lorenz, N., 2020. The "red herring" after 20 years: ageing and health care expenditures. *The European Journal of Health Economics*, 22, pp. 661-667. https://doi.org/10.1007/s10198-020-01203-x
14. Brick, A. and Connolly, S., 2021. Waiting Times for Publicly Funded Hospital Treatment: How does Ireland Measure Up? *The Economic and Social Review*, 52(1), pp. 41-52.

15. Burau, V. and Blank, R. H., 2006. Comparing Health Policy: An Assessment of Typologies of Health Systems. *Journal of Comparative Policy Analysis*, 8(1), pp. 63-76. https://doi.org/10.1080/13876980500513558
16. Burke, S. [et al.], 2018. Sláintecare – A ten-year plan to achieve universal healthcare in Ireland. *Health Policy*, 122(12), pp. 1278-1282. https://doi.org/10.1016/j.healthpol.2018.05.006
17. Calinski, R. B. and Harabasz, J., 1974. A dendrite method for cluster analysis. *Communications in Statistics*, 3(1), pp. 1-27. https://doi.org/10.1080/03610927408827101
18. Carreira Da Silva, F., 2017. Welfare State. In: W. Outwhaite and S. Turner, eds. *The SAGE Handbook of Political Sociology*. https://doi.org/10.17863/CAM.8022
19. Castles, F. G. and Obinger, H., 2008. Worlds, Families, Regimes: Country Clusters in European and OECD Area Public Policy. *West European Politics*, 31(1-2), pp. 321-344. https://doi.org/10.1080/01402380701835140
20. Castles, F. G. (ed.), 2007. *The Disappearing State? Retrenchment Realities in an Age of Globalisation*. Cheltenham: Edward Elgar. https://doi.org/10.4337/9781847205292
21. Castles, F. G., 2002. Developing new measures of welfare state change and reform. *European Journal of Political Research*, 41(5), pp. 613-641. https://doi.org/10.1111/1475-6765.00024
22. Castles, F. G., 1999. *Comparative Public Policy Patterns of Post-War Transformation*. Cheltenham: Edward Elgar.
23. Cerami, A., 2010. The Politics of Social Security Reforms in the Czech Republic, Hungary, Poland and Slovakia. In: B. Palier, ed. *A Long Goodbye to Bismarck?: The Politics of Welfare Reform in Continental Europe*. Amsterdam: Amsterdam University Press, pp. 233-254. https://doi.org/10.1017/9789048512454.010
24. Daly, M., 1997. Welfare states under pressure: cash benefits in European welfare states over the last ten years. *Journal of European Social Policy*, 7(2), pp. 129-146. https://doi.org/10.1177/095892879700700204
25. Danforth, B., 2014. Worlds of welfare in time: A historical reassessment of the three-world typology. *Journal of European Social Policy*, 24(2), pp. 164-182. https://doi.org/10.1177/0958928713517919
26. De Simone, E., Gaeta, G. L. and Ercolano, S., 2012. Exploring Public Social Expenditure Trends in the Globalization Era. *European Research Studies*, 15(1), pp. 23-42. https://doi.org/10.35808/ersj/342
27. Den Teuling, N. G. P., Pauws, S. C. and van den Heuvel, E. R., 2021. A comparison of methods for clustering longitudinal data with slowly changing trends. *Communications in Statistics – Simulation and Computation*, 52(3), pp. 621-648. https://doi.org/10.1080/03610918.2020.1861464
28. Dubas-Jakóbczyk, K. [et al.], 2020. Hospital reforms in 11 Central and Eastern European countries between 2008 and 2019: a comparative analysis. *Health Policy*, 124(4), pp. 368-379. https://doi.org/10.1016/j.healthpol.2020.02.003

29. Duda, R. O. and Hart, P. E., 1973. *Pattern Classification and Scene Analysis.* New York: John Wiley and Sons.
30. Emmenegger, P. [et al.], 2015. Three Worlds of Welfare Capitalism: The making of a classic. *Journal of European Social Policy*, 25(1), pp. 3-13. https://doi.org/10.1177/0958928714556966
31. Ermakoff, I., 2019. Causality and history: Modes of causal investigation in historical social sciences. *Annual Review of Sociology*, 45(1), pp. 581-606. https://doi.org/10.1146/annurev-soc-073117-041140
32. Esping-Andersen, G., 1990. *The Three Worlds of Welfare Capitalism*. Princeton: Princeton University Press.
33. Everitt, B. S. [et al.], 2011. *Cluster Analysis*. 5th ed. London: Wiley. https://doi.org/10.1002/9780470977811
34. Ferge, Z., 1997. The Changed Welfare Paradigm: The Individualization of the Social. *Social Policy and Administration*, 31(1), pp. 20-44. https://doi.org/10.1111/1467-9515.00035
35. Ferragina, E. and Seeleib-Kaiser, M., 2011. Welfare regime debate: past, present, futures? *Policy and Politics*, 39(4), pp. 583-611. https://doi.org/10.1332/030557311X603592
36. Filipovic, H. and Dobrotic, I., 2022. Eastern European Welfare States. In: B. Greve, ed. *Handbook of Contemporary Welfare States*. Berlin: De Gruyter, pp. 119-134. https://doi.org/10.1515/9783110721768-008
37. Fosse, E., 2011. Different welfare states different policies? An analysis of the substance of national health promotion policies in three European countries. *International Journal of Health Services*, 41(2), pp. 255-273. https://doi.org/10.2190/HS.41.2.e
38. Freeman, R., 1999. Institutions, states and cultures: health policy and politics in Europe. In: J. Clasen, ed. *Comparative Social Policy*. Oxford: Blackwell.
39. Giaimo, S. and Manow, P., 1999. Adapting the Welfare State – The case of Health Care Reform in Britain, Germany, and the United States. *Comparative Political Studies*, 32(8), pp. 967-1000. https://doi.org/10.1177/0010414099032008003
40. Gilardi, F., Füglister, K. and Luyet, S., 2009. Learning From Others – The Diffusion of Hospital Financing Reforms in OECD countries. *Comparative Political Studies*, 42(4), pp. 549-573. https://doi.org/10.1177/0010414008327428
41. Giuliani, A. and Vici, A., 2024. On the (Apparently) Paradoxical Role of Noise in the Recognition of Signal Character of Minor Principal Components. *Stats*, 7(1), pp. 54-64. https://doi.org/10.3390/stats7010004
42. Goodin, R. E. and Le Grand, J., 1987. *Not Only the Poor – The Middle Classes and the Welfare State.* London: Routledge.
43. Gough, I., 2001. Social assistance regimes: a cluster analysis. *Journal of European Social Policy*, 11(2), pp. 165-170. https://doi.org/10.1177/095892870101100205

44. Hacker, J. S., 1998. The historical logic of national health insurance: structure and sequence in the development of British, Canadian, and U.S. medical policy. *Studies in American Political Development*, 12(Spring), pp. 57-130.
45. Hall, P. A. and Soskice, D. (ed.), 2001. *Varieties of capitalism: the institutional foundations of comparative advantage*. Oxford: Oxford Academic Press. https://doi.org/10.1093/0199247757.001.0001
46. Hemerijck, A., 2012. When Changing Welfare States and the Eurocrisis Meet. *Sociologica*, 1(1), pp. 1-58.
47. Hermann, C., 2010. The Marketisation of Health Care in Europe. *Socialist Register*, 46, pp. 125-144.
48. Higgins, J., 1981. *States of Welfare – Comparative Analysis in Social Policy.* Oxford: Basil Blackwell; Martin Robertson.
49. Huber, E. and Stephens, J., 2001. *Development and Crisis of the Welfare State: Parties and Policies in Global Market*s. Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226356495.001.0001
50. Jensen, C., 2008. Worlds of welfare services and transfers. *Journal of European Social Policy*, 18(2), pp. 151-162. https://doi.org/10.1177/0958928707087591
51. Joumard, I., André, C. and Nicq, C., 2010. Health care systems: efficiency and institutions. *Economics Department Working Papers,* No. 769. https://doi.org/10.2139/ssrn.1616546
52. Kangas, O., 1994. The politics of social security: on regressions, qualitative comparisons, and cluster analysis. In: T. Janoski and A. M. Hicks. *The Comparative Political Economy of the Welfare State.* Cambridge: Cambridge University Press, pp. 346-364.
53. Kasza, G. J., 2002. The Illusion of Welfare ‘Regimes’. *Journal of Social Policy*, 31(2), pp. 271-287. https://doi.org/10.1017/S0047279401006584
54. Kaufman, L. and Rousseeuw, P. J., 1990. *Finding Groups in Data. An Introduction to Cluster Analysis.* New York: John Wiley and Sons. https://doi.org/10.1002/9780470316801
55. Kautto, M., 2002. Investing in services in West European welfare states. *Journal of European Social Policy*, 12(1), pp. 53-65. https://doi.org/10.1177/0952872002012001636
56. Kringos, D. S. [et al.], 2013. The strength of primary care in Europe: an international comparative study. *British Journal of General Practice*, 63(616), e742-e750. https://doi.org/10.3399/bjgp13X674422
57. Kringos, D. S. [et al.], eds., 2015. Building primary care in a changing Europe. *Observatory Studies Series*, No. 38.
58. Kuitto, K., 2011. More than just money: Patterns of disaggregated welfare expenditure in the enlarged Europe. *Journal of European Social Policy*, 21(4), pp. 348-364. https://doi.org/10.1177/0958928711412223
59. Kuitto, K., 2016. *Post-Communist welfare states in European context: Patterns of welfare policies in Central and Eastern Europe*. Cheltenham: Edward Elgar.

60. Lamping, W. and Rüb, F. W., 2004. From the Conservative Welfare State to an Uncertain Else: German Pension Politics in Comparative Perspective. *Policy and Politics*, 32(2), pp. 169-191. https://doi.org/10.1332/030557304773558134
61. Lee, J. Y. [et al.], 2016. Trajectories of cigarette smoking beginning in adolescence predict insomnia in the mid thirties. *Substance Use and Misuse*, 51(5), pp. 616-624. https://doi.org/10.3109/10826084.2015.1126747
62. Leiter, A. M. and Theurl, E., 2012. The convergence of health care financing structures: empirical evidence from OECD-countries. *The European Journal of Health Economics*, 13(1), pp. 7-18. https://doi.org/10.1007/s10198-010-0265-z
63. Losada, A. and Ares, C., 2021. The resistance of the four European worlds of welfare from the births of the euro. *Brazilian Journal of Political Economy*, 41(2), pp. 271-291. https://doi.org/10.1590/0101-31572021-3066
64. Lu, Z., 2024. Clustering Longitudinal Data: A Review of Methods and Software Packages. *International Statistical Review*, 93(3), pp. 425-458. https://doi.org/10.1111/insr.12588
65. Lundberg, O. [et al.], 2015. Welfare State and Health Inequalities. *Canadian Public Policy*, 41(Sip. 2). https://doi.org/10.3138/cpp.2014-079
66. McGeady, J. [et al.], 2024. Social Justice Matters – 2024 guide to a fairer society. *Social Justice Ireland*, April 29, 2024.
67. Milligan, G. W. and Cooper, M. C., 1985. An Examination of Procedures for Determining the Number of Clusters in a Data Set. *Psychometrika*, 50(2), pp. 159-179. https://doi.org/10.1007/BF02294245
68. Minas, C., Jacobson, D. and McMullan, C., 2014. Welfare regime, welfare pillar and Southern Europe. *Journal of European Social Policy*, 24(2), pp. 135-149. https://doi.org/10.1177/0958928713517917
69. Moran, M., 2000. Understanding the Welfare State: The Case of Health Care. *British Journal of Politics and International Relations*, 2(2), pp. 135-160. https://doi.org/10.1111/1467-856X.00031
70. Movahed, M., 2023. Varieties of capitalism and income inequality. *International Journal of Comparative Sociology*, 64(6), pp. 621-657. https://doi.org/10.1177/00207152231174
71. Musgrove, P., 2003. Judging health systems: reflections on WHO's methods. *The Lancet*, 361(9371), pp. 1817-1820. https://doi.org/10.1016/S0140-6736(03)13408-3
72. Obinger, H. and Wagschal, U., 2001. Families of nations and public policy. *West European Politics*, 24(1), pp. 99-114. https://doi.org/10.1080/01402380108425419
73. OECD, 2018. *Health at a glance: Europe 2018: State of health in the EU cycle*. Paris: OECD. https://doi.org/10.1787/health_glance_eur-2018-en
74. OECD, 2025a. *Health at a Glance 2025*. Paris: OECD. https://doi.org/10.1787/8f9e3f98-en
75. OECD, 2025b. *How Do Health System Features Influence Health System Performance?* Paris: OECD. https://doi.org/10.1787/7b877762-en

76. Palier, B., 2010. Ordering Change: Understanding the 'Bismarckian' Welfare Reform Trajectory. In: B. Palier, ed. *A Long Goodbye to Bismarck?* Amsterdam: Amsterdam University Press, pp. 19-44. https://doi.org/10.1017/9789048512454.002
77. Panic, M., 2003. *Globalization and national economic welfare*. London: Palgrave; MacMillan. https://doi.org/10.1057/9780230512481
78. Pestieau, P. and Lefebvre, M., 2018. *The Welfare State in Europe: Economic and Social Perspectives*. 2nd ed. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198817055.001.0001
79. Pierson, C., 2006. *Beyond the Welfare State – The new political economy of welfare*. Cambridge: Polity Press.
80. Pierson, P., 2004. *Politics in Time – History, Institutions, and Social Analysis*. Princeton and Oxford: Princeton University Press. https://doi.org/10.1515/9781400841080
81. Pinker, R., 1971. *Social Theory and Social Policy.* London: Heinemann.
82. Ploug, N., 1995. The welfare state in liquidation? *International Social Security Review*, 48(2), pp. 61-71. https://doi.org/10.1111/j.1468-246X.1995.tb00430.x
83. Polin, K. [et al.], 2021. "Top-Three" health reforms in 31 high-income countries in 2018 and 2019: an expert informed overview. *Health Policy*, 125(7), pp. 815-832. https://doi.org/10.1016/j.healthpol.2021.04.005
84. Powell, M. and Barrientos, A., 2004. Welfare regimes and the welfare mix. *European Journal of Political Research*, 43(1), pp. 83-105. https://doi.org/10.1111/j.1475-6765.2004.00146.x
85. Powell, M., Yörük, E. and Bargu, A., 2019. Thirty years of the Three Worlds of Welfare Capitalism: A review of reviews. *Social Policy Administration*, 54(1), pp. 60-87. https://doi.org/10.1111/spol.12510
86. Ranci, C. and Pavolini, E., 2015. Not all that glitters is gold: Long-term care reforms in the last two decades in Europe. *Journal of European Social Policy*, 25(3), pp. 270-285. https://doi.org/10.1177/0958928715588704
87. Reibling, N., 2010. Healthcare systems in Europe: towards an incorporation of patient access. *Journal of European Social Policy*, 20(1), pp. 5-18. https://doi.org/10.1177/0958928709352406
88. Reibling, N., Ariaans, M. and Wendt, C., 2019. Worlds of Healthcare: A Healthcare System Typology of OECD Countries. *Health Policy*, 123(7), pp. 611-620. https://doi.org/10.1016/j.healthpol.2019.05.001
89. Rothgang, H. [et al.], 2010. *The State and Healthcare: Comparing OECD Countries*. London: Palgrave; MacMillan. https://doi.org/10.1057/9780230292345
90. Rothgang, H., 2021. Health. In: D. Béland [et al.], eds. *The Oxford Handbook of the Welfare State.* 2nd ed. Oxford: Oxford University Press, pp. 506-523. https://doi.org/10.1093/oxfordhb/9780198828389.013.29
91. Ruggles, P. and O'Higgins, M., 1987. Retrenchment and the New Right: A Comparative Analysis of the Impacts of the Thatcher and Reagan Administration. In: M. Rein, G. Esping-Andersen and L. Rainwater, eds. *Stagnation and Renewal in Social Policy – The Rise and Fall of Policy Regimes*. London: M.E. Sharpe Inc., pp. 160-190.

92. Saint-Arnaud, S. and Bernard, P., 2003. Convergence or resilience? A hierarchial cluster analysis of the welfare regimes in advanced countries. *Current Sociology*, 51(5), pp. 499-527. https://doi.org/10.1177/00113921030515004
93. Shalev, M., 2007. Limits and alternatives to multiple regression in comparative research. *Comparative Social Research*, 24, pp. 261-308.
94. Sorenson, C., Drummond, M. and Kahn, B. B., 2013. Medical technology as a key driver of rising health expenditures: disentangling the relationship. *ClinicoEconomics and Outcomes Research,* 5, pp. 223-234. https://doi.org/10.2147/CEOR.S39634
95. Sowula, J. [et al.], 1993. Against Positive Rights Feature. *East European Constitutional review*, 35, pp. 35-38.
96. Sowula, J. [et al.], 2023. The end of welfare states as we know them? A multidimensional perspective. *Social Policy Administration*, 58(5), pp. 785-799. https://doi.org/10.1111/spol.12990
97. Sunstein, C. R., 1993. Against positive rights feature. *East European Constitutional Review,* 2, pp. 35-38.
98. Swank, D., 2002. *Global Capital, Political Institutions, and Policy Change in Developed Welfare States*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511613371
99. Szebehely, M. and Meagher, G., 2017. Nordic eldercare – Weak universalism becoming weaker? *Journal of European Social Policy*, 28(3), pp. 294-308. https://doi.org/10.1177/0958928717735062
100. Taylor Brown, C. and Ben Brik, A., 2024. Welfare regime typologies: The six worlds of social inclusion. *Journal of International and Comparative Social Policy*, 40(1), pp. 1-23. https://doi.org/10.1017/ics.2024.14
101. Taylor-Gooby, P., 1996. The Future of Health Care in Six European Countries: The Views of Policy Elites. *International Journal of Health Policy*, 26(2), pp. 203-219. https://doi.org/10.2190/BC70-6QY7-TAKH-NGT0
102. Taylor-Gooby, P., Leruth, B. and Chung, H. (eds.), 2017. *After Austerity: Welfare State Transformation in Europe after the Great Recession*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198790266.001.0001
103. The Lancet Regional Health – Europe, 2025. Ireland: Europe's outlier in primary health care. *The Lancet Regional Health – Europe*, 50, 101253. https://doi.org/10.1016/j.lanepe.2025.101253
104. Thomas, S. [et al.], 2021. Sláintecare implementation status in 2020: Limited progress with entitlement expansion. *Health Policy*, 125(3), pp. 277-283. https://doi.org/10.1016/j.healthpol.2021.01.009
105. Tiemann, O., Schreyogg, J. and Busse, R., 2012. Hospital ownership and efficiency: A review of studies with particular focus on Germany. *Health Policy*, 104(2), pp. 163-171. https://doi.org/10.1016/j.healthpol.2011.11.010
106. Tine, R. [et al.], 2022. Revisiting the Nordic long-term care model for older people - still equal? *European Journal of Ageing*, 19, pp. 201-210. https://doi.org/10.1007/s10433-022-00703-4

107. Titmuss, R. M., 1974. *Social Policy*. London: Allen and Unwin.
108. Vaughan-Whitehead, D., 2017. The European Social Model in times of crisis: An overview. In: D. Vaughan-Whitehead, ed. *The European Social Model in Crisis – Is Europe Losing Its Soul?* Cheltenham: Edward Elgar, pp. 1-65.
109. Walsh, B. [et al.], 2022. The impact of inpatient bed capacity on length of stay. *The European Journal of Health Economics*, 23, 499-510. https://doi.org/10.1007/s10198-021-01373-2
110. Wendt, C. and Kohl, J., 2010. Translating Monetary Inputs into Health Care Provision: A Comparative Analysis of the Impact of Different Modes of Public Policy. *Journal of Comparative Policy Analysis: Research and Practice*, 12(1), pp. 11-31. https://doi.org/10.1080/13876980903076161
111. Wendt, C., 2009. Mapping European healthcare systems: a comparative analysis of financing, service provision and access to healthcare. *Journal of European Social Policy*, 19(5), pp. 432-445. https://doi.org/10.1177/0958928709344247
112. Wendt, C., 2014. Changing Healthcare System Types. *Social Policy and Administration*, 48(7), pp. 864-882. https://doi.org/10.1111/spol.12061
113. Weybright, E. H. [et al.], 2016. Trajectories of adolescent substance use development and the influence of healthy leisure: A growth mixture modeling approach. *Journal of Adolescence,* 49(1), pp. 158-69. https://doi.org/10.1016/j.adolescence.2016.03.012
114. WHO, 2000. *The World Health Report 2000 – Health systems: Improving performance*. Geneva: World Health Organization.
115. Whyte, R., Connolly, S. and Wren, M., 2020. Insurance status and waiting times for hospital-based services in Ireland. *Health Policy*, 124(11), pp. 1174-1181. https://doi.org/10.1016/j.healthpol.2020.07.001
116. Wilensky, H. and Lebeaux, C., 1965. *Industrial Society and Social Welfare*. New York: Free Press.
117. Wilsford, D., 1994. Path Dependency, or Why History Makes It Difficult but Not Impossible to Reform Health Care Systems in a Big Way. *Journal of Public Policy*, 14(3), pp. 251-283. https://doi.org/10.1017/S0143814X00007285
118. Witt, M. [et al.], 2018. Mapping the business systems of 61 major economies: a taxonomy and implications for varieties of capitalism and business systems research. *Socio-Economic Review*, 16(1), pp. 5-38. https://doi.org/10.1093/ser/mwx012
119. Zhou, J., Zhang, Y. and Tu, W., 2023. clusterMLD: An efficient hierarchical clustering method for multivariate longitudinal data. *Journal of Computational and Graphical Statistics*, 32(3), pp. 1131-1144. https://doi.org/10.1080/10618600.2022.2149540
120. Zweifel, P., Steinmann, L. and Eugster, P., 2005. The Sisyphus Syndrome in Health Revisited. *International Journal of Health Care Finance and Economies*, 5(2), pp. 127-145. https://doi.org/10.1007/s10754-005-1864-6

APPENDIX

Table A1

Cluster composition (stopping at 3)

Cluster 1	Cluster 2	Cluster 3
Austria	Iceland	Netherlands (5)
Belgium		
Czechia		
Denmark		
Estonia		
France		
Finland		
Germany		
Greece		
Hungary		
Ireland		
Italy		
Latvia		
Lithuania		
Luxembourg		
Netherlands (16)		
Norway		
Poland		
Portugal		
Slovakia		
Slovenia		
Spain		
Sweden		
Switzerland		
United Kingdom		

Note: Average linkage – Euclidean distance.

Table A2

Cluster composition (stopping at 4)

Cluster 1	Cluster 2	Cluster 3	Cluster 4
Austria	Iceland	Ireland	Netherlands (5)
Belgium			
Czechia			
Denmark			
Estonia			
France			
Finland			
Germany			
Greece			
Hungary			
Italy			
Latvia			
Lithuania			
Luxembourg			
Netherlands (16)			
Norway			
Poland			
Portugal			
Slovakia			
Slovenia			
Spain			
Sweden			
Switzerland			
United Kingdom			

Note: Average linkage – Euclidean distance.

PUBLIC SECTOR ECONOMICS
50 (2) 167-213 (2026)

Table A3

Cluster composition (stopping at 5)

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Austria	Sweden (6)	Iceland	Ireland	Netherlands (5)
Belgium				
Czechia				
Denmark				
Estonia				
Finland				
France				
Germany				
Greece				
Hungary				
Italy				
Latvia				
Lithuania				
Luxembourg				
Netherlands (16)				
Norway				
Poland				
Portugal				
Slovakia				
Slovenia				
Spain				
Sweden				
Switzerland				
United Kingdom				

Note: Average linkage – Euclidean distance.

Table A4

Cluster composition (stopping at 6)

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Austria	Switzerland	Sweden (6)	Iceland	Ireland	Netherlands (5)
Belgium					
Czechia					
Denmark					
Estonia					
Finland					
France					
Germany					
Greece					
Hungary					
Italy					
Latvia					
Lithuania					
Luxembourg					
Netherlands (16)					
Norway					
Poland					
Portugal					
Slovakia					
Slovenia					
Spain					
Sweden (15)					
United Kingdom					

Note: Average linkage – Euclidean distance.

Table A5
Cluster 1: Average values

Year	health	old pop	death rate	life exp	V1A	V1B	V1C	V1D	tot emp	disch	gov ins	srtsw26s
2001	8.2	16.1	9.6	78.8	0.17	0.83	0	0	49.9	16776	96.0	0.58
2002	8.5	16.2	9.7	78.9	0.17	0.83	0	0	50.9	16721	95.7	0.58
2003	8.8	16.3	9.8	79.0	0.16	0.84	0	0	51.8	16720	95.7	0.58
2004	8.8	16.4	9.3	79.5	0.15	0.84	0	0	52.5	16868	95.7	0.58
2005	8.9	16.3	9.2	79.8	0.15	0.84	0	0	53.9	16881	95.9	0.58
2006	8.8	16.8	9.3	80.0	0.14	0.85	0	0	54.0	16898	98.7	0.59
2007	8.8	17.0	9.3	80.1	0.14	0.85	0	0.01	54.7	16968	98.7	0.58
2008	9.0	16.9	9.3	80.4	0.15	0.85	0	0	59.3	17224	98.9	0.62
2009	9.8	17.2	9.2	80.6	0.15	0.85	0	0	60.2	17379	99.7	0.61
2010	9.4	16.8	8.9	80.9	0.15	0.85	0	0	61.4	16988	99.8	0.63
2011	9.8	17.6	9.1	81.1	0.15	0.85	0	0	61.9	17211	99.8	0.61
2012	9.9	17.9	9.4	81.1	0.14	0.85	0	0	62.2	17106	99.8	0.61
2013	9.9	18.3	9.3	81.4	0.14	0.85	0	0	62.7	16557	99.8	0.61
2014	9.9	18.6	9.1	81.7	0.14	0.85	0	0	62.9	16509	99.8	0.61
2015	9.6	18.2	9.2	81.6	0.15	0.85	0	0	64.1	15756	99.8	0.63
2016	10	19.1	9.3	81.8	0.14	0.85	0	0	64.3	16066	99.8	0.61
2017	9.9	19.4	9.5	81.9	0.14	0.85	0	0	64.9	15894	99.8	0.61
2018	9.9	19.6	9.5	82.0	0.14	0.85	0	0	65.8	15820	99.9	0.61
2019	10.1	19.8	9.4	82.3	0.14	0.85	0	0	66.6	15683	99.9	0.60
2020	10.6	19.3	9.8	81.8	0.15	0.84	0	0	68.1	13576	99.9	0.62
2021	11	20.3	10.1	82.0	0.14	0.85	0	0	69.8	14311	99.9	0.60

Cluster 1: Average values (continued)

Year	swaiting	sduratio	scovratl	H568	H580	H540	fed	gov	vturn	rae	pres
2001	1.2	92.8	0.85	2.4	62.5	9.5	0.75	3.3	74.5	0.77	0.33
2002	1.2	93.7	0.86	2.5	63.7	9.4	0.75	2.8	75	0.77	0.33
2003	1.2	94.5	0.86	2.5	63.8	9.3	0.75	2.8	75.4	0.76	0.33
2004	1.2	95.4	0.87	2.5	64	9	0.75	2.8	76.9	0.76	0.33
2005	2.2	92.8	0.88	2.4	62.8	8.6	0.69	2.5	77.6	0.77	0.38
2006	1.2	88.5	0.87	2.3	67	8.7	0.75	2.7	76.7	0.77	0.33
2007	1.2	80.7	0.87	2.3	67.1	8.6	0.75	2.7	76.4	0.78	0.33
2008	1.1	71.2	0.88	2.2	65.9	8.4	0.69	2.6	76.2	0.78	0.31
2009	1.1	64	0.88	2.3	65.8	8.2	0.69	2.6	75.3	0.78	0.31
2010	1.9	56.2	0.88	2.1	65.1	7.9	0.6	2.5	76.9	0.78	0.33
2011	1.1	57.8	0.87	2.3	64.2	7.8	0.69	2.2	75.1	0.79	0.31
2012	1.1	58.8	0.87	2.3	64.1	7.7	0.69	2.2	73.5	0.8	0.31
2013	1.1	59.7	0.87	2.3	62	7.6	0.69	2.4	72.9	0.81	0.31
2014	1.1	60.7	0.87	2.3	61.3	7.5	0.69	2.4	73	0.81	0.31
2015	1.9	60.3	0.88	2.2	60.5	7.4	0.6	2.5	74.6	0.81	0.33
2016	1.1	61.7	0.87	2.3	60.8	7.3	0.69	2.3	72.2	0.81	0.31
2017	1.1	61.7	0.87	2.3	60.1	7.2	0.69	2.3	73.8	0.81	0.31
2018	1	61.8	0.87	2.2	59.6	7.1	0.69	2.1	73.7	0.81	0.31
2019	1	61.9	0.86	2.2	58.7	7.1	0.69	2.4	73.3	0.82	0.31
2020	1.7	62.3	0.87	2.1	57.1	7.1	0.6	2.6	75	0.82	0.33
2021	0.9	62	0.86	2.2	57.1	6.9	0.69	2.7	73.1	0.82	0.31

Table A6
Cluster 2: Average values

Year	health	old pop	death rate	life exp	V1A	V1B	V1C	V1D	tot emp	disch	gov ins	srtsw26s
2001	8.2	14.5	9	77.5	0.02	0.19	0	0.03	34.5	15997	99.5	0.35
2002	8.4	14.6	9	77.8	0.02	0.19	0	0.04	35.6	15776	99.5	0.36
2003	8.6	14.8	9.1	77.9	0.02	0.19	0.01	0.04	36.4	15294	99.7	0.36
2004	8.7	14.9	8.6	78.7	0.02	0.19	0.01	0.04	37.1	15300	99.7	0.36
2005	8.9	15.1	8.7	78.8	0.02	0.19	0.01	0.04	38	15317	99.7	0.37
2006	8.8	15.1	8.5	79.3	0.02	0.19	0.01	0.04	39	15435	99.7	0.41
2007	8.7	15.2	8.4	79.6	0.02	0.18	0.01	0.04	39.7	15489	99.7	0.46
2008	9.3	15.4	8.5	79.9	0.02	0.18	0.01	0.04	40.5	15476	100	0.5
2009	10.1	15.6	8.5	80.1	0.02	0.18	0.01	0.04	41.3	15264	100	0.55
2010	10.1	15.8	8.4	80.4	0.03	0.17	0.01	0.04	42.6	15158	100	0.59
2011	10	16	8.4	80.7	0.02	0.18	0.01	0.04	43.2	15076	100	0.59
2012	10.1	16.4	8.7	80.8	0.02	0.18	0.01	0.04	43.5	15011	100	0.58
2013	10	16.7	8.7	81	0.02	0.18	0.01	0.04	43.8	15367	100	0.58
2014	9.7	17.1	8.5	81.5	0.02	0.18	0.01	0.04	44.3	15401	100	0.57
2015	9.2	17.5	8.9	81.4	0.02	0.18	0.01	0.04	44.9	15311	100	0.57
2016	9.2	17.9	8.9	81.6	0.02	0.18	0.01	0.04	45.6	15279	100	0.56
2017	9	18.3	9	81.7	0.02	0.19	0.01	0.03	46.4	15032	100	0.56
2018	8.9	18.7	9.1	81.9	0.02	0.19	0.01	0.03	46.8	14933	100	0.56
2019	9	19.1	9.1	82.1	0.02	0.19	0.01	0.03	47.2	14755	100	0.56
2020	9.8	19.4	10	81.5	0.02	0.19	0.01	0.03	47.5	13080	100	0.56
2021	9.9	19.8	9.9	81.6	0.02	0.2	0.01	0.03	49.2	13780	100	0.56

Cluster 2: Additional variables (continued)

Year	swaiting	sduratio	scovratl	H568	H580	H540	fed	gov	vturn	rae	pres
2001	2.3	195	0.9	2.2	63.3	8.9	0	3.3	66.5	0.77	1
2002	2.3	195	0.9	2.2	63.7	8.8	0	2.3	64.7	0.76	1.5
2003	2.3	195	0.89	2.3	63.8	8.8	0	1.5	64.7	0.76	1.5
2004	2.3	195	0.89	2.3	64.5	9	0	1.3	62.3	0.77	1.5
2005	2.3	195	0.89	2.3	64.7	8.8	0	2	63	0.77	1.5
2006	2.3	174	0.89	2.4	64.6	8.7	0	2	63	0.77	1.5
2007	2.3	153	0.89	2.4	64.7	8.5	0	2.3	63.1	0.75	1.5
2008	2.3	133	0.89	2.5	64.7	8.5	0	2.5	63.7	0.74	1.5
2009	2.3	112	0.89	2.7	64.7	8.2	0	2.8	62.6	0.76	1.5
2010	2.3	91	0.9	2.7	66	8.1	0	2.8	62.6	0.76	1.5
2011	2.5	114	0.89	2.7	64.3	8.1	0	2.5	63.5	0.77	1.5
2012	2.7	138	0.89	2.8	61	8.2	0	2.3	62.7	0.79	1.5
2013	2.9	161	0.89	2.8	58.9	8	0	3	62.7	0.79	1.5
2014	3.1	185	0.89	2.8	57	8	0	2.8	59.2	0.79	1.5
2015	3.3	208	0.89	2.8	55.7	8	0	3	58.7	0.78	1.5
2016	3.2	208	0.9	2.8	54	8	0	3.5	57.5	0.8	1.5
2017	3.2	208	0.91	2.8	52.4	8.1	0	2.8	55.4	0.81	1.5
2018	3.1	208	0.92	2.8	50.9	8.1	0	2.3	55.6	0.82	1.5
2019	3.1	208	0.92	2.8	50.9	8.1	0	2	53.8	0.84	1.5
2020	3	208	0.93	2.8	51.1	8.1	0	2.8	53.2	0.85	1.5
2021	3	208	0.94	2.8	51.3	8.2	0	2.8	53.2	0.85	1.5

Table A7
Cluster 3: Average values

Year	health	old pop	death rate	life exp	V1A	V1B	V1C	V1D	tot emp	disch	gov ins	srtsw26s
2001	5.9	13.7	11.4	72.9	0.15	0.84	0	0	24.4	20703	95.6	0.7
2002	6.1	13.9	11.5	73.1	0.15	0.85	0	0	24.4	20804	95.9	0.69
2003	6.3	14.1	11.6	73.2	0.15	0.85	0	0	25.3	20964	96.3	0.69
2004	6.4	14.3	11.5	73.5	0.13	0.86	0	0	24.9	21195	96.5	0.68
2005	6.4	14.5	11.9	73.5	0.13	0.86	0	0	24.7	20873	97.2	0.67
2006	6.4	14.8	11.9	73.7	0.13	0.86	0	0	25.3	20886	97.4	0.66
2007	6.1	15.3	12.2	73.8	0.14	0.85	0	0	25.3	20246	97	0.67
2008	6.4	15.5	11.9	74.3	0.14	0.85	0	0	25.3	20210	96.6	0.66
2009	7	15.7	11.8	74.8	0.15	0.85	0	0	25.4	19804	96.6	0.65
2010	7	15.9	11.8	75.1	0.12	0.87	0	0.01	25.3	19170	96.4	0.64
2011	6.7	16.1	11.7	75.6	0.11	0.88	0	0.01	25.8	19335	96.2	0.64
2012	6.7	16.4	11.8	75.8	0.11	0.88	0	0	26.1	19304	95.4	0.64
2013	6.6	16.7	11.8	76.1	0.12	0.88	0	0	26.5	19267	95.4	0.64
2014	6.6	17	11.7	76.5	0.12	0.88	0	0	27.1	19336	95.2	0.64
2015	6.6	17.5	12.1	76.5	0.13	0.87	0	0	27.7	19306	95.2	0.64
2016	6.8	17.8	11.9	76.8	0.13	0.86	0	0	28.4	19374	95.4	0.64
2017	6.6	18.3	12.2	76.9	0.14	0.85	0	0	28.8	19058	95.6	0.64
2018	6.6	18.6	12.3	77	0.14	0.85	0	0	29.4	18835	96.3	0.64
2019	6.8	18.9	12	77.4	0.14	0.85	0	0	29.9	18553	96.5	0.64
2020	7.5	19.3	13.2	76.7	0.16	0.84	0	0	30.3	14840	96.6	0.64
2021	7.9	19.7	15.2	75.4	0.14	0.85	0	0	31	14614	97.0	0.64

PUBLIC SECTOR ECONOMICS
50 (2) 167-213 (2026)

Cluster 3: Additional variables (continued)

Year	swaiting	sduratio	scovratl	H568	H580	H540	fed	gov	vturn	rae	pres
2001	0.17	45	0.82	1.6	52.8	10.1	0	2.3	65	0.81	0.5
2002	0.17	44.2	0.81	1.7	54	9.8	0	3	62.5	0.8	0.5
2003	0.17	43.4	0.81	1.7	54.3	9.5	0	2.7	62.5	0.8	0.5
2004	0.17	42.6	0.81	2	52.2	9.5	0	2.7	60.4	0.8	0.5
2005	0.17	41.8	0.81	1.9	52.4	9.3	0	2.8	59.4	0.81	0.5
2006	0.17	41	0.8	2	52.3	9.1	0	2.8	55.5	0.79	0.5
2007	0.4	38.2	0.81	1.9	47.9	8.8	0	2.7	58.7	0.77	0.43
2008	0.46	37.5	0.81	2	49.1	8.8	0	2.7	59.1	0.78	0.43
2009	0.51	36.8	0.8	2.2	50.1	8.5	0	2.4	59.1	0.78	0.43
2010	0.57	36.1	0.8	2	49.6	8.5	0	1.7	59	0.78	0.43
2011	0.66	36.1	0.81	1.9	49.4	8.4	0	1.1	58	0.79	0.43
2012	0.74	36.1	0.82	2	48.8	8.3	0	1.7	58.7	0.78	0.43
2013	0.83	36.1	0.83	1.9	48.2	8.1	0	2.1	58.2	0.78	0.57
2014	0.91	36.1	0.85	1.9	48	8.1	0	2.6	57.9	0.79	0.57
2015	1.00	36.1	0.86	1.9	47.6	8.2	0	2.6	58.3	0.8	0.57
2016	0.91	36.1	0.86	1.9	47.9	8.2	0	2.3	58	0.81	0.57
2017	0.83	36.1	0.86	1.9	48.1	8.2	0	2.3	58.2	0.81	0.57
2018	0.74	36.1	0.87	1.9	48.3	8.2	0	2.1	58.7	0.82	0.57
2019	0.66	36.1	0.87	1.8	49.2	8.3	0	2	60.1	0.81	0.57
2020	0.57	36.1	0.87	1.8	49.9	8.3	0	1.7	60.6	0.81	0.57
2021	0.57	36.1	0.88	1.8	50.6	8.3	0	1.1	61.2	0.81	0.57

The effect of physical and intangible capital on labour productivity: role of institutional and development factors

VALENTIN LOVRIĆ, univ.mag.oec.*

Article**
JEL: O31, O34, O43, O47
https://doi.org/10.3326/pse.50.2.2

* The author is grateful to two anonymous referees who have contributed to the quality of the final version of the paper.

Views and opinions expressed in the paper are the author's own and do not necessarily reflect the policies or views of the Croatian National Bank.

** Received: June 1, 2025
Accepted: February 9, 2026

Valentin LOVRIĆ
Croatian National Bank, Trg hrvatskih velikana 3, 10000 Zagreb, Hrvatska
e-mail: valovric93@gmail.com
ORCiD: 0009-0008-5168-8112

Abstract

Economists agree that physical and intangible capital can have a positive impact on productivity growth. However, the dependence of intangible capital on adequate institutions may lead to its weak impact on productivity in countries with weak institutions. The aim of this paper is to assess the link between investments in physical and intangible capital on the one hand and productivity on the other. For determining differences in the level of effect of physical and intangible capital on productivity depending on institutional and developmental characteristics of countries, the research is divided into two parts. First, we assessed regression models for Croatia, after which we used panel analysis to evaluate a number of models for highly developed countries with quality institutions. The results suggest a significant positive effect of physical and intangible capital in advanced economies, while in Croatia only physical capital has a significant impact.

Keywords: labour productivity, intangible capital, panel analysis, institutions, economic development

1 INTRODUCTION

Determinants of labour productivity are important because they greatly influence economic competitiveness and growth of living standards (real wages) of individuals. Increase in labour productivity should necessarily lead to more output being produced in one hour of work with a more efficient use of production factors, resulting from workers being more equipped with physical and intangible (intellectual) capital per hour of work. Also, the increase in labour productivity is reflected in the growth of the overall technology level and other productivity factors, influenced by the quality and quantity of highly skilled workforce, as well as formal (public sector, laws, regulations, penalties) and informal (customs, norms) institutions, whereby well-designed economic policies can stimulate these factors.

Physical capital (machines, devices) and its role in productivity growth are well defined in the literature, but the accelerated development of intangible capital requires more detailed analyses of it and its role. One of the first references to intangible capital has been made by Griliches (1981), implying total stocks of intangible value (or know-how) in the economy (R&D investments, patents, organisational and human capital). Researchers also provided more detailed explanations of the positive effect of intangible capital on labour productivity on both microeconomic (e.g. Brynjolfsson, Hitt and Yang, 2002) and macroeconomic level (Solow, 1987; Nakamura, 2010).

Although the effect of intangible capital on labour productivity has already been investigated in more detail both in theory and empirically, the link with institutional and development characteristics has been studied only partially. In theory, a good institutional environment is necessary for intangible capital to have an effective impact on productivity (Acemoglu, Aghion and Zilibotti, 2006; Aghion and Howitt, 2006). But what is needed even more are quality and stimulating economic policies, developed financial markets and laws on the protection of property rights and efficient public administration. The most developed countries will base productivity growth on original innovation and the use of innovative intangible capital. In other words, they are at the technological frontier, i.e. their good institutional conditions lead to an increase in the efficiency of such capital (Acemoglu, Aghion and Zilibotti, 2006).

High levels of human capital can also lead to more efficient use of intangible capital in the economy. Clearly, there is a significant link between productivity and intangible capital in countries with well-defined institutional and development characteristics and an environment stimulating investments.

The impact of intangible capital characteristics on productivity can be analysed by comparing countries with different institutional and development characteristics, which is the objective of this paper. In the first part, we conduct a regression analysis of aggregate data and examine the impact of physical and intangible capital on labour productivity in Croatia. According to WIPO data (2025), Croatia is classified as a country with a relatively inadequate institutional framework for stimulating intangible investments, and the stagnant productivity dynamics in recent years serve as a good example of a country in which the impact of intangible capital may not be significant due to inadequate institutional support in the transition period toward higher development levels.

The second part is a panel analysis of sectoral and aggregate data of highly developed countries with quality institutional characteristics. The objective is to identify differences between the two parts of the analysis, in particular the importance of the long-term effects of intangible capital on labour productivity. By doing so, we could determine potential differences in the impact of institutional and development characteristics on the effect of intangible capital on productivity. This paper, modelled on Griliches (1995), uses regression analysis for estimating and calculating long-term coefficients of effect of (in-)tangible capital on productivity. Indeed, due to a possible time lag in the impact of capital on productivity, any analysis of standard estimated coefficients might not provide adequate conclusions. In the second part of the analysis, we also use an estimator which takes into account certain properties in the panel data that may potentially impair the impartiality of model results (primarily cross-sectional dependence). In the second part of the analysis, we also employ an estimator that accounts for certain properties of panel data that may potentially compromise the unbiasedness of the model results (primarily cross-sectional dependence). In this context, we estimate individualized average (as well as pooled) effects that are isolated from spillover effects arising from the statistical mechanism of cross-sectional dependence.

This research offers a more flexible design (comparative analysis of country with weaker institutional features and those with superior institutions) that can better interpret the causes of links between different types of capital and productivity. The review of empirical research provided below did not find a satisfactory number of studies dealing with the impact of physical and, more importantly, intangible capital

PUBLIC SECTOR
ECONOMICS
50 (2) 215-240 (2026)

on labour productivity in various institutional circumstances. Moreover, there have been no empirical surveys of determinants of labour productivity in relation to intangible capital for Croatian case. Furthermore, a partial novelty is the use of models that estimate the long-term effects of links between selected variables (as well as the consideration of certain statistical issues in the part of the analysis with panel data related to cross-sectional dependence), but in a different context than in other earlier studies of productivity determinants using similar estimators.

The results of this research primarily imply a significant impact of both physical and intangible capital on labour productivity in countries with developed institutions. On the other hand, a significant effect of intangible capital did not emerge in the analysis for Croatia. The result points to the importance of institutions that would support the creation of a foundation for the accumulation of intangible capital which, due to its complexity of application, is highly sensitive to institutional support in the country. Support is essential through well-designed laws, reducing political and economic costs that will allow already established but lagging market players to exit the market and facilitate the entry of fast-growing companies into the market. It also facilitates the financing of such companies, for example through subsidies, the development of financial markets and the creation of adequate human capital.

The following chapter provides the most important findings of previous empirical research on the impact of intangible capital on labour productivity; third chapter provides an overview and description of the data, their sources and the econometric methods used; fourth chapter presents the results of empirical research and the fifth chapter provides discussion of results and conclusions.

2 LITERATURE OVERVIEW

Theoretical research accumulated to this day mainly dealt with intangible capital analyses in the context of endogenous emergence of technological advances affecting productivity growth (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992; Benhabib and Spiegel, 1994; Griliches and Mairesse, 1998).

Recent decades have seen weaker labour productivity growth dynamics, especially in the most developed countries, which some researchers attribute to lower growth rates of intangible capital investment (Van Ark, 2016; Van Ark and O'Mahony, 2016). However, the effect of intangible capital on productivity remains relatively significant in the empirical sense, not only in the theoretical sense, as suggested by the above authors.

Among the empirical research so far, interesting findings are offered by Corrado, Haskel and Jona-Lasinio (2017), analysing the impact of intangible capital investments on productivity growth in ten EU member states in the period 1998-2007. They find that productivity is higher in sectors with greater use of ICT technology and investments in intangible capital and knowledge.

Furthermore, in the extensive empirical literature providing evidence of a statistically significant positive link between intangible (and certainly physical) capital and the level of production per worker, surveys by Bontempi and Mairesse (2015) for Italy, Roth and Thum (2013), Roth (2020), Hintzmann, Lladós-Masllorens and Ramos (2021) for EU countries, Corrado, Hulten and Sichel (2009) for the USA in the period 1973-2003 stand out. Similar results are also reported in Baldwin et al. (2009) for Canada, and Fukao et al. (2009) for Japan. The latter research is particularly interesting because they find a significant positive link for the period 1985-2005, although investments in intangible capital in Japan in the late 1980s and early 2000s recorded slower growth, which can explain the difference in labour productivity growth between Japan and the USA. Slow investment in intangible capital in Japan can also be linked to general formal, and even more so, informal Japanese institutions that are more generally rigid in a socio-cultural sense (slow to accept new values and prone to traditional values) and do not allow for subtle adaptation to innovations resulting from intangible capital investments (Rosser and Rosser, 2018).

The differences in dynamics have also been dealt with by Hao, Manole and Van Ark (2009) for France, Germany, Italy, Spain, USA and the UK, implying that countries with relatively higher intangible capital investments have better productivity indicators. Again, differences between countries can be explained by differences in their public policies and institutions that do not channel the private sector enough toward innovation and primary research.

Furthermore, a positive link is also found in sector analyses, as Eberhardt, Helmers and Strauss (2013) show for 13 industries in 10 highly developed countries; Crass, Licht and Peters (2015) for sectors in Germany, where the effect of physical capital on productivity increases parallel to an increase in intangible capital investment.

While, on the one hand, the literature points to the importance of physical, and especially intangible capital, on the other hand there is a lack of research into institutional and developmental determinants of the relationship between physical and intangible capital and labour productivity.

The only piece of research we found that goes into more detail regarding the institutional and development determinants of the link between intangible capital and labour productivity is the one by Castelli et al. (2024). They introduce certain heterogeneities and assess the impact of intangible capital on productivity in western and eastern European countries, identifying complementarity of the institutional environment as the missing link for explaining differences. For the period 2000-2017, they deepen their analysis further by adding interactive members to the panel model equations for eastern EU countries, assessing the differences between EU core countries (with good institutional conditions) and peripheral countries (i.e. eastern), with a weaker development environment. The results suggest a positive and significant link between labour productivity and intangible capital, for all EU countries. On the other hand, coefficients expanded with interaction members of eastern EU countries with intangible capital do not show a significant positive link with productivity, suggesting that lack of capacity and an adequate institutional environment is the main reason for the weak impact of intangible capital on

PUBLIC SECTOR ECONOMICS
50 (2) 215-240 (2026)

productivity, resulting in a more pronounced productivity gap between the eastern and western EU countries.

Moreover, Roth (2019) argues that the literature on the determinants of intangible capital investment decisions, as well as on the efficiency of the impact of physical and intangible capital on productivity, depends strictly on the quality of public infrastructure that helps such investment to be used more efficiently. Public investment in intangible capital then encourages companies to invest in such capital themselves, thereby ensuring productivity growth.

3 DATA AND METHODOLOGY

The paper uses data from the EU KLEMS database (*EU level analysis of capital (K), labour (L), energy (E), materials (M) and service (S) inputs*) (Bontadini et al., 2023).

For the purpose of econometric analysis, we collected sectoral and aggregate data on net cumulative investments (stocks) of total physical capital (*Total tangible capital stock – K*), which serve as a control variable. For the main independent variable, we collected data on net cumulative investments in intangible capital (*Total intangible capital stock – IN*), such as computer operating systems and databases, research and development, organisational and human capital (Bontadini et al., 2023).

K and IN were calculated using the geometric PIM method (*Perpetual Inventory Method*), while details on K and IN calculation (with depreciation rate included) are explained in detail in Bontadini et al. (2023). For the calculation of the effect of capital equipment per hours of work, K and IN are then divided by the total annual hours of work.

Net cumulative investments (stocks) were selected for analysis, rather than current or GFCF (*Gross Fixed Capital Formation*). Namely, stocks are better observed over time, i.e. the impact of K or IN on productivity depends on the composition of investments in present and past periods, and capital, such as the intangible one, needs a certain period in order to be efficiently applied in companies, sectors or the whole economy. GFCF is not so suitable for observation over time; even when included in regression specifications as lagged coefficient, it will not well represent the effect of time flow as it ignores the impact of time depreciation of capital (Eberhardt, Helmers and Strauss, 2013). Cumulative capital and accumulated knowledge from the beginning of the period means little if it does not continue to increase, i.e. depreciation must be introduced in the calculation (e.g. the use of new working methods and new knowledge from twenty years ago will not mean much for current productivity if new knowledge is not introduced over time).

For the calculation of dependent variable – labour productivity – we first collected Gross value added (GVA) data. After that, GVA was divided by data on total annual hours of work of employees in a specific sector or country in order to obtain a comparable measure of *Labour productivity* (LP). There are several reasons for selecting GVA per hour of work. While GVA per employee is also used often, such indicator may be more difficult to compare internationally due to differences in working hours of employees over certain periods in different countries and sectors. Labour productivity levels measured by GVA per hour of work in two countries in comparison could be the same, while GVA per number of employees could be higher in the country where employees work more hours per year, and one employee can produce more annually, leading to apparently higher labour productivity. Also, the use of GVA, unlike GDP, avoids the unintended effect of taxes.

In order to reflect the real values, all data are reported in constant prices (*chained linked volumes*), with price deflators adjusted for each item of national accounts (i.e. different deflator for GVA, physical and intangible capital).

All data – both sectoral and aggregated – refer to the period 1995-2021. Aggregated data includes all economic sectors according to the usual NACE Rev. 2 classification, while sectoral data includes 15 service and manufacturing industries (shown in table 1, coded by ordinal number). While for 13 highly developed economies (Austria, Czechia, Denmark, Finland, France, Italy, Japan, the Netherlands, Germany, the United States, Spain, Sweden and the United Kingdom) both sectoral and aggregated data were available, for Croatia only aggregated data is obtainable.

Thus, the first part analyses observations of aggregate data for Croatia. In the second part, a panel analysis was carried out for 13 countries – first of aggregate data for 12 countries (UK missing), over 27 years, i.e. 324 observations. Sectoral data were then analysed, for the country-sector (180 units), over 27 years, totalling 4,860 observations.

Table 1

Description of sectors used in the analysis

Ordinal No.	Sector name	Database code
1.	Mining and quarrying	B
2.	Manufacture of food products, beverages and tobacco products	C10-C12
3.	Manufacture of textiles and wearing apparel	C13-C15
4.	Manufacture of wood and paper	C16-C18
5.	Manufacture of chemicals and chemical products	C20
6.	Manufacture of basic pharmaceutical products and preparations	C21
7.	Manufacture of rubber and plastic products	C22-C23
8.	Manufacture of computer, electronic and optical products	C26
9.	Manufacture of machinery and equipment n.e.c.	C28
10.	Manufacture of motor vehicles and other transport equipment	C29-C30
11.	Construction	F
12.	Transportation	H
13.	Information and communication	J
14.	Computer programming and consultancy, information service activities	J62-J63
15.	Public administration and defence, education, human health and social care activities	O-Q

Source: Author, based on Bontadini et al. (2023).

PUBLIC SECTOR ECONOMICS
50 (2) 215-240 (2026)

All variables have been log transformed to better adapt to the linear assumptions of the regression models used.

For this type of analysis, we optioned to use some of the methods that calculate the long-term effects of certain variables on the variable of interest. In this sense, model should also provide the use of time shifts of all variables, bearing in mind the bias of empirical tests of a static nature that do not take into account statistically significant dynamic and time-persisting relationships between variables.

The formal explanation of the link between labour productivity and physical and intellectual capital can be explained by a neoclassical production function. Most commonly used is the Cobb-Douglas production function, which, for example, Roth and Thum (2013) represent as:

$$Y = A\mathcal{K}^{\alpha} L^{\eta} \mathcal{IN}^{\beta} \varepsilon \tag{1}$$

where Y is the production level represented by GVA, A is a constant, defined as the technology or total factor productivity that is neutral in Hicks' sense, i.e. does not alter the distribution of labour and capital in production but increases their productivity. $\mathcal{K}$ are stocks of a sector's physical capital, L are the total hours of work of employees, $\mathcal{IN}$ are stocks of intangible capital used in the production of the sector and ε is the stochastic error component. Assuming linear homogeneity of the production function (sum of the elasticity coefficients α, β and η equals 1), equation (1) can be corrected with the total hours of labour:

$$\frac{Y}{L} = A\left(\frac{\mathcal{K}}{L}\right)^{\alpha}\left(\frac{\mathcal{IN}}{L}\right)^{\beta} \varepsilon \tag{2}$$

with $L^{1-\alpha-\beta} / L = L^{1-\alpha-\beta-1} = L^{-\alpha-\beta} = 1/ L^{\alpha+\beta}$ and $\eta = 1-\alpha-\beta$. By log transforming expression (2), it can be expressed as:

$$ln\left(\frac{Y}{L}\right) = lnA + \alpha ln\left(\frac{\mathcal{K}}{L}\right) + \beta ln\left(\frac{\mathcal{IN}}{L}\right) + \varepsilon \tag{3}$$

where $ln\left(\frac{Y}{L}\right)$ has already been defined by LP, $ln\left(\frac{\mathcal{K}}{L}\right)$ is K, while $ln\left(\frac{\mathcal{IN}}{L}\right)$ is IN. Equation (3) provides a good starting point for regression analysis of data for Croatia and panel data for 13 highly developed countries.

The following is an explanation of the model estimation for Croatia. For an adequate estimation of long-term impact of IN on LP, we use an autoregression model with distributed lags (*AutoRegressive Distributed Lag – ARDL*). With the help of the ARDL model, the dynamic specification provides long-term coefficients that best determine the effect of IN and K on LP. Due to the small sample for Croatia, this model is also useful because it allows different shifts of (non-)dependent variables, which leads to savings in degrees of freedom. Defining t as a time unit, such a dynamic model can, as demonstrated by Kripfganz and Schneider (2023), be represented as:

$$LP_t = c_0 + \sum_{i=1}^{p} \phi_i LP_{t-i} + \sum_{i=0}^{q} \beta_i' x_{t-i} + u_t \tag{4}$$

where c_0 is a constant, $\sum_{i=1}^{p} \phi_i LP_{t-i}$ represents auto-regression structure of LP with p shifts, while $\sum_{i=0}^{q} \beta_i' x_{t-i}$ shows the structure of the time distribution of effects of independent variables with q shifts, whereby β_i comprises coefficients of effect of K and IN on LP to be estimated, while $x_t = (K_t, IN_t)'$ is the independent variables vector comprising K and IN. The specification in (4) makes it possible to estimate the long-term effects of K and IN on LP, just as Griliches (1995) suggested in the analysis of the impact of capital investments on productivity. Such long-term coefficients can be obtained by redefining the ARDL model in (4), and the model can be represented as:

$$\Delta LP_t = c_0 - \alpha(LP_{t-1} - \theta' x_{t-1}) + \sum_{i=1}^{p-1} \phi_i \Delta LP_{t-i} + \sum_{i=0}^{q-1} \beta_i' \Delta x_{t-i} + u_t \tag{5}$$

where α is the coefficient with the *error-correction member (EC*), while θ comprises long-term coefficients for K and IN. α and θ can be defined as:

$$\alpha = (1 - \sum_{i=1}^{p} \phi_i) \tag{6}$$

$$\boldsymbol{\theta} = \frac{\sum_{i=0}^{q} \beta_i}{\alpha} \tag{7}$$

For the ARDL analysis, the Pesaran, Shin and Smith (2001) boundary test (PSS Bounds test) is also important, indicating the existence of a genuine long-term connection and not one driven by similar common trends through a process of spurious regression. The null hypothesis of the test assumes that there is no long-term relationship between variables, and the test itself contains two steps. The first step measures the aggregate statistical significance of all coefficients of independent variables via the group F-test, i.e. the significance of long-term coefficients, and the second step measures the statistical significance of the error correction member parameter α. At the same time, the t-test for α is not valid in the estimated model as t statistics for α does not have a standard distribution at null hypothesis, and the second step of the boundary test uses adjusted critical values according to Kripfganz and Schneider (2020) and approximated p-values for the t-test.

We now continue with the description of panel model estimations for thirteen highly developed countries. Since the focus of the paper is on long-term impact assessment, some of the methods of cointegration analysis can be used, adapted to the panel data with a longer time component, which is potentially non-stationary. Ditzen (2021) states that the model for assessing impact of IN and K on LP should take into account the potentially significant dynamic structure of variables, i.e. it should be able to shift all variables. Also, if not removed from the specification, there may be cross-sectional dependence in model residuals, which may lead to biased estimates. Therefore, an econometric analysis of the panel data was selected, which takes all these problems into account.

The panel ARDL model, identical to the one from the first part of the analysis, but adapted to the panel data, is initially expressed, as per Chudik et al. (2013), as:

$$LP_{i,t} = c_i + \sum_{l=1}^{p} \phi_{i,l} LP_{i,t-l} + \sum_{l=0}^{q} \beta'_{i,l} x_{i,t-l} + e_{i,t} \tag{8}$$

whereby:

$$e_{i,t} = \varrho'_i f_t + u_{i,t} \tag{9}$$

$$x_{i,t} = \Gamma'_i f_t + v_{i,t} \tag{10}$$

$i = 1, ..., N$ and $t = 1, ..., T$ are cross-sectional and time units. In addition, f_t is a $m \times 1$ vector of unobserved common factors (m is the number of time shifts of these factors), ϱ_i is a vector of so-called heterogeneous factor loading associated with LP, i.e. Γ_i is a matrix of heterogeneous factor loading associated with independent variables (k is the number of independent variables). c_i represents fixed effects specific to cross-sectional units. Idiosyncratic components $v_{i,t}$ of $x_{i,t}$ are under assumption distributed independently of $u_{i,t}$. $LP_{i,t}$ and vector $x_{i,t}$ are identical to the first part of the analysis, only in the notation for the panel data. Unlike the first part of the analysis, the model in (4) cannot be redefined directly to obtain long-term coefficients, as the estimation of the equation by using the standard least squares method will lead to a biased statistical estimation. This bias is created by the structure of common factors defined by (9) and (10) (which is why cross-sectional dependence appears in the panel data), i.e. the bias problem arises due to the omission of an important independent variable (in this case f_t) (Chudik et al., 2013; Ditzen, 2021).

Chudik and Pesaran (2015) therefore propose a method that can estimate equation (4) consistently, approximating common factors by adding cross-sectional averages as additional members to this equation. It can be shown that with the cross-sectional component increasing, the addition of cross-sectional averages and its shifts to the dynamic model, it encompasses f_t and OLS (*Ordinary Least Squares*) estimation becomes impartial (Chudik and Pesaran (2015) offer mathematical evidence). According to their framework rule, at least $\sqrt[3]{T}$ of the shifts of such averages should be added, where T is the time component for the available panel data. An additional problem with the dynamic specification (4) is the shift of the dependent variable leading to endogeneity due to the correlation of shifts with previous realizations of the relation error, leading to inconsistencies of the OLS estimator. However, Chudik and Pesaran (2015) explain that a specification with additional members of cross-sectional averages also addresses the issue of endogeneity and becomes impartial if a sufficient number of shifts of such averages are associated. In such cases, equation (4) can be formulated as:

$$LP_{i,t} = c_i + \sum_{l=1}^{p} \phi_{i,l} LP_{i,t-l} + \sum_{l=0}^{q} \beta'_{i,l} x_{i,t-l} + \sum_{l=0}^{p_{\bar{z}}} \gamma'_{i,l} \bar{z}_{t-l} + e_{i,t} \tag{11}$$

The vector $\overline{z}_t = (N^{-1}\sum_{i=1}^{N} x_{it}, N^{-1}\sum_{i=1}^{N} LP_{it})$ contains cross-sectional averages of all variables connected to the model, where coefficients with cross-sectional averages within the vector $\gamma_{i,l}$ are usually omitted in the model results because they do not have a clear interpretation. Therefore, equation (11) becomes well specified to obtain long-term coefficients from such specification by applying the OLS method. A model can then be written for the second part of the analysis as demonstrated by Chudik et al. (2013):

$$\Delta LP_{i,t} = c_i + \theta_i' x_{i,t} + \sum_{l=1}^{p-1} \phi_{i,l} \Delta LP_{i,t-l} + \sum_{l=0}^{q-1} \beta_{i,l}' \Delta x_{i,t-l} + \sum_{l=0}^{p_{\overline{z}}} \gamma_{i,l}' \overline{z}_{t-l} + e_{i,t} \quad (12)$$

where long-term coefficients are contained inside $\theta_i = \alpha_i^{-1}\sum_{l=0}^{q} \beta_{i,l}$, while coefficient with EC member is defined as $\alpha_i = 1 - \sum_{l=1}^{p} \phi_{i,l}$. For now, θ_i are shown heterogeneously by i; however, an explanation is in order. To obtain more information on how variables are connected, equation (12) is estimated in two ways for both data levels.

The first concerns the estimation of model (12) with the assumption of homogeneity of long-term coefficients, i.e. θ. In such cases, equation (12) is estimated through the DCCE-PMG (*Dynamic Common Correlated Effects – Pooled Mean Group*) estimator, which calculates one pooled long-term coefficient for K and IN based on panel data, while short-term coefficients for K and IN are individually estimated by cross-sectional units, and their significance is measured on the average of these short-term coefficients per unit.

The other method refers to the estimation of equation (12) with the DCCE-MG (*Dynamic Common Correlated Effects – Mean Group*) estimator, which evaluates the individual equations for each cross-sectional unit, taking the average for each coefficient per unit and measuring the correlation of variables with this average coefficient value. This ensures the heterogeneity of model coefficients, i.e. the possible significant heterogeneity between cross-sectional units is controlled. The significance of effect of K and IN on LP is measured by average long-term coefficients $\overline{\theta}_{MG} = N^{-1}\sum_{i=1}^{N} \theta_i$.

We should also explain how the level of cross-sectional dependence is determined in panel data. The estimation of cross-sectional dependence is tested using the Pesaran CD (*Cross-sectional Dependence*) test, whose results are important because they show how well the influence of cross-sectional dependence has been removed in the model by the addition of cross-sectional averages. Test may also be important before the model estimation, as an indicator of the need to include cross-sectional averages in the model. The null hypothesis of the Pesaran test is that the errors of the relation $e_{i,t}$ of the test equation without cross-sectional averages are weakly cross-sectionally dependent, while the alternative hypothesis assumes strong dependence. Pesaran (2004) proposes the following test statistics:

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \right) \quad (13)$$

where the estimated correlation coefficient is:

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^{T} \hat{e}_{it} \hat{e}_{jt}}{\left(\sum_{t=1}^{T} \hat{e}_{it}^{\,2}\right)^{1/2} \left(\sum_{t=1}^{T} \hat{e}_{jt}^{\,2}\right)^{1/2}} \tag{14}$$

Chudik, Pesaran and Tosetti (2011) define the first type of weak cross-sectional dependence of errors as follows:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \left| \varrho_i \right| = 0 \tag{15}$$

Another, stronger version at a specific constant K, independent of N, where for any large enough N, is defined as:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \left| \varrho_i \right| \geq K > 0 \tag{16}$$

Cross-sectional independence is here defined as $\theta\, \varrho_i = 0 \, \forall i$. Other factors are formulated in a similar manner. Since cross-sectional independence is in itself an excessively restrictive assumption, only a strong dependence defined with an asymptotic definition in (16) is problematic (Pesaran, 2015). In such a case, unbiasedness and consistency properties of the estimator will be violated, because the structure of the model residuals will show a strong correlation through cross-section units, and these errors will no longer be independently and identically distributed. In fact, unobserved common and heterogeneous factor loadings will be part of the error member of the relation $e_{i,t}$. In this case, the problem of bias arises because important independent variables have been omitted (*omitted-variable bias*) and the observed explanatory variables will be correlated with unobserved common factors. By rejecting the null hypothesis of the Pesaran test, the OLS estimation becomes inconsistent (Everaert and de Groote, 2016).

The Pesaran test shall be performed prior to the use of an estimator with CCE factors to determine whether there are cross-sectional dependencies in the data (shown in the main results table). A model with CCE factors is then used and the test is applied in a standard way after modelling has been performed in order to obtain information on the representativeness of estimates. Model estimation follows.

4 RESULTS OF EMPIRICAL RESEARCH

4.1 RESULTS OF ANALYSIS FOR CROATIA

Before the model estimation, it is important to obtain basic information about the time series of variables, and above all about the presence of unit root in the series. Indeed, the estimation of the ARDL model allows a combination of variables that are integrated and of orders 0 and 1, i.e. both variables that are stationary in levels and those that are stationary after the transformation in the first differences but does not allow analysis with variables that are also I(2) (integrated to the second order). In case of I(2), the OLS method encounters estimation difficulties. Also, if all variables are I(0), the purpose of calculating long-term coefficients is questionable. For this purpose, the Augmented Dickey-Fuller test (ADF) of the unit root is used, where the null hypothesis of the test assumes non-stationarity of the time series. The test results (available on request) suggest that the K, IN and LP variables are all I(1), allowing the analysis to continue. Table 2 shows model results with estimated long-term coefficients.

PUBLIC SECTOR ECONOMICS
50 (2) 215-240 (2026)

For Model (1), the results indicate a long-term significant relationship between K and LP at all standard levels of significance, with EC being of the correct sign and implying a 60.4% reduction of imbalance in the first year. The results of the PSS test also indicate long-term correlation of variables, with long-term parameters being significant at the level of 10%, with the EC member reaching significance at the level of 5%. On the other hand, the results indicate that the long-term coefficient with IN is not significant. In the model variant without K, the long-term coefficient with IN is still not significant, while the null hypothesis of the PSS test cannot be rejected, which indicates that there is no real long-term link between intangible capital and labour productivity in Croatia.

Table 2

Model estimation with long-term coefficients, Croatia

	(1)	(2)
Long-term coefficient (IN)	0.263 (0.082)	0.57 (0.401)
Long-term coefficient (K)	0.391*** (0.168)	
$\hat{\alpha}$	-0.604*** (0.163)	-0.159* (0.084)
Short-term coefficient (IN, t)	-0.461* (0.241)	
Short-term coefficient (IN, $t-1$)	0.057 (0.243)	
Short-term coefficient (IN, $t-2$)	-0.509* (0.244)	
Constant	0.398** (0.147)	0.301* (0.16)
R^2	0.499	0.148
Autocorrelation	0.077	0.354
Heteroskedasticity	0.404	0.579
PSS test (F test)	4.789*	1.827
PSS test (t test)	-3.698**	-1.898
Observations	24	24

*Note: ***, **, * refers to rejection of the null hypothesis at 1%, 5% and 10% level of significance respectively. The optimal number of shifts for models was selected based on the Akaike information criterion. For diagnostic model verification, we used Breusch-Godfrey LM test for autocorrelation of errors (null hypothesis: no autocorrelation) and Cameron and Trivedi decomposition of IM-test for heteroskedasticity (null hypothesis: homoskedasticity). For diagnostic tests, p-values are shown. For the PSS test, test sizes are specified.*

Source: Author's own calculations.

PUBLIC SECTOR
ECONOMICS
50 (2) 215-240 (2026)

To further establish the relationship between variables, the Engle-Granger (1987)-type procedure was carried out in two phases. The results (available on demand), in the variant with and without factor K, indicate that the variables are not long-term related, as the residual time series in both variants of long-term relationship are non-stationary and indicate different long-term dynamics. Also, the results of the error correction model suggest a short-term significant effect of K on LP, but not on IN. Although the results obtained are based on a limited sample, they nevertheless systematically imply that K significantly influences labour productivity in Croatia. Conversely, the effect of IN on LP is not statistically significant in any variant of different specifications, indicating that there are no adequate conditions for additional productivity growth based on intangible capital. The results obtained will be compared with estimates for highly developed countries.

4.2 RESULTS OF PANEL ANALYSIS FOR HIGHLY DEVELOPED ECONOMIES

Here again, performing the unit root test is useful for analysing long-term relationship of variables, because the reasons explained in the first part also apply to the panel analysis. Here we use the Pesaran CIPS test (*Cross-sectionally augmented Im-Pesaran-Shin*), which assumes the possibility of cross-sectional dependence. Test results (available on demand) suggest combinations of I(0) and I(1) variables, i.e. none of the variables are I(2). The results of regression models follow, first for sectoral data of the estimated DCCE-PMG and DCCE-MG methods (table 3).

Model results based on DCCE-PMG estimator indicate that long-term coefficients at K and IN are positive and significant (IN is significant even without the inclusion of K). The estimator with the EC is significant and with a negative sign, as expected. The results of the Pesaran test indicate that the decision to remove cross-sectional dependence from the residuals of the estimated models in the PMG variant and to include cross-sectional averages was correct. In the model variant without the CCE factor, the null hypothesis is rejected, which implies a significant level of dependence in model residuals. Conversely, after the inclusion of cross-sectional averages, the null hypothesis cannot be rejected, and the resulting estimates can be considered unbiased.

For the DCCE-MG model, which assumes the heterogeneity of long-term coefficients (Model (3)), coefficients with K and IN are no longer significant. However, as the null hypothesis of the Pesaran test is rejected, these results cannot be considered reliable. On the other hand, in the model variant without K (4), the coefficient with IN is significant and positive, and cross-sectional averages effectively remove dependence from residuals, as indicated by the Pesaran test.

Table 3

Model estimations with long-term coefficients, sectoral panel data for thirteen highly developed countries

	DCCE-PMG		DCCE-MG	
	(1)	**(2)**	**(3)**	**(4)**
Long-term coefficient (IN)	0.144*** (0.02)	0.295*** (0.014)	-0.405 (1.118)	0.666** (0.299)
Long-term coefficient (K)	0.326*** (0.018)		0.289 (0.56)	
$\hat{\alpha}$	-0.493*** (0.019)	-0.423*** (0.023)	-1.041*** (0.088)	-0.661*** (0.0381)
Short-term coefficient (IN)	0.158* (0.095)	0.331*** (0.037)	0.23 (0.194)	0.297*** (0.0391)
Short-term coefficient (K)	0.104 (0.094)		-0.262 (0.196)	
R^2	0.3	0.47	0.82	0.65
Pesaran test (no CCE factor)	53.6***	83.04***	42.93***	57***
Pesaran test	0.82	0.26	-2.03**	-1.09
Country-sector	180	180	180	180
Observations	3,780	3,780	3,780	3,780
Pedroni test				
P	**Combined**		**Heterogeneous (group)**	
t-test	-3.882***		-4.493***	
ADF-test	-3.087***		-5.004***	

*Note: ***, **, * refers to rejection of the null hypothesis at 1%, 5% and 10% level of significance respectively. The estimated model coefficients are provided. Standard errors are specified in parentheses (for pooled coefficients at (1) and (2), clustered robust standard errors (per unit) were calculated). R^2 is the coefficient of determination. The optimal number of shifts was selected based on the Schwarz information criterion for each individual dynamic model of 180 countries-sectors, and the most frequent shifts were selected for all specifications. All details of model results, tests and sequential use of the information criterion for obtaining information on optimal shift structure are available from the author upon inquiry. Three shifts of cross-sectional averages of each variable were added for each model as per rule $\sqrt[3]{T}$. Due to the pronounced cross-sectional component in relation to time, a recursive average correction is used to reduce bias resulting from a smaller pattern of time series in such a panel structure modelled on Chudik and Pesaran (2015). This correction removes the partial mean from all variables, where $\tilde{\omega}_{it} = \omega_{it} - \frac{1}{t-1}\sum_{s=1}^{t-1}\omega_{is}$ and ω_{it} includes all variables except the constant. The Pedroni (2004) test provides information on stationarity of residuals of the initial model, in order to determine the significance of long-term relationship. The null hypothesis assumes that there is no long-term relationship for all panels, while the alternative of the test is divided into two types: pooled (by rejecting the null hypothesis, there is a long-term relationship for all panels) and heterogeneous (by rejecting the null hypothesis, there is a long-term connection only for some panels). The Pedroni test included a constant and was performed on all included variables.*

Source: Author's own calculations.

The results of the Pedroni cointegration test are shown at the bottom of the table. In both the combined and heterogeneous model variants, the test statistics indicate the rejection of the null hypothesis which assumes the absence of long-term relationship (assuming the non-stationarity of model residuals). Based on all these results, we can conclude that for sectoral data in the panel of highly developed countries, the impact of K and IN on LP is significant. For further examination of sectoral results, the appendix (table A1) provides additional dynamic and static estimations, which mainly support additional findings of DCCE-MG and DCCE-PMG estimators, especially as regards the long-term impact of IN on LP.

The following are estimates of aggregate data from the same set of countries (only without the UK) (table 4), which are used to further compare the results. As with sectoral data, aggregate data estimation models using PMG variant methods point to long-term significance of coefficients IN and K, and the results of the Pesaran test again point to effective elimination of cross-sectional dependence by adding cross-sectional averages to equations. The results of the Pedroni test in the pooled variant support the conclusions of the model on significant long-term relation.

On the other hand, results of the version using MG estimator are not so clear. Specification (3) with K and IN included shows no significance at θ_i. Variant (4) without K shows significance of the positive long-term coefficient, but the results of the Pesaran test fail to point to a single conclusion. However, the results of the Pedroni test in the heterogeneous version support the conclusion on long-term connection between variables. All of this, together with findings of the models using sectoral data from highly developed countries, point to a positive and significant long-term effect of both K and IN on LP.

Table 4

Model estimations with long-term coefficients, aggregate panel data for twelve highly developed countries

	DCCE-PMG		**DCCE-MG**	
	(1)	**(2)**	**(3)**	**(4)**
Long-term coefficient (IN)	0.583*** (0.017)	0.922*** (0.014)	-1.224 (4.169)	0.625*** (0.133)
Long-term coefficient (K)	0.524*** (0.022)		3.717 (3.635)	
$\hat{\alpha}$	-0.417*** (0.007)	-0.176*** (0.012)	-0.677*** (0.179)	-0.501*** (0.104)
Short-term coefficient (IN)	-0.24 (0.144)	0.537*** (0.12)	-0.677*** (0.179)	0.321** (0.137)
Short-term coefficient (K)	0.645*** (0.142)		-0.576*** (0.176)	
R^2	0.02	0.06	0.98	0.97
Pesaran test (no CCE factor)	13.67***	15.12***	14.21***	14.33***
Pesaran test	-0.62	1.36	-1.14	1.84*
Number of countries	12	12	12	12
Observations	264	264	264	264
Pedroni test				
	Combined		**Heterogeneous (group)**	
t-test	-1.932**		-1.729**	
ADF-test	-2.152**		-2.876***	

*Note: ***, **, * refers to rejection of the null hypothesis at 1%, 5% and 10% level of significance respectively. Models with aggregate data do not use recursive correction of average values. For other details, see note under table 4.*

Source: Author's own calculations.

5 DISCUSSION AND CONCLUSION

The methods used and the results obtained from the estimated models with data for Croatia and panel data from highly developed countries indicate potentially important differences in the contribution of intangible capital to the growth of labour productivity in these economies. Estimates indicate that labour productivity in Croatia is affected more by physical rather than intangible capital per hour. On the other hand, comprehensive model estimates with panel data for highly developed economies indicate a significant impact of both types of capital on long-term productivity. This means that in an equilibrium state (where variables are not liable to short-term shocks), the unit increase in intangible capital per hour of work does not have a significant lasting effect on labour productivity in Croatia but does so in advanced economies. Clearly, the role of institutional and development factors that enable productivity to grow is important, both through physical and intangible capital. More favourable institutional and development conditions and environment favour the use of intangible capital, which leads to faster productivity growth and thus faster real income growth per capita. This can partly explain differences in the level of economic development of Croatia and selected highly developed countries.

In the case of Croatia, only significant channel of productivity growth through K can lead to a stagnation of productivity growth when the country reaches the level of development at which further investment in physical capital brings little marginal benefits for productivity. Therefore, without the significant role of the IN channel, neither continued long-term productivity growth nor higher income per worker can be expected. This is also indicated by the research referred to in the literature review.

Furthermore, applying intangible capital to the economy requires a high level of human capital and a development environment that will encourage companies to make the structural investments necessary to increase productivity. The state should therefore establish a stimulating institutional and development environment necessary for the growth of labour productivity and economic development. The results of this empirical survey can be compared to the latest edition of the Global Innovation Index for 2025 (GII), which can partially mirror the institutional success of a country, as it explains the development of its innovative capacities (WIPO, 2025).

Croatia ranks 40th among 139 countries in the Global Innovation Index. According to the sub-components of this Index, Croatia has particularly poorly developed key institutional support (65th place), which indicates that more developed formal institutions, including practical efficiency of regulatory processes, would facilitate investment and growth of investments in intangible capital and lead to potential productivity growth. Furthermore, equality of all companies before the law (which again indicates the importance of formal institutions) would allow for a greater number and capacity of newly created fast-growing enterprises whose intangible capital investments would lead to their higher productivity and returns, which would have a positive impact on the whole economy.

Furthermore, Croatia ranks relatively poorly both in business (53rd place) and market (54th place) sophistication, indicating a lack of adequate basic sophisticated infrastructures for further productivity growth based on intangible capital. Therefore, the public sector's contribution to the development of intangible capital is essential, or as Aghion and Howitt (2006) state, well-designed economic policies that will stimulate growth of IN (subsidies, tax breaks and reliefs) are essential.

In more practical sense, Croatia as EU member state should consider a more efficient and structured decision-making process on how to spend assets from EU development funds. Related to the market sophistication rank, the development of financial markets could contribute to the relocation of labour and capital to the most efficient and most innovative but potentially risky companies is also very important (Aghion and Howitt, 2006).

This would stimulate the growth of IN and hence LP. Furthermore, 40th position in human capital rankings indicates that further reforms of the higher education system leading to higher levels of human capital, necessary for effective use of IN, are needed, as indicated by the papers cited in the literature review.

The introduction of such development policies in the long run could help reduce the productivity gap and speed up Croatia's convergence towards the average real income per capita of the most developed EU member states. On the other hand, the sampled advanced countries from this study are ranked as follows: Sweden (2), United States (3), United Kingdom (6), Finland (7), Netherlands (8), Denmark (9), Germany (11), Japan (12), France (13), Austria (19), Italy (28), Spain (29); and Czechia (32). These rankings point to their superior institutional development, which is then evident in the results of the econometric analysis which are in line with important theoretical findings on the role of the institutional environment on productivity growth and the economic development (Aghion and Howitt, 2006).

Therefore, arguments built by this research, according to which institutional development positively influences the transition of productivity growth based on accumulation of physical capital per worker (or hours) with a parallel imitation of technologies already created towards intangible capital and the creation of new innovations (Acemoglu, Aghion and Zilibotti, 2006), are partially confirmed by the results of econometric analysis and supported by the results of the GII. However, without further empirical testing and adequate analysis of institutions and the historical and cultural context, we should be careful when discussing the causal role of institutions and intangible capital in productivity growth.

However, looking at A1 charts, the results and model estimates in this paper are not surprising, because in the last available year (2021) and during the period 1995-2021, Croatia was almost continuously at the bottom when looking at the intangible capital stock when compared to highly developed countries in the sample. These are significant deficiencies in infrastructure and development conditions necessary for

PUBLIC SECTOR ECONOMICS
50 (2) 215-240 (2026)

a significant effect of IN on productivity, which can also be seen in the trends in labour productivity per hour during this period (figure A2). Croatia is constantly at the bottom relative to this criterion as well in comparison with highly developed countries.

Finally, the shortcomings of this, as well as possible recommendations for further research, should be noted. First, an important control variable – the level of human capital explaining the important link with intangible capital in the contribution to productivity – was omitted. However, as sectoral data are also used in the paper, the calculation of such an indicator would be too demanding at the moment. Therefore, the impact of human capital levels on labour productivity and the impact of patents and innovations should be analysed in future works, especially in Croatia where this has been poorly investigated. Furthermore, with the help of sectoral data, the link between K and IN with LP could be investigated in a more detailed and precise way, for which it is necessary to build adequate databases for sectors or even companies in Croatia. Finally, it is necessary to note that the data used in this study from the EU KLEMS database, while highly useful, are undoubtedly subject to statistical errors and inaccuracies. They should therefore be regarded as approximations of capital stock and productivity. Nevertheless, researchers will likely recognize that the data presented in this database do not differ substantially from other comparable sources available.

Disclosure statement
The author has no conflicts of interest to declare.

REFERENCES

1. Acemoglu, D., Aghion, P. and Zilibotti, F., 2006. Distance to Frontier, Selection, and Economic Growth. *Journal of the European Economic Association*, 4(1), pp. 37-74. https://doi.org/10.1162/jeea.2006.4.1.37
2. Aghion, P. and Howitt, P., 1992. A Model of Growth through Creative Destruction. *Econometrica*, 60(2), pp. 323-351. https://doi.org/10.2307/2951599
3. Aghion, P. and Howitt, P., 2006. Appropriate Growth Policy: A Unifying Framework. *Journal of the European Economic Association*, 4(2-3), pp. 269-314. https://doi.org/10.1162/jeea.2006.4.2-3.269
4. Baldwin, J. [et al.], 2009. Investment in Intangible Assets in Canada: R&D, Innovation, Brand, and Mining, Oil and Gas Exploration Expenditures. *The Canadian Productivity Review*, No. 26. https://doi.org/10.2139/ssrn.1517255
5. Benhabib, J. and Spiegel, M. M., 1994. The Role of Human Capital in Economic Development – Evidence from Aggregate Cross-Country Data. *Journal of Monetary Economics*, 34(2), pp. 143-173. https://doi.org/10.1016/0304-3932(94)90047-7
6. Blundell, R. and Bond, S. R., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics,* 87(1), pp. 115-143. https://doi.org/10.1016/S0304-4076(98)00009-8
7. Bontadini, F. [et al.], 2023. Euklems and IntanProd: industry productivity accounts with intangibles. *MPRA Paper*, No. 126218.
8. Bontempi, M. E. and Mairesse, J., 2015. Intangible capital and productivity at the firm level: a panel data assessment. *Economics of Innovation and New Technology*, 24(1-2), pp. 22-51. https://doi.org/10.1080/10438599.2014.897859
9. Brynjolfsson, E., Hitt, L. M. and Yang, S., 2002. Intangible Assets: Computers and Organizational Capital. *Brookings Papers on Economic Activity,* 1, pp. 137-198. https://doi.org/10.1353/eca.2002.0003
10. Castelli, C. [et al.], 2024. Tangible, intangible assets and labour productivity growth. *Journal of Economic Studies*, 51(9), pp. 272-289. https://doi.org/10.1108/JES-11-2023-0620
11. Chudik, A. [et al.], 2013. Debt, Inflation and Growth: Robust Estimation of Long-Run Effects in Dynamic Panel Data Models. *CAFE Research Paper*, No. 13.23.
12. Chudik, A. and Pesaran, M. H., 2015. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics,* 188(2), pp. 393-420. https://doi.org/10.1016/j.jeconom.2015.03.007
13. Chudik, A., Pesaran, M. H. and Tosetti, E., 2011. Weak and strong cross-section dependence and estimation of large panels. *Econometrics Journal,* 14(1), pp. 45-90. https://doi.org/10.1111/j.1368-423X.2010.00330.x
14. Corrado, C., Haskel, J. and Jona-Lasinio, C., 2017. Knowledge Spillovers, ICT and Productivity Growth. *Oxford Bulletin of Economics and Statistics*, 79(4), pp. 592-618. https://doi.org/10.1111/obes.12171

15. Corrado, C., Hulten, C. and Sichel, D., 2009. Intangible Capital and U.S. Economic Growth. *The Review of Income and Wealth*, 55(3), pp. 661-685. https://doi.org/10.1111/j.1475-4991.2009.00343.x
16. Crass, D., Licht, G. and Peters, B., 2015. Intangible Assets and Investments at the Sector Level: Empirical Evidence for Germany. In: A. Bounfour and T. Miyagawa, eds. *Intangibles, Market Failure and Innovation Performance.* Cham: Springer, pp. 57-111. https://doi.org/10.1007/978-3-319-07533-4_4
17. Ditzen, J., 2021. Estimating long run effects and the exponent of cross-sectional dependence: an update to xtdcce2. *The Stata Journal,* 21(3), pp. 687-707. https://doi.org/10.1177/1536867X211045560
18. Eberhardt, M., Helmers, C. and Strauss, H., 2013. Do Spillovers Matter When Estimating Private Returns to R&D? *The Review of Economics and Statistics*, 95(2), pp. 436-448. https://doi.org/10.1162/REST_a_00272
19. Engle, R. F. and Granger, C., 1987. Co-integration and error correction: Representation, estimation, and testing. *Econometrica,* 55(2), pp. 251-276. https://doi.org/10.2307/1913236
20. Everaert, G. and De Groote, T., 2016. Common correlated effects estimation of dynamic panels with cross-sectional dependence. *Econometric Reviews,* 35, pp. 428-463. https://doi.org/10.1080/07474938.2014.966635
21. Fukao, K. [et al.], 2009. Intangible Investment in Japan: Measurement and Contribution to Economic Growth. *The Review of Income and Wealth,* 55(3), pp. 717-736. https://doi.org/10.1111/j.1475-4991.2009.00345.x
22. Griliches, Z. and Mairesse, J., 1998. Production functions: The search for identification. In: S. Steinar, ed. *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*. Cambridge University Press, pp. 169-203.
23. Griliches, Z., 1981. Market value, R&D, and patents. *Economics Letters,* 7(2), pp. 183-187. https://doi.org/10.1016/0165-1765(87)90114-5
24. Griliches, Z., 1995. R&D and productivity: econometric results and measurement issues. In: P. Stoneman, ed. *Handbook of the Economics of Innovation and Technological Change*. Oxford: Blackwell Publishers Ltd.
25. Grossman, G. and Helpman, E., 1991. Quality ladders in the theory of growth. *Review of Economic Studies,* 58(1), pp. 43-61. https://doi.org/10.2307/2298044
26. Hao, J. X., Manole, V. and Van Ark, B., 2009. Intangible Capital and Growth: An International Comparison. *COINVEST Working Paper* D3.6.
27. Hintzmann, C., Lladós-Masllorens, J. and Ramos, R., 2021. Intangible assets and labour productivity growth. *Economies*, 9(2), 82. https://doi.org/10.3390/economies9020082
28. Kripfganz, S. and Schneider, D. C., 2020. Response surface regressions for critical value bounds and approximate p-values in equilibrium correction models. *Oxford Bulletin of Economics and Statistics,* 82(6), pp. 1456-1481. https://doi.org/10.1111/obes.12377.

29. Kripfganz, S. and Schneider, D. C., 2023. ardl: Estimating autoregressive distributed lag and equilibrium correction models. *The Stata Journal*, 23(4), pp. 983-1019. https://doi.org/10.1177/1536867X231212434
30. Nakamura, L. I., 2010. Intangible Assets and National Income Accounting. *The Review of Income and Wealth,* 56(1), pp. s135-s155. https://doi.org/10.1111/j.1475-4991.2010.00390.x
31. Pedroni, P., 2004. Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory,* 20(3), pp. 597-625.
32. Pesaran, M. H., 2004. General diagnostic tests for cross section dependence in panels. *Cambridge Working Papers in Economics,* No. 0435. http://dx.doi.org/10.2139/ssrn.572504
33. Pesaran, M. H., 2007. A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics,* 22(2), pp. 265-312. https://doi.org/10.1002/jae.951
34. Pesaran, M. H., 2015. Testing weak cross-sectional dependence in large panels. *Econometric Reviews,* 34, pp. 1089-1117. https://doi.org/10.1080/07474938.2014.956623
35. Pesaran, M. H., Shin, Y. and Smith, R. J., 2001. Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics,* 16, pp. 289-326. https://doi.org/10.1002/jae.616
36. Romer, P., 1990. Endogenous Technological Change. *Journal of Political Economy,* 98(5), pp. 71-101. https://doi.org/10.1086/261725
37. Roodman, D., 2009. A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics,* 71(1), pp. 135-158. https://doi.org/10.1111/j.1468-0084.2008.00542.x
38. Rosser, J. B. and Rosser, M. V., 2018. *Comparative Economics in a Transforming World Economy.* The MIT Press.
39. Roth, F. and Thum, A.-E., 2013. Intangible Capital and Labor Productivity Growth: Panel Evidence for the EU from 1998-2005. *Review of Income and Wealth*, 59(3), pp. 486-508. https://doi.org/10.1111/roiw.12009
40. Roth, F., 2019. Intangible Capital and Labour Productivity Growth: A Review of the Literature. *Hamburg Discussion Papers in International Economics*, No. 4.
41. Roth, F., 2020. Revisiting intangible capital and labour productivity growth, 2000-2015: Accounting for the crisis and economic recovery in the EU. *Journal of Intellectual Capital*, 21(5), pp. 671-690. https://doi.org/10.1108/JIC-05-2019-0119
42. Solow, R. M., 1987. *We'd better watch out*. New York Times Book Review, 36.
43. Van Ark, B. and O'Mahony, M., 2016. Productivity Growth in Europe before and since the 2008/2009 economic and financial crisis. In: D. Jorgensen, K. Fukao and M. P. Timmer, eds. *The World Economy: Growth Stagnation*. Cambridge: Cambridge University Press, pp. 111-152. https://doi.org/10.1017/9781316534502.004

44. Van Ark, B., 2016. Europe's Productivity Slowdown Revisited: A comparative perspective to the United States. In: P. Askenazy et al., eds. *Productivity Puzzles Across Europe*. Oxford: Oxford University Press, pp. 26-48. https://doi.org/10.1093/acprof:oso/9780198786160.003.0002
45. Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics,* 126(1), pp. 25-51. https://doi.org/10.1016/j.jeconom.2004.02.005
46. WIPO, 2025. *GII 2025 results*.

Table A1

Additional dynamic and static estimations of the equation of effect of IN and K on LP

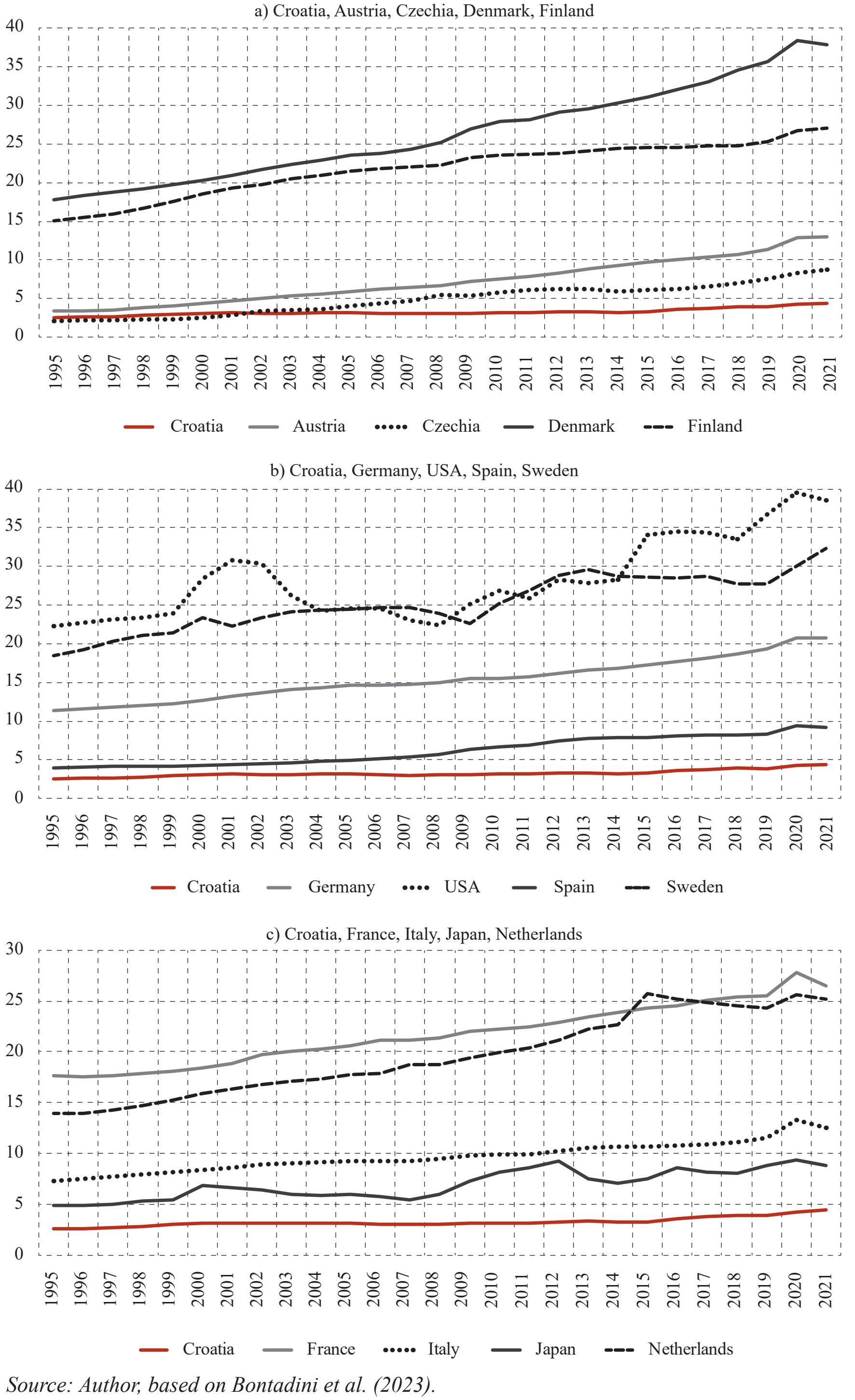
	Data by sector			
	GMM		OLS	
	(1)	(2)	(3)	(4)
LP_{t-1}	0.691*** (0.147)	0.669*** (0.0530)		
K	-0.154 (0.594)	0.00562 (0.0222)	0.438*** (0.0842)	0.191*** (0.00771)
IN	0.211* (0.124)	0.166*** (0.0279)	0.376*** (0.0565)	0.263*** (0.00628)
R^2			0.438	0.509
Instruments	8	10		
AR(1)	0.002	0.000		
AR(2)	0.187	0.172		
Sargan test	0.241	0.093		
Hansen test	0.141	0.116		
Country-sector	180	180	180	180
Observations	1,440	1,440	4,860	4,860

*Note: ***, **, * refers to rejection of the null hypothesis at 1%, 5% and 10% level of significance respectively. Estimated model coefficients are presented (numbers in brackets represent standard errors corrected by Windmeijer's (2005) two-stage variance and covariance matrix corrections for (1) and (2), while clustered robust standard errors are used for (3) and (4)). AR(1) and AR(2) refer to the first- and second-order serial correlation Arellano-Bond test. Sargan and Hansen are tests of preidentified restrictions (the null test hypothesis is the validity of instruments, i.e. conditions per moment, meaning there is no expected link between instruments and relation errors). P-values are shown for the tests.*

R^2 is the coefficient of determination. Models (1) and (2) were estimated based on the system Generalized Method of Moments (GMM) estimator of the dynamic panel model using the method with the expanded instrument variation proposed by Blundell and Bond (1998). Model (3) is an OLS estimator with fixed effects, while (4) is a pooled OLS estimator. For dynamic models (1) and (2), the transformation of forward-orthogonal deviations was used to address the impact of fixed effects. Due to the issue of endogenous shift of the dependent variable in (1) and (2), internal instruments of the dependent variable shift transformed by orthogonal deviations are used for (1), while internal instruments are used for (2) in addition to the standard ones (cumulative physical capital of the sector, total annual hours of work of the sector's employees). For GMM estimators, the collapse option is used to stop over-proliferation of instruments (Roodman, 2009). For models (1) and (2), 3-year averages for the period 1995-2021 in 180 countries-sectors are used to improve the functionality of GMM estimators at the short time component.

Source: Author's own calculations.

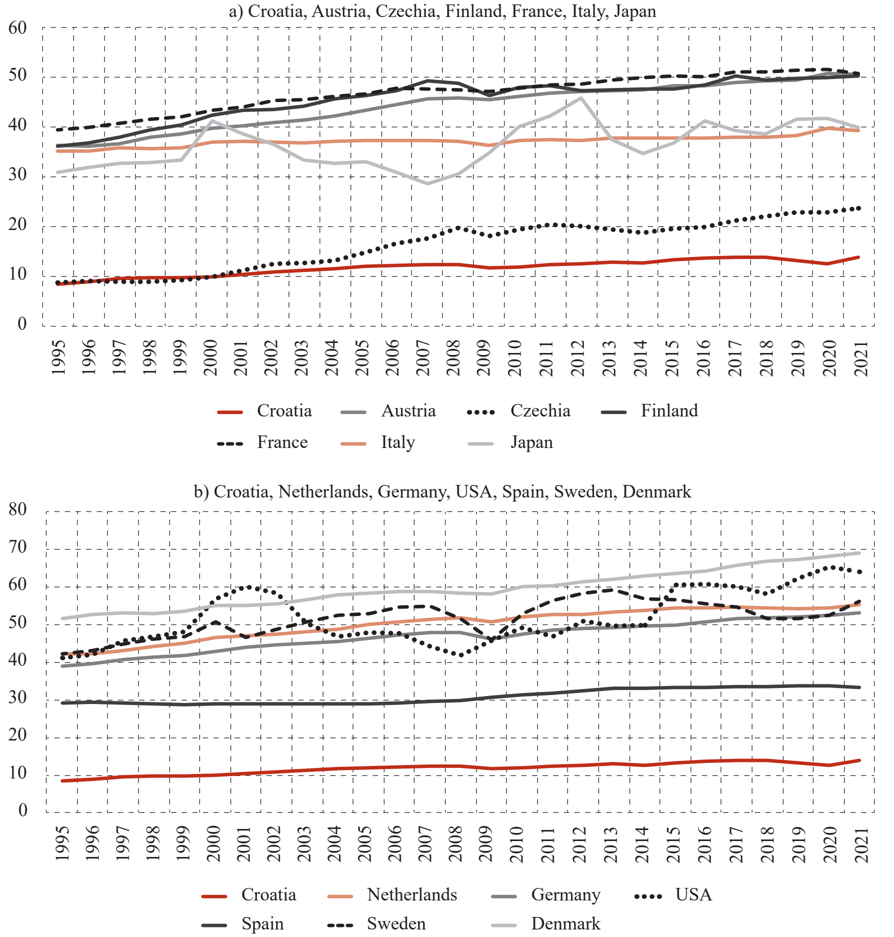
Figure A1
Intangible capital stock per hour of work, in euros and permanent prices (1995-2021), all NACE Rev. 2 activities



Source: Author, based on Bontadini et al. (2023).

Figure A2

Labour productivity measured in gross added value per hour of work, in euros and permanent prices (1995-2021), all NACE Rev. 2 activities



Source: Author, based on Bontadini et al. (2023).

Bridging the gap: socioeconomic inequalities in the use of formal and informal home care

MAJA MATANIĆ VAUTMANS, Ph.D.*

Article**
JEL: I0, H4, R2
https://doi.org/10.3326/pse.50.2.3

* The publication of this paper has been supported by University of Rijeka under the project "Smart cities in function of the development of the national economy" (uniri-drustv-18-255-1424). The author is grateful to two anonymous referees who have contributed to the quality of the final version of the paper.

** Received: June 1, 2025
Accepted: December 15, 2025

Maja MATANIĆ VAUTMANS
University of Rijeka, Faculty of Economics and Business, Ivana Filipovića 4, 51000 Rijeka, Croatia
e-mail: majamatanic5@gmail.com
ORCiD: 0000-0002-5071-2634

PUBLIC SECTOR ECONOMICS
50 (2) 241-257 (2026)

Abstract

This paper examines socioeconomic inequalities in formal and informal home care among Europeans aged 65+, using SHARE data from 27 countries. Results show that health needs – especially age and functional limitations – are the main drivers of care use. However, informal care is more common among lower-income groups, while formal care is concentrated among wealthier and better-educated individuals. Most disparities are explained by horizontal inequities linked to income, education, and marital status, rather than differences in actual care needs. Even after adjusting for health conditions, poorer populations rely more on informal care, indicating unequal access to formal services. The findings underline the need for policies that improve access to formal care, support informal caregivers, and reduce socioeconomic barriers.

Keywords: inequality, socioeconomics, formal care, informal care, SHARE data

1 INTRODUCTION

Europe's rapidly aging population poses profound challenges for long-term care (LTC) systems. As life expectancy rises, more older adults (65+) live with chronic conditions and functional limitations that require daily support. Traditionally, much of this support has been provided in institutional settings, but in recent decades home care has gained increasing importance. This shift is not only a response to demographic pressures and the limitations of institutional care but also reflects active policy efforts to promote the "ageing in place" agenda, which emphasizes enabling older adults to remain in their homes and communities for as long as possible (WHO, 2015; European Commission, 2021). Home care services typically include personal hygiene assistance, meal preparation, household chores, and other forms of daily support. Despite this policy emphasis, access to home care across Europe remains highly uneven. In many countries, eligibility for publicly supported LTC is means-tested, with income or asset assessments determining whether and to what extent older adults qualify for support (Colombo et al., 2011; Rodrigues, Ilinca and Schmidt, 2017; OECD, 2020). These rules often restrict access for low-income households and contribute to reliance on informal caregiving – family members, friends, or neighbours, especially where formal services are limited. Geographic disparities compound this inequality[1], creating what is often described as a "postal code lottery" (Ilinca et al., 2017) where urban residents typically enjoy better access to services than those in rural areas. Research has consistently shown that socioeconomic factors such as income, education, and household composition shape the use of both formal and informal care, and that cross-national differences follow a clear North – South and East – West divide: formal care is more widely available in Northern Europe, while Southern and Eastern regions depend more heavily on informal provision (Haberkern and Szydlik, 2009; Rodrigues, Ilinca and Schmidt, 2017). The present study, however, has analysed inequalities in terms of horizontal equity (whether individuals with equal needs receive equal levels of care) and vertical inequity (whether individuals with greater needs receive proportionally more care) (Vallejo-Torres, 2013). In theory, care provision should increase in line with need, but in practice socioeconomic and contextual factors may distort this

[1] Inequality = what exists (health, age, gender).
Inequity = what is unfair or unjustified, adjusting for need (income, education, living environment, etc.).

relationship. For example, a high-need individual in a low-income rural household may receive less support than a lower-need individual in a high-income urban setting. Understanding vertical inequity therefore requires examining not only the presence of care but whether its *intensity* appropriately reflects differences in need. Existing studies have considered horizontal and vertical inequities separately (Wagstaff and van Doorslaer, 2000; Morris, Sutton and Gravelle, 2005; Rodrigues, Ilinca and Schmidt, 2017), but little work has examined their combined effect. This paper introduces the concept of *total inequality*, the joint consideration of horizontal and vertical inequities – as a way to more comprehensively assess fairness in LTC. The present study has three main aims, to: (1) identify and summarize the key theoretical features of home care systems in Europe, (2) measure vertical inequity in long-term home care using a newly developed methodological approach, and (3) assess horizontal and vertical inequalities in the use of home care among older adults after accounting for need factors.

To address these aims, the following research questions are posed: (1) What is the prevalence of formal and informal home care use among older adults (65+) in Europe? (2) Which socioeconomic factors (income, education, household composition, and living environment) enable or constrain access to home care? (3) To what extent do horizontal and vertical inequalities persist after controlling need for factors such as age, gender, and health status?

This paper draws on data from the Survey of Health, Ageing and Retirement in Europe (SHARE), Waves 8 and 9, covering 26 EU countries and Switzerland. Using regression-based approaches, it explores both horizontal and vertical inequities in access to LTC. By explicitly integrating both dimensions into a measure of total inequality, the study fills a key research gap and provides a more nuanced understanding of how socioeconomic disparities shape home care use across Europe.

The structure of this paper is as follows: the second part presents a literature review. The third part presents an overview of the data and methods that will be used in the research. The research results are presented in the fourth part. Conclusion and discussion of the research are contained in the fifth part of this paper.

2 LITERATURE REVIEW

Socioeconomic inequalities play a central role in shaping access to LTC across Europe, influencing both formal and informal care provision. From a theoretical standpoint, these inequalities can be understood through frameworks such as the Andersen-Newman model, which distinguishes between predisposing factors (e.g., age, gender), enabling factors (e.g., income, education, family support), and need factors (e.g., health status, functional limitations) (Andersen and Newman, 1973). Together, these dimensions provide a lens for understanding why certain groups are more likely to access formal services, rely on informal care, or face gaps in support.

Gender emerges as a particularly salient predisposing factor. Women not only live longer than men on average but also experience higher rates of chronic illness and disability, which increases their demand for care (Jiménez-Martín and Prieto, 2011;

Portrait, Lindeboom and Deeg, 2000). At the same time, gender intersects with household composition and socioeconomic resources. For example, single older women, who are more likely to live alone, face elevated poverty risks (Boudet et al., 2018) and may lack access to family-based informal care, while simultaneously shouldering caregiving responsibilities themselves (OECD, 2017; Da Roit and Le Bihan, 2010). This interplay illustrates how gendered social roles and family structures can amplify inequalities in care utilization. Income and education further shape these dynamics, acting as key enabling factors. Higher-income individuals can afford co-financed formal care, while lower-income households often depend on informal care, incurring significant financial and emotional costs (Theobald, 2003; European Commission, 2021). Income disparities are tightly linked to gender: older women, particularly those who are single or have interrupted work histories, frequently have lower lifetime earnings and reduced pension entitlements, placing them at a disadvantage in accessing formal services. Education reinforces these inequalities by influencing health behaviours, planning for aging, and knowledge of available services (Mirowsky and Ross, 2005; Ilinca et al., 2017). Better educated individuals are not only healthier but also more able to navigate care systems, increasing their likelihood of receiving services proportionate to need.

Health status and age operate as central need factors, interacting with socioeconomic resources to shape care outcomes. Individuals with more severe functional limitations or chronic conditions require greater support, yet evidence suggests that care often fails to match needs, particularly among disadvantaged groups (Rodrigues, Ilinca and Schmidt, 2017). For instance, a low-income older adult with multiple health limitations may receive less formal care than a wealthier individual with fewer needs, highlighting both horizontal and vertical inequities in LTC. Table 1 provides an overview of horizontal (in)equality and vertical (in)equity in the use of formal LTC.

Table 1

Display of horizontal (in)equality and vertical (in)equity in the use of LTC services defined by the number of visits of professional caregivers (formal type of care)

Person	Health status	Number of visits by a formal caregiver			
		A	B	C	D
1	Good	5	5	5	5
2	Good	10	5	10	5
3	Bad	5	10	15	5
4	Bad	10	10	20	5
HI*		Inequality	Equality	Inequality	Equality
VI**		Inequity	Equity	Equity	Inequity

Note: HI shows horizontal (in)equality*[2] *VI** shows vertical (in)equity.*

Source: Author based on Vallejo-Torres (2013).

Only situation B achieves complete fairness in LTC use. It satisfies both components of equity: (a) vertical equity – people with poorer health receive *more* professional

[2] Inequality = what exists (health, age, gender). These factors are unchangeable.
Inequity = what is unfair or unjustified, adjusting for need factors (income, education, living environment, etc.). This can be changed based on life circumstances.

care visits than those in better health; (b) horizontal equality – people with similar levels of need receive the *same* number of visits. Because situation B fulfils both principles at the same time, it is the only scenario that represents full equity in care use.

PUBLIC SECTOR ECONOMICS
50 (2) 241-257 (2026)

Geography and institutional context further compound these inequalities. Northern and Western European countries typically rely more heavily on formal care, supported by higher public expenditure and smaller family structures, whereas Southern and Eastern European countries depend primarily on informal family care due to cultural norms and lower public support (Fernandez and Forder, 2015; Zigante, 2018; Igel et al., 2009). Within countries, urban – rural disparities also affect access, as rural residents often face limited-service availability and greater travel distances. These geographic and institutional factors intersect with socioeconomic and demographic characteristics, amplifying disparities for women, low-income households, and older adults with high care needs.

Taken together, these studies suggest that inequalities in LTC are not the result of any single factor but emerge from the interaction of gender, income, education, health, household composition, and geographic context, mediated by public policy and social norms. Women with lower income and education, living alone in rural areas, are particularly vulnerable to both unmet care needs and reliance on informal provision. Despite this rich evidence, most studies treat formal and informal care separately and focus primarily on horizontal inequity, neglecting vertical inequity – the extent to which care provision scales with need. Addressing this gap requires a framework that simultaneously considers need, enabling resources, and structural context, forming the theoretical foundation for this paper.

3 DATA AND METHODS

This paper uses mixed-methods exploratory sequential design, combining qualitative interviews with quantitative analysis. The quantitative component draws on data from the SHARE survey, a large European panel study of individuals aged 50+ across 28 countries and Israel. The analysis uses cross-sectional data from two waves (2019/2020 and 2021/2022) and focuses on individuals aged 65 and older, with country-level results presented for 26 EU countries and Switzerland.

Formal LTC has been measured as professional support received in the past 12 months. Informal LTC captures non-professional support provided inside or outside the household. The two forms of care are analysed in parallel because they represent distinct components of the care system. Although the two are not treated as substitutes in this study, they are also not fully independent: informal care is more common at lower levels of need, and patterns of use can reflect cultural norms and the availability and capacity of family caregivers (e.g., their employment status, proximity, and other responsibilities). This contextual information is included here only to clarify why the two forms of care are modelled separately yet interpreted with awareness of their potential interaction.

Selected dependent and independent variables related to formal and informal care use are examined to assess inequalities and inequities in LTC utilization across countries and over time. Table 2 presents a list of independent variables used.

PUBLIC SECTOR
ECONOMICS
50 (2) 241-257 (2026)

Table 2

Explanations of independent variables that are used in the examination of inequality and inequity in home care

Independent variables	**Explanations**
Age	It is assumed that older people will be more likely to use formal care since they often live alone (without a spouse or children in their immediate environment), and their health condition becomes increasingly complicated with age, requiring professional care. Age is a quantitative variable that measures the age of the respondents[3] at the time of the survey.
Gender	Women are more likely to be widowed. With a longer life expectancy and having cared for their spouses throughout their lives, they themselves become dependent on other people's help. Since the complexity of health conditions increases with age, it is assumed that they will use formal care. Gender is a binary variable, indicating whether the respondent is female.
Education	Higher levels of education are associated with greater access to information, higher income, and a tendency to choose formal care. Education is a quantitative variable that measures the total number of years of formal education completed by respondents.
Income	Higher income provides greater opportunities and greater use of formal care. Household income is determined by the total average income of all family members. As income has an asymmetric distribution with most respondents having lower values and a smaller number of respondents having high and very high values, the variable will be transformed by its natural logarithm before inclusion in the analyses. This quantitative variable in the paper is marked as Income.
Marital status	Married people are more likely to choose informal care, as one of the spouses can take on the role of informal caregiver. The marital status variable is a binary variable, where the first category includes the original categories: "never married", "divorced", "widows/widowers"; and the second categories: "married, living with spouse", "registered partnership" and "married, not living with spouse".
Children	People living in families with children are more likely to use informal care than people without children. The children variable is a quantitative variable of the number of children of the respondents.
Area	Area is a binary variable of the place where the respondents live, where the original variable had 5 categories that were combined into two: urban and rural.
Health status	Based on the person's own assessment (*self-perceived health* with a scale from 1 – excellent to 5 – poor and given that it takes on 5 different values with an approximately symmetrical distribution, this is used in the analyses as a quantitative variable with numerical values from 1 to 5.
Chronic diseases	Health conditions that last longer than 3 months. Chronic conditions records the number of chronic conditions that the respondent lives with and is also treated as a quantitative variable.

[3] Respondent is the person who is receiving the care (not the informal carer).

Independent variables	Explanations
ADL	Activities of daily living limitations (ADL) refers to six parameters using the Katz index to record a person's state of health and refers to: (1) walking across the room; (2) dressing; (3) bathing or showering; (4) dining; (5) getting on/off the bed; (6) use of the toilet. ADL was measured on scale from 0 to 6.
IADL	Instrumental activities of daily living limitations (IADL) refers to 8 parameters using the Lawtaon index to assess independent functioning in daily life. They relate to: (1) preparing meals; (2) shopping for groceries; (3) making telephone calls; (4) taking medications; (5) managing finances; (6) using transportation; (7) household activities; and (8) dressing and doing laundry. IADL is measured on a scale from 0 to 9 (laundry is a separate category 9).

Source: Author.

Correlation analysis has shown that ADL and IADL are highly correlated, and it can be said that they measure very similar constructs. To avoid the potential problem of multicollinearity in the analyses, by including both variables, a synthetic variable (I)ADL will be created as the average of the values of these two variables, which will be included in further analyses (Gross, Jones and Inouye, 2014; Hong et al., 2019). Since the scales of these variables are different, before calculating the average, the IADL variable will be multiplied by 6/9 to equalize the ranges of the two variables. In this way, the mean values and standard deviations of these variables are approximately equal, so the calculation of the synthetic variable is well founded. The (I)ADL variable was treated as a quantitative variable.

Apart from defining dependent and independent variables to assess inequality in care use, this research distinguishes between: (1) need (legitimate) factors: characteristics that justifiably affect care use, such as health status and functional limitations; (2) non-need (illegitimate) factors: characteristics that should not influence care use, such as income or education, which contribute to inequity if they do.

Inequality is measured using the concentration index (CI), which captures the distribution of care across socioeconomic status (SES). Horizontal inequity (HI) quantifies unequal treatment among individuals with the same needs, while vertical inequity (VI) captures inappropriate differences among individuals with unequal needs. Total inequality is the sum of horizontal and vertical inequity.

For binary care outcomes, nonlinear models (e.g., logistic regression) and marginal effects are applied. SES is measured by equalized household income, adjusted for household size and purchasing power parity (PPP). Sample sizes per country are determined from the SHARE dataset (SHARE-ERIC, 2024a-2024d), with post hoc calculations ensuring sufficient precision for CI estimates.

PUBLIC SECTOR ECONOMICS
50 (2) 241-257 (2026)

3.1 CONCENTRATION INDEX (CI)

Wagstaff, van Doorslaer and Paci (1991) introduced the concentration index (CI) to measure health inequalities. Derived from concentration curves, it illustrates the relationship between a health variable and socioeconomic status (SES). The x-axis ranks SES, while the y-axis shows the cumulative health outcome. CI quantifies inequality, assessing whether health care resources and outcomes are fairly distributed across SES groups. According to O'Donnell et al. (2007), the CI can be calculated as follows:

$$C = \frac{2}{n\mu}\sum_{i=1}^{n} h_i r_i - 1 - \frac{1}{n} \tag{1}$$

where h_i is the health variable, μ its mean value, and $r_i = \frac{1}{n}$ is the relative rank of the respondent i in the SES distribution, where $i = 1$ denotes the poorest and $i = n$ the richest respondent. An equivalent formula for CI uses the covariance:

$$C = \frac{2}{\mu} cov\left(h, r\right) \tag{2}$$

The concentration index (CI) ranges from -1 to 1, with negative values indicating higher prevalence of the health variable among the poor and positive values showing greater prevalence among the wealthy. A negative CI reflects disparities affecting lower socioeconomic groups, while a positive CI highlights inequalities favouring higher-income groups. The concentration index (CI) is influenced by the mean value of the health variable, which can complicate comparisons across populations or countries. Variations in average health outcomes may lead to misleading interpretations of inequality in different contexts. Erreygers (2009) provided a corrected version of the CI for binned variables that does not depend on the mean value of the variable:

$$CC = \frac{4\mu}{a-b} C \tag{3}$$

where a and b are minimum and maximum values of the health variable. The variables of formal and informal care are binary, and for them, $a = 0$ and $b = 1$.

3.1.1 Decomposition of the concentration index

The concentration index can be broken down into the contributions of different factors, which can also be quantified, that is, a CI decomposition can be carried out. The CI decomposition method uses regression models to break down the factor contributions. In its most basic form, where the health variable is quantitative (continuous), a linear regression model is appropriate:

$$h_i = \alpha + \sum_k \beta_k x_{ki} + \varepsilon_i \tag{4}$$

PUBLIC SECTOR ECONOMICS
50 (2) 241-257 (2026)

where x_k represents individual factors, ε_i random errors, and α regression β_k coefficients. Based on this model, the expression for the concentration index can be written:

$$C = \sum_k \frac{\beta_k \overline{x_k}}{\mu} C_k + \frac{GC_\varepsilon}{\mu} \quad (5)$$

where $\overline{x_k}$ is the mean value of x_k, C_k is the (partial) concentration index of the variable x_k with respect to the SES measure, and GC_ε is the generalized concentration index of the random error. The concentration index is the weighted sum of partial concentration indices of independent variables, with weights determined by elasticity. The generalized random error concentration index (residual) represents unexplained inequality. These factor contributions can be positive, indicating a greater representation among the wealthy, or negative, reflecting a higher prevalence among the poor. The contribution's sign depends on the elasticity and partial CI, where matching signs (++ or --) result in a positive contribution, and opposite signs (+- or -+) lead to a negative contribution. The decomposition method can also be applied to the corrected concentration index (Erreygers, 2009), and thus the following is valid:

$$CC = 4\left(\sum_k \beta_k \overline{x_k} C_k + GC_\varepsilon\right) \quad (6)$$

Full formulas, derivations, and technical details for the decomposition of concentration index and inequity calculations are provided upon request.

4 RESULTS

This analysis of older adults (65+) across 26 European countries and Switzerland addresses key questions about the use and distribution of LTC. In response to the first research question, it finds that informal care is generally more prevalent than formal care, with informal care use exceeding 30% in countries such as the Czech Republic, Austria, and several Nordic nations, while formal care ranges from just 1.9% in Romania to 28% in Belgium.[4]

These findings highlight the key theoretical features of LTC in Europe, namely that informal care often supplements formal services and is shaped by cultural norms, family structures, and state support, fulfilling the first aim of this research, to identify and summarize the key theoretical features of home care systems in Europe.

4.1 DECOMPOSITION ANALYSIS

Addressing the second research question, decomposition analyses reveal that socioeconomic factors – income, education, household composition, and living environment – significantly influence care utilization, but differently for formal and informal care. Higher-income and better-educated individuals are more likely to use formal care, particularly in Western Europe, whereas informal care often

[4] Descriptive statistics are available upon request.

PUBLIC SECTOR ECONOMICS
50 (2) 241-257 (2026)

supports lower-income populations, especially in Sweden, Denmark, France, Germany, Spain, Italy, Greece, and Malta. Married individuals generally use slightly less care, rural residents rely more on informal care, and age consistently predicts higher care use. These findings relate directly to the third aim of this research: to assess horizontal and vertical inequalities in the use of home care among older adults after accounting for need factors such as age, gender, and health status.

In terms of horizontal inequity, formal care is largely equitably distributed when health needs are accounted for, whereas informal care remains concentrated among lower-income groups, indicating persistent reliance on family and social networks where formal services are limited. Vertical inequity, which measures whether care allocation is appropriate to varying levels of need, is generally small across Europe, with exceptions such as Sweden and the Czech Republic in the 9th wave, where formal care shows inequities favouring higher-income individuals. The inclusion of vertical inequity only modestly changes total inequality, confirming that horizontal inequities driven by socioeconomic factors are the main contributors to unfair disparities in LTC access, which directly addresses the third and fourth research questions and the second aim of this research: measure vertical inequity in long-term home care using a newly developed methodological approach. Health-related factors remain the strongest predictors of care use, with functional limitations (I)ADL, chronic conditions, and self-rated health increasing both formal and informal care utilization, while socioeconomic and demographic factors contribute additional but smaller effects. The decomposition confirms that older adults with poorer health rely more heavily on care, particularly informal care, supporting existing literature and validating the models used. Country-specific variations, shaped by cultural norms, policy structures, and state support, explain differences in formal and informal care patterns across Northern, Western, Southern, and Eastern Europe.

It is shown that health is the main driver of inequality in LTC, with poorer populations relying more on informal care. Income and education favour wealthier groups in formal care, while informal care mainly supports poorer groups.[5] Regional differences are strong: Western Europe shows higher formal care use, Eastern Europe more informal care. Age increases care needs but often correlates with lower socioeconomic status. Non-health factors (income, education, marital status, rural residence) also shape disparities, though less than health. Overall, poorer populations bear a disproportionate burden, highlighting the need for targeted, equitable LTC policies.

[5] Graphical representations for the concentration indexes (CI) of current care utilization, horizontal, vertical, and total inequality, contributions of individual factors to the concentration indexes, diagrams of elasticity and partial concentration indexes (CI) and contributions of the necessity factor to vertical inequity available upon request for the countries Belgium, Bulgaria, Denmark, Germany, Estonia, Greece, Spain, France, Croatia, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, the Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden, Switzerland and the Czech Republic.

The last part of the decomposition is about vertical inequity. Formal care contributions favouring the poor were strongest in Hungary, Bulgaria, Cyprus, and Luxembourg, mainly linked to health factors. Age generally favoured the rich (notably in Denmark, Sweden, Hungary, Latvia, and Slovakia), but in Estonia and Slovakia it favoured the poor. Gender effects varied by country and wave, with notable contributions in Latvia, the Czech Republic, Austria, Denmark, the Netherlands, and Hungary. Croatia showed minimal contributions, except for health in the 8th wave.

In informal care, SP health reduced contributions while (I)ADL increased them, with strong effects in several European countries. Age showed unstable patterns, and overall vertical inequity was minor and inconsistent, mainly driven by (I)ADL, Age, and SP health. Since vertical inequity explains little of total disparities, policy efforts should focus on horizontal inequity from income, education, and other socioeconomic factors. These results demonstrate that both horizontal and vertical inequities exist, but horizontal inequity driven by socioeconomic differences remains the primary source of disparities, emphasizing the need for policy interventions targeting income, education, and geographic barriers to achieve fairer LTC access.

In summary, the findings answer the research questions posed and fulfil the study's aims by showing the prevalence of formal and informal care, identifying the key socioeconomic determinants, quantifying horizontal and vertical inequities, and providing an empirical methodology to measure vertical inequity, thereby contributing new insights into how LTC use is structured and distributed across Europe.

4.2 LIMITATIONS

This study is subject to several important limitations. First, the analytical sample excludes individuals younger than 65 years. Second, constraints related to sample size, particularly in smaller countries such as Slovakia, Cyprus, and Malta, affect the precision of concentration indices and logistic regression estimates. Vertical inequity analyses were especially sensitive to these limitations, as regression samples were effectively reduced by half and non-essential covariates were removed to maintain appropriate events-per-variable ratios. Consequently, results for countries with low care utilization or small sample sizes may be less reliable and must be interpreted with caution; in cases where reliability is substantially compromised, estimates may be omitted. By contrast, logistic regression models used for horizontal inequity decomposition generally satisfy standard sample size criteria and are therefore considered more robust. Third, indicators of health status rely on participants' subjective self-assessments, which may introduce measurement bias. Finally, the analysis is cross-sectional in nature, restricting the ability to draw causal inferences regarding the relationships between socioeconomic status, health needs, and care use. Future research employing longitudinal designs could more effectively capture the dynamic interactions among these factors.

5 CONCLUSION AND DISCUSSION

This paper analysed horizontal and vertical socioeconomic inequalities in the use of formal and informal home care across 27 European countries, using data from the 8th and 9th waves of the SHARE database (SHARE-ERIC, 2024a-2024d). The research aimed to: (1) identify and summarize the key theoretical features of home care systems in Europe, (2) measure vertical inequity in long-term home care using a newly developed methodological approach, and (3) assess horizontal and vertical inequalities in the use of home care among older adults after accounting for need factors.

The findings show that health-related needs, particularly functional limitations and age, remain the primary determinants of care utilization, consistent with previous studies (Rodrigues, Ilinca and Schmidt, 2017; Ilinca et al., 2017; Bakx et al., 2015). Informal care is disproportionately utilized by individuals with lower incomes and education levels, reflecting its relative accessibility compared to formal care, which is more commonly used by individuals with higher socioeconomic status. This pattern is broadly compatible with earlier work suggesting that formal and informal care may operate as both substitutes and complements, depending on the institutional and family context (Broese van Groenou and De Boer, 2016; Liu, 2021). The relationship between the two types of care is part of a wider theoretical debate on the interaction between public and private transfers originating in Becker's work on altruism and family economics where crowding-out and crowding-in effects may coexist. In many European settings, informal care tends to be more prevalent at lower levels of need or where formal service provision is limited, but this does not necessarily imply that higher informal care mechanically leads to higher formal care use. Rather, existing studies show that substitution and complementarity can occur simultaneously across different levels of need, family structures, and welfare-state arrangements. A more nuanced interpretation is therefore required: while the results of this research align with evidence of partial complementarity, the balance between substitution and complementarity is context-dependent and shaped by both household resources and the availability of publicly funded services.

Horizontal inequity, reflecting inequalities not justified by differences in health needs, was present in several countries, especially in informal care, whereas vertical inequity capturing inappropriate differential treatment based on needs contributed only marginally to total inequities. Interestingly, in some cases, formal care use showed patterns favouring higher-income and better-educated individuals even after adjusting for needs, particularly in Austria, Belgium, and Luxembourg. These results suggest persistent systemic barriers to formal care access, including financial, organizational, and cultural factors, confirming earlier findings from European LTC research (Barbieri and Ghibelli, 2018; Broese van Groenou and De Boer, 2016). The observed concentration of informal care among lower-income groups may reflect both necessity and cultural expectations regarding family-provided care, consistent with prior studies in Southern and Eastern Europe.

The methodological approach, combining concentration indexes with decomposition using logistic regression models, offers additional insights beyond standard logit analyses. The use of the corrected Erreygers CI allowed quantification of the degree of inequality across socioeconomic groups while accounting for the bounded nature of binary care variables, providing a more nuanced picture of disparities in care utilization. The decomposition approach highlighted the relative contributions of health needs versus socioeconomic factors to observed inequalities, confirming that health needs largely drive utilization, while income and education contribute to residual inequities. This reinforces the value of combining traditional regression models with inequality measures to assess LTC access comprehensively.

Several findings warrant particular attention. The strong association of informal care with lower-income groups confirms expected patterns, but the persistence of horizontal inequity in countries with developed LTC systems (e.g., Denmark, Sweden, and the Netherlands) was somewhat surprising and suggests that non-need factors, such as family structure and local availability of services, continue to influence care distribution. Age-related contributions to formal care use were highest in Northern Europe, reflecting generous state-supported LTC programs, whereas in Southern and Eastern Europe, both formal and informal care were more closely tied to income and educational status. These patterns demonstrate that policy context interacts with socioeconomic factors to shape care access.

The study makes several contributions to the literature. It confirms the continued importance of health needs in driving LTC utilization, quantifies the magnitude of socioeconomic inequalities across a broad set of European countries, and demonstrates the utility of combining concentration indexes with decomposition techniques to distinguish between justified and unjustified disparities. By highlighting the persistence of horizontal inequities, particularly in informal care, the research informs policy interventions aimed at reducing socioeconomic barriers, supporting informal caregivers, and ensuring equitable access to both formal and informal care services. Using data from the 8th and 9th waves of the SHARE database (SHARE-ERIC, 2024a-2024d), this paper included, to my knowledge, the largest number of countries analysed with traditionally developed methods for horizontal inequality. In addition, the data from the 9th wave provide a clear overview of the current state of formal and informal care use in the European Union, which is crucial for planning care policies for older people.

In conclusion, the findings of this research underscore the complexity of the interplay of health needs, socioeconomic factors, and policy contexts in shaping home care utilization. Policies that expand access to formal care, provide targeted support for informal caregivers, and consider both health and socioeconomic circumstances are essential for promoting equitable LTC across Europe. Moreover, the methodological approach adopted here offers a replicable framework for assessing inequality and inequity in LTC, which can be applied to future studies and other contexts to monitor progress toward fairer care systems.

Acknowledgements

This paper uses data from SHARE Waves 1, 2, 3, 4, 5, 6, 7, 8 and 9 (DOIs: 10.6103/SHARE.w1.900, 10.6103/SHARE.w2.900, 10.6103/SHARE.w3.900, 10.6103/SHARE.w4.900, 10.6103/SHARE.w5.900, 10.6103/SHARE.w6.900, 10.6103/SHARE.w7.900, 10.6103/SHARE.w8.900, 10.6103/SHARE.w8ca.900, 10.6103/SHARE.w9.900, 10.6103/SHARE.w9ca900) see Börsch-Supan et al. (2013) for methodological details.(1) The SHARE data collection has been funded by the European Commission, DG RTD through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARE-LIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782, SHARE-COVID19: GA N°101015924) and by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, VS 2020/0313 and SHARE-EUCOV: GA N°101052589 and EUCOVII: GA N°101102412. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, BSR12-04, R01_AG052527-02, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-eric.eu).

Disclosure statement

The author has no conflicts of interest to declare.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the author used ChatGPT to assist with summarizing the text. The author reviewed and edited any AI-generated content as needed and take full responsibility for the final manuscript.

REFERENCES

1. Andersen, R. and Newman, J. F., 1973. Societal and individual determinants of medical care utilization in the United States. *Milbank Memorial Fund Quarterly. Health and Society*, 51(1), pp. 95-124. https://doi.org/10.1111/j.1468-0009.2005.00428.x
2. Bakx, P. [et al.], 2015. Going formal or informal, who cares? The influence of public long-term care insurance. *Health Economics*, 24(6), pp. 631-643. https://doi.org/10.1002/hec.3050
3. Barbieri, D. and Ghibelli, P., 2018. *Formal versus informal long-term care: Economic and social impacts*. Zenodo. https://doi.org/10.5281/zenodo.1410379
4. Börsch-Supan, A. [et al.], 2013. Data resource profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, 42(4), pp. 992-1001. https://doi.org/10.1093/ije/dyt088
5. Boudet, A. M. M. [et al.], 2018. Gender differences in poverty and household composition through the life cycle: A global perspective. *Policy Research Working Papers*, No. 8360. https://doi.org/10.1596/1813-9450-8360
6. Broese van Groenou, M. I. and De Boer, A., 2016. Providing informal care in a changing society. *European Journal of Ageing*, 13(3), pp. 271-279. https://doi.org/10.1007/s10433-016-0370-7
7. Colombo, F. [et al.], 2011. Help wanted? Providing and paying for long-term care. Paris: OECD. https://doi.org/10.1787/9789264097759-en
8. Da Roit, B. and Le Bihan, B., 2010. Similar and yet so different: Cash-for-care in six European countries' long-term care policies. *The Milbank Quarterly*, 88(3), pp. 286-309.
9. Erreygers, G., 2009. Correcting the concentration index. *Journal of Health Economics*, 28(2), pp. 504-515. https://doi.org/10.1016/j.jhealeco.2008.02.003
10. European Commission, 2021. Green paper on ageing (COM(2021) 350 final).
11. Fernandez, J. L. and Forder, J., 2015. Local variability in long-term care services: Local autonomy, exogenous influences and policy spillovers. *Health Economics*, 24(S1), pp. 146-157. https://doi.org/10.1002/hec.3151
12. Gross, A. L., Jones, R. N. and Inouye, S. K., 2014. Development of a composite measure of physical functioning for older persons. *Research on Aging,* 37(7), pp. 671-694. https://doi.org/10.1177/0164027514550834
13. Haberkern, K. and Szydlik, M., 2009. State care provision, societal opinion and children's care of older parents in 11 European countries. *Ageing and Society*, 30(2), p. 299-323. https://doi.org/10.1017/S0144686X09990316
14. Hong, H. G. [et al.], 2019. New Composite Measure for ADL Limitations: Application to Predicting Nursing Home Placement for Michigan MI Choice Clients. *Medical Care Research and Review,* 78(4), pp. 413-422. https://doi.org/10.1177/1077558719886735
15. Igel, C. [et al.], 2009. Specialization between family and state intergenerational time transfers in Western Europe. *Journal of Comparative Family Studies*, 40(2), pp. 203-226. https://www.jstor.org/stable/41604275

16. Ilinca, S. [et al.], 2017. Fairness and eligibility to long-term care: An analysis of the factors driving inequality and inequity in the use of home care for older Europeans. *International Journal of Environmental Research and Public Health*, 14(10), 1224. https://doi.org/10.3390/ijerph14101224
17. Jiménez-Martín, S. and Prieto, C. V., 2011. The trade-off between formal and informal care in Spain. *The European Journal of Health Economics*, 13(4), pp. 461-490. https://doi.org/10.1007/s10198-011-0317-z
18. Liu, H., 2021. Formal and Informal Care: Complementary or Substitutes in Care for Elderly People? Empirical Evidence from China. *Sage Open*, 11(2). https://doi.org/10.1177/21582440211016413
19. Mirowsky, J. and Ross, C. E., 2005. Education, cumulative advantage, and health. *Ageing International*, 30, pp. 27-62. https://doi.org/10.1007/BF02681006
20. Morris, S., Sutton, M. and Gravelle, H., 2005. Inequity and inequality in the use of health care in England: an empirical investigation. *Social Science & Medicine*, 60(6), pp. 1251-1266. https://doi.org/10.1016/j.socscimed.2004.07.016
21. O'Donnell, O. [et al.], 2007. *Analyzing health equity using household survey data*. Washington: The World Bank. https://doi.org/10.1596/978-0-8213-6933-3
22. OECD, 2017. *The pursuit of gender equality: An uphill battle*. Paris: OECD. https://doi.org/10.1787/9789264281318-en
23. OECD, 2020. *Affordability of long-term care services among older people in the OECD and the EU*. Paris: OECD. https://doi.org/10.1787/f066a74c-en
24. Portrait, F., Lindeboom, M. and Deeg, D. J. H., 2000. The use of long-term care services by the Dutch elderly. *Health Economics*, 9(6), pp. 513-531. https://doi.org/10.1002/1099-1050(200009)9:6<513::AID-HEC534>3.0.CO;2-R
25. Rodrigues, R., Ilinca, S. and Schmidt, A. E., 2017. Income-rich and wealth-poor? The impact of measures of socio-economic status in the analysis of the distribution of long-term care use among older people. *Health Economics,* 27(3), pp. 637-646. https://doi.org/10.1002/hec.3607
26. SHARE-ERIC, 2024a. Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8 (Release version 9.0.0) [Data set]. https://doi.org/10.6103/SHARE.w8.900
27. SHARE-ERIC, 2024b. Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8 COVID-19 Survey 1 (Release version 9.0.0) [Data set]. https://doi.org/10.6103/SHARE.w8ca.900
28. SHARE-ERIC, 2024c. Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 9 (Release version 9.0.0) [Data set]. https://doi.org/10.6103/SHARE.w9.900
29. SHARE-ERIC, 2024d. Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 9 COVID-19 Survey 2 (Release version 9.0.0) [Data set]. https://doi.org/10.6103/SHARE.w9ca.900
30. Theobald, H., 2003. Care for the elderly: Welfare system, professionalization and the question of inequality. *International Journal of Sociology and Social Policy*, 23(4-5), pp. 159-185. https://doi.org/10.1108/01443330310790561

31. Vallejo-Torres, L., 2013. *An economic analysis of vertical equity in the delivery of health care in England*. Doctoral dissertation, University of York.
32. Wagstaff, A., van Doorslaer, E. and Paci, P., 1991. On the measurement of horizontal inequity in the delivery of health care. *Journal of Health Economics*, 10(2), pp. 169-205. https://doi.org/10.1016/0167-6296(91)90003-6
33. Wagstaff, A. and van Doorslaer, E., 2000. Income inequality and health: What does the literature tell us? *Annual Review of Public Health*, 21, pp. 543-567. https://doi.org/10.1146/annurev.publhealth.21.1.543
34. WHO, 2015. *World report on ageing and health*. Geneve: World Health Organization.
35. Zigante, V., 2018. *Informal care in Europe: Exploring formalization, availability and quality*. European Commission, Directorate-General for Employment, Social Affairs and Inclusion. https://doi.org/10.2767/78836

Analysing the impact of migration flows on regional per capita GDP in Türkiye: a spatial panel data approach

ALİ OSMAN ÖZTOP, Ph.D.*
TUNA KÖSE, Ph.D.*

Article**
JEL: F22, R11, R23
https://doi.org/10.3326/pse.50.2.4

* The authors would like to thank two anonymous referees for their comments and suggestions.

** Received: August 28, 2025
Accepted: December 21, 2025

Ali Osman ÖZTOP
Department of Economics and Finance, Muğla Sıtkı Koçman University, Fethiye, 48300 Muğla, Türkiye
e-mail: aliosmanoztop@mu.edu.tr
ORCiD: 0000-0002-7568-5927

Tuna KÖSE
Independent Researcher, Muğla, Türkiye
e-mail: tunakose48@gmail.com
ORCiD: 0000-0002-8703-0215

Abstract

This paper investigates the effects of net international and internal migration on per capita GDP in Türkiye's 81 provinces. Unlike prior research, it analyses both migration types within the same spatial panel model and accounts for spatial dependence. The first model (2016-22) examines overall migration effects, the second (2009-22) focuses on the effects of migrants' educational levels. We find that net international immigration significantly boosts per capita, GDP, with highly educated migrants contributing more to growth. The robustness checks further highlight the complementary role of financial deepening in boosting regional economic performance.

Keywords: internal migration, international migration, financial inclusion, per capita GDP, spatial panel data analysis

1 INTRODUCTION

Migration is a complex phenomenon influenced by – and at the same time influencing – many economic, social, cultural and political developments. Over recent years, international migration has gained considerable traction in public and political debates around the world. However, it is important to remember that migration not only takes place internationally, i.e., across national borders, but also internally, i.e., between different regions of a given country. The literature typically focuses on one or the other form of migration: most often, human capital growth models are used to study international migration, and convergence models are used for internal migration. This paper looks at the effects of both international and internal migration on regional economic growth in Türkiye.

Due to its location at the nexus between Asia and Europe, Türkiye has long been a significant transit and destination state for migrants. In recent years, the country has experienced a rise in international migration: the number of international migrants increased from around 381,000 in 2016 to 494,000 in 2022 (TurkStat), amounting to approximately 0.6% of the Turkish population. This increase was mostly a result of the war in Ukraine – the number of migrants from Russia and Ukraine increased from around 13,000 in 2016 to 132,000 in 2022 (TurkStat) – and emigration from the Middle Eastern countries. This is a relatively modest number compared to some other countries: Poland, for example, had migration equal to 2.7% of its population in 2022. According to Eurostat, there were 2 million Ukrainians in Poland before the war in Ukraine, whereas 5.4 million came after the start of the war in 2022, to a population of 36 million. In 2023, the number of new migrant arrivals in Spain was approximately 2.5% of the total population, while in Greece it was just over 1%. While the change in immigration to Türkiye has been less pronounced than elsewhere, this is precisely what makes it a good case study: immigration numbers have actually remained relatively stable, and there are unlikely to be outsize effects from irregular migrant flows.

Given the scale of international and internal migration, it is natural to ask what its macroeconomic effects have been. Most of the literature considers the effects of international and internal migration separately, at aggregate country level, and does not distinguish between different types of migrants. We consider the effects of both types of migration simultaneously, and study them at the regional level, using a unique dataset covering all of Türkiye's 81 provinces. This setup provides an opportunity to analyse the relationship between migration and economic growth in greater depth. Moreover, we look at how migrants with different education levels affect regional growth.

Empirical studies directly testing this effect are rare (Chen, 2006; Orefice, 2010; Di Maria and Lazarova, 2012; Boubtane, Dumont and Rault, 2016; Șerban et al., 2020). We believe that this study is among the first to analyse the effect of settled migrants' educational levels on per capita GDP using the spatial panel data approach.

Our main finding is that migration contributes positively to Türkiye's regional economic performance, although the magnitude of the effect differs by migrant type. Specifically, international migration has a small but positive impact on per capita GDP, while internal migration driven by educated individuals significantly enhances regional income levels. In contrast, less educated internal migrants have an insignificant effect, implying that their movement neither hampers nor improves regional per capita income. These results highlight the importance of human capital in shaping the economic impact of migration.

To check the robustness of our baseline result, we added a lagged dependent variable (to account for trend population growth) and indicators of financial inclusion such as population per bank branch and per capita bank deposits, which have been shown in the literature to affect per capita GDP growth (e.g. Kim, Yu and Hassan, 2018; Sonkar and Sarkar, 2020; Kanga et al., 2022; Afonso and Blanco-Arana, 2024). The robustness checks confirm the stability of our main empirical results. The inclusion of a lagged dependent variable, time trend, and financial inclusion indicators does not alter the direction or significance of estimated relationships. Moreover, when spatial dependencies are accounted for through spatial autoregressive models with additional autoregressive error structure (SARAR), the positive link between migration and per capita GDP remains intact. This consistency across model specifications suggests that our baseline findings are robust to variations in model design.

The remainder of this paper is structured as follows. Section 2 reviews the literature on migration and growth. Section 3 describes the data and lays out the estimating framework. Section 4 presents the estimation results. Section 5 discusses the robustness of our baseline model specification. Section 6 concludes.

2 LITERATURE REVIEW

The relationship between economic development and migration has long been regarded as a foundational yet intricate area of inquiry. This complexity stems from different ways in which migration impacts economies, including labour market dynamics, human capital accumulation, and spatial interactions. To provide an overview of the existing empirical research, studies examining the relationship between migration and economic growth are summarised in appendix table A1. This summary table serves as a framework for positioning the present study within the broader literature. As seen in the table, most existing studies analyse either international or internal migration separately and generally focus on country-level data. To our knowledge, there are no studies that have addressed both dimensions simultaneously or examined how migrants' characteristics influence regional economic outcomes – an area to which this study contributes. Moreover, studies employing the spatial panel data analysis method, which accounts for spatial dependence, are quite limited (e.g. Vakulenko, 2014; Manavgat and Saygılı, 2016; Yürük and Batmaz, 2023).

In the broader study of immigration, research dating back to Ravenstein (1885) both explores and explains the complexity of international and internal migration. More recently, Greenwood (1975; 1997) examined internal migration in advanced economies, Lucas (1997) addressed internal migration from the perspective of developing economies, Borjas (1999) analysed the phenomenon of migration on an international scale, and Etzo (2008) surveyed the literature on internal migration.

In studies of migration and economic growth specifically, three dominant frameworks have emerged: early Solow-Swan growth models, human capital models, and convergence models. The early models treated migration as an extended version of the traditional Solow-Swan growth model with a production technology that features declining marginal productivity of labour. In such models, population growth is traditionally associated with negative effects on per capita productivity and per capita income growth (see e.g. Dolado, Goria and Ichino, 1994).

Human capital and convergence models both evolved from the Solow-Swan growth model, but they have been typically used in different contexts: the former is favoured in studies of international migration, and the latter in those of internal migration. This is because international migration often implies a skill or human capital differential between migrants and locals, whereas internal migration processes are often understood as a rebalancing or convergence between regional labour markets and wages. For clarity, this literature review will first look at research on international and internal migration separately, with the latter focussing on Türkiye, and then identify some of the gaps in the literature.

2.1 INTERNATIONAL MIGRATION

There is little consensus on whether international migrants contribute to the economic growth of the host country. Some evidence is positive: Mariya and Tritah (2009) found that migration generally led to an increase in income of the host country, and Docquier, Peri and Ruyssen (2014) established a positive correlation between GDP growth and migration rates. Others showed more modest effects: Bashier and Siam (2014) found that migrants had a positive but insignificant effect on real GDP, and Öztop, Köse and Ünlü (2024) identified a statistically significant yet quantitatively small effect of foreign nationals arriving in Türkiye on province-level GDP. Another set of studies suggests the existence of a negative relationship between net migration and economic growth (Güner and Yalınız, 2013; Sevinç et al., 2016).

When account is taken of factors such as skill level, time, and age distribution, the picture becomes even more varied. Some models found that skilled migration flows had a less negative impact on per capita income growth than overall population growth (Barro and Sala-i-Martin, 2004; Dolado, Goria and Ichino, 1994). Boubtane, Dumont and Rault (2016), Orefice (2010) and Şerban et al. (2020) measured migrants' human capital based on their educational attainment and found it to be positively correlated with per capita GDP growth in host countries. Kang and Kim (2018) highlighted the role of migration from developed economies in enhancing growth of developing countries through the transfer of advanced knowledge and skills. But Chen (2006) and Di Maria and Lazarova (2012) found that skilled labour migration, regardless of origin, negatively affected economic growth of the host country.

With regard to the time dimension, Brunow, Nijkamp and Poot (2015) concluded that net migration helped raise growth in the long run in both developed and developing economies, but they could not detect a statistically significant relationship between migration and growth in the short run. By contrast, Ortega and Peri (2009) showed that migration increased the overall GDP level of the host country in the short run and was neutral in the long run.

In a study that considered the age of migrants, Aleksynska and Tritah (2015) found that a high proportion of young immigrants had a negative impact on host country per capita income, whereas a high proportion of older immigrants had a positive effect.

Another strand of the literature, often using the cointegrating framework, focused on the direction of causality between migration and growth, as the decision to emigrate was shown to be largely driven by differences in GDP levels (Draženović, Kunovac and Pripužić, 2018). Felbermayr, Hiller and Sala (2010) and Göv and Dürrü (2017) thus claimed that the causality ran from migration to growth; Boubtane, Coulibaly and Rault (2013a) and Morley (2006) that it ran from growth to migration; while Boubtane, Coulibaly and Rault (2013b) and Altunç, Uçan and Akyıldız (2017) emphasised a positive bidirectional relationship.

Our study addresses both internal and international migration, and distinguishes between domestic and foreign citizens among international migrants, as well as between skilled and unskilled internal migrants. Using a spatial panel data rather than a cointegrating framework we assess how these different types of migrants affect regional per capita GDP. This perspective complements the predominantly country-level analysis and contributes to a better understanding of the spatial dimension of migration and economic growth.

2.2 INTERNAL MIGRATION

Studies examining the relationship between internal migration and economic growth have also been inconclusive. Theoretical models generally predict that a migrant will move from low- to higher-wage regions, or to regions where the expected wage is higher (Bauer and Zimmermann, 1999).

The empirical work has mostly relied on convergence models to explain the effects of migration on growth. In early contributions, Barro and Sala-i-Martin (1992; 2004) found that net migration had a positive but insignificant effect on growth. This finding was echoed in several other studies (Bayraktar and Özyılmaz, 2019; Foldvari, van Leeuwen and van Zanden, 2013; Huber and Tondl, 2012; Özgen, Nijkamp and Poot, 2010; Peri, 2010). Studies of spatial interactions in migration, triggered for example by income differentials and labour market conditions in neighbouring regions, have been less common (see e.g. LeSage and Fischer, 2009).

When looking at Türkiye specifically, Özyılmaz and Bayraktar (2021) established that net internal migration increased growth in all of its NUTS-2 regions. Manavgat and Saygılı (2016) found that, as surplus labour moved from lower-productivity provinces, per capita income increased both in the provinces the migrants left and in those to which they relocated. Ulucan (2022) also noted that increases in per capita income in a given province contained emigration from that province and attracted a significant immigration from other provinces. This finding supports the view that income differences are the primary driver of internal migration, while demographic structure and education level of workers explain in addition which provinces receive or release domestic migrants. Yürük and Batmaz (2023) found that there was no statistically significant relationship between the extent of internal immigration and economic growth, but provinces that lost part of their workforce experienced a significant slowdown in growth. As in the case of international migration, it is important consider the direction of causality: there are studies suggesting that economic growth leads to internal migration (Bunea, 2012; Ghatak, Mulhern and Watson, 2008; Zhang and Song, 2003) rather than the opposite.

In contrast to these studies, Kırdar and Saraçoğlu (2012) found that internal migration had a negative effect on regional growth in Türkiye. Huber and Tondl (2012) found a similar negative long-term effect on per capita GDP of migrant-sending regions in the EU. Vakulenko (2014) looked at regional migration and per capita income differences in Russia and found a negative relationship between these two variables. Ortega (2008) reported that that regional migration in Spain slowed per capita income growth. Borozan (2017) found that the effect of net migration on long-term economic growth in Croatia was insignificant.

The level of human capital was found to be quite important in explaining the relationship between migration and growth. Several studies found a statistically significant and positive relationship between the share of highly educated settled migrants and per capita GDP (Akbari and Haider, 2018; Özyılmaz and Bayraktar, 2021). Pouliakas et al. (2009) found that migration of highly skilled workers led to productivity losses in certain regions in Greece, while migration of unskilled workers had a positive effect on regional GDP.

Building on these insights, our research explores how internal migration interacts with regional economic performance once spatial dependence and spillover effects – i.e. the influence of migration and economic conditions in one region on growth outcomes of geographically neighbouring regions – are explicitly modelled. While previous studies have generally relied on non-spatial models or national-level data, our study incorporates spatial interactions, regional data, as well as the education level of migrants.

3 DATA AND EMPIRICAL FRAMEWORK

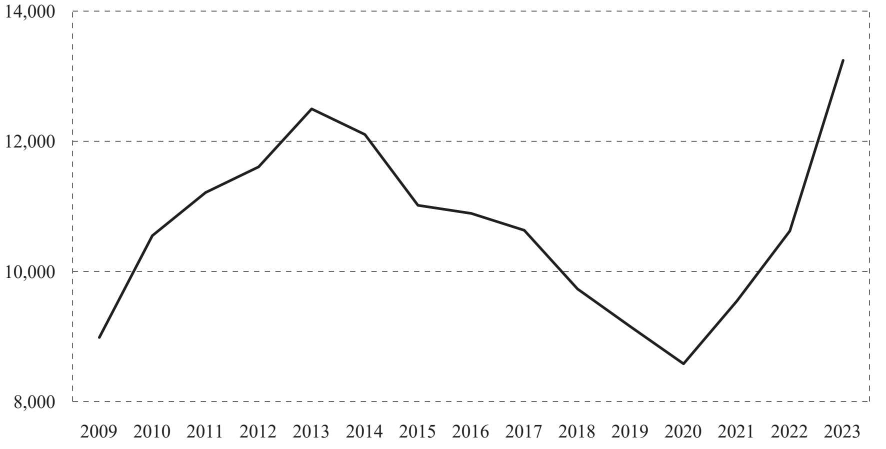
The dataset covers two periods, 2009-22 and 2016-22, for which the most comprehensive migration data are available and when intensive migration flows into and within Türkiye took place. The migration data comprise net international and net internal migration; the latter is also broken down by education status for all of Türkiye's 81 NUTS-3 (Nomenclature of Territorial Units for Statistics) provinces. The migration data were obtained from the Turkish Statistical Institute (TÜİK). Net internal migration is defined as the difference between the in-migration and out-migration in a province, while net international migration is defined analogously for the country as a whole. Since the data are regularly compiled from the address-based population registration system, foreigners with short-term visas or residence permits, as well as Syrians present in the country with temporary protection status, are not considered.[1]

We use the level of per capita GDP in US dollar terms as the dependent variable. Figure 1 shows the evolution of per capita GDP at the country level.

[1] According to the Presidency of Migration Management of Türkiye (2024), around 2.5 million Syrians were living in the country in 2023 with temporary protection status. In addition, over 75,000 Syrians had residence permits, while about 239,000 people of Syrian origin had acquired Turkish citizenship. In addition, around 1.1 million foreigners had short-term residence permits. These groups add up to some 3% of the total population of Türkiye but are not included in the statistics of the address-based population registration system.

Figure 1

Republic of Türkiye per capita GDP in US dollars, 2009-23



Source: Turkish Statistical Institute; authors' estimates.

We checked the stationarity of per capita GDP levels with unit root tests at both provincial and country levels. The augmented Dickey-Fuller (ADF) test, incorporating a deterministic trend and employing the Akaike Information Criterion (AIC) for lag selection, showed that within the 2009-22 panel, a trend was present in 8 of the 81 provinces at the 5% significance level; in the remaining 73 provinces there was no statistically significant trend. For the shorter 2016-22 subsample, only 5 provinces had a significant trend, with the other 76 exhibiting none. The cross-sectionally augmented Im, Pesaran, and Shin test (CIPS) test developed by Pesaran (2007) indicated that the panel dataset was stationary in first differences for both subsamples (with $p < 0.01$). Nevertheless, we checked for robustness of the baseline specification in levels by adding a lagged per capita GDP level and time trend to the regressions (see table 3).

Our first model covers both international (*INM*) and internal net migration (*IM*), with net migration measured as a percentage of the population in the country or a province period *t*. Among people migrating internationally, we consider separately both domestic (INM^D) and foreign (INM^F) citizens. For the data set containing observations on Turkish provinces, denoted $i = (1, ..., N)$, assumed to be observable over the entire period, $t = (1, ..., T)$ (with no attrition or missing data), our first model can be represented as:

$$ln(GDPpc)_{it} = \beta_0 + \beta_1 IM_{it} + \beta_2 INM_{it}^D + \beta_3 INM_{it}^F + v_{it} \tag{1a}$$

where the composite error term $v_{it} = \mu_i + \lambda_t + \varepsilon_{it}$ includes time-invariant province-specific effects (μ_i), time-specific effects capturing common macroeconomic shocks affecting all provinces simultaneously (λ_t), and an error term ε_{it} assumed to be independently and identically distributed with zero mean and constant variance σ_i^2.

We test the following three hypotheses:

H1: $\beta_1 > 0 \rightarrow$ Internal net migration positively affects the per capita GDP of a province.

H2: $\beta_2 > 0 \rightarrow$ International net migration of the domestic population has a positive effect on the per capita GDP of a province.

H3: $\beta_3 > 0 \rightarrow$ International net migration of foreign nationals has a positive effect on the per capita GDP of a province.

Equation (1a) describes the overall effect of net migration on per capita GDP. Other things equal, net migration inflows increase labour supply and raise potential output and hence per capita GDP (Mariya and Tritah, 2009; Docquier, Peri and Ruyssen, 2014; Özyılmaz and Bayraktar, 2021). However, as internal migration in Türkiye comprises mainly low-skilled rural-urban movements, it might also affect per capita GDP of a province negatively (Kırdar and Saraçoğlu, 2012). The sign of the β_1 coefficient is thus a priori indeterminate.

Similarly, the signs of β_2 and β_3 coefficients are *a priori* indeterminate as they depend, among other factors, on the skills composition of domestic workers leaving and foreign workers entering Türkiye. For different samples, Kang and Kim (2018), for example, estimated these coefficients to be positive, while Chen (2006) and Di Maria and Lazarova (2012) estimated them to be negative.

In the second model we analyse internal migration. We consider separately internal migrants with at least a bachelor's degree (IM^B) and those without a bachelor's degree (IM^{NB}):

$$ln\left(GDPpc\right)_{it} = \alpha_0 + \varnothing_1 IM_{it}^{NB} + \varnothing_2 IM_{it}^{B} + u_{it} \tag{1b}$$

where u_{it} is the composite error term $u_{it} = \mu_i + \lambda_t + \varepsilon_{it}$ defined analogously to equation (1a).

For this model, we test two additional hypotheses:

H4: $\varnothing_1 < 0 \rightarrow$ Low-skilled net internal migration negatively affects the per capita GDP of a host province.

H4: $\varnothing_2 > 0 \rightarrow$ High-skilled net internal migration positively affects the per capita GDP of a host province.

In the empirical literature there is a somewhat greater consensus on the expected signs of coefficients $\varnothing_1$ and $\varnothing_2$: net inflow of low-skilled workers is generally estimated to dampen per capita GDP of the host province (e.g. Özyılmaz and Bayraktar, 2021), while net inflow of high-skilled workers is generally estimated to raise the per capita GDP of the host province through the positive effect on its potential output (e.g. Akbari and Haider, 2018).

3.1 SPATIAL PANEL MODEL

Spatial panel models help exploit both observable and unobservable heterogeneity among provinces. The former comprises differences in per capita GDP levels, the latter, e.g. differences in spatial structure such as the nature of provincial borders, location and distance of settlements, transportation and agglomeration infrastructures, industrial structures, etc. These characteristics influence migration decisions in origin regions and result in variability in regression coefficients, which could result in biased estimates if not accounted for, as e.g. in non-panel or non-spatial settings used in cross-sectional regressions (Anselin, Gallo and Jayet, 2008).

The use of the spatial panel model also improves the accuracy of estimates for larger sample sizes. As spatial dependence increases in large sample sizes, test power decreases and rejection rates in right-tailed tests approach zero. The spatial panel model prevents this loss of power by accounting for effects of neighbouring units (Kang, 2018). Spatial panel models also allow a better modelling of dynamic relationships across multiple time periods. For example, Pu et al. (2019) demonstrate that interprovincial migration flows in China exhibit strong temporal dependence and spatial spillovers. They use a spatial dynamic panel model to decompose migration effects over contemporaneous, short-run, and long-run periods, showing how spatial panel models can be efficient in capturing dynamic relationships over time.

4 ESTIMATION RESULTS

To assess the impact of migration dynamics on per capita GDP in Türkiye, two spatial panel data models are estimated. Panel A analyses the effects of internal migration (*IM*), international domestic and foreign migration (INM^D, INM^F) on the per capita GDP of Türkiye's 81 provinces. Panel B disaggregates internal migration by education level, examining the effects of migrants with higher (IM^B) and lower education (IM^{NB}) on per capita GDP at the provincial level. All variables associated with internal and international migration are expressed in terms of net migration. Descriptive statistics and definitions of variables are shown in appendix table A2.

Table 1 presents estimates for pooled data, fixed effects (within), and random effects models without considering spatial autocorrelation. These estimates serve as a diagnostic step in analysis, enabling us to assess whether ignoring spatial dependence introduces a significant bias in estimates.

Intuitively, the pooled data model treats the entire sample as a single cross-section repeated over time and estimates the *average* effect of migration on per capita GDP (equations (1a) and (1b)). The fixed effects model eliminates any fixed differences between provinces via province-specific intercepts α_i, which are potentially correlated with regressors, and examines only changes over time, assuming that $\varepsilon_{it} \sim N\left(0,\sigma^2\right) i.i.d.$:

$$\ln\left(GDPpc\right)_{it} = \beta_{10} + \beta_{20} IM_{it} + \beta_{31} INM_{it}^{D} + \beta_{32} INM_{it}^{F} + \alpha_i + \varepsilon_{it} \tag{2a}$$

$$\ln\left(GDPpc\right)_{it} = \alpha_0 + \alpha_1 IM_{it}^{NB} + \alpha_2 IM_{it}^{B} + \alpha_i + \varepsilon_{it} \tag{2b}$$

The random effects model treats the unobserved heterogeneity as random effects ($\varepsilon_{it} \sim N\left(0,\sigma^2\right) i.i.d.$) drawn from a larger sample, enabling efficient estimates that leverage both cross-sectional and time-series variation in data:

$$\ln\left(GDPpc\right)_{it} = \beta_{10} + \beta_{20} IM_{it} + \beta_{31} INM_{it}^{D} + \beta_{32} INM_{it}^{F} + \alpha + u_{it} \tag{3a}$$

$$u_{it} = \alpha_i + \varepsilon_{it} \tag{3b}$$

$$\ln\left(GDPpc\right)_{it} = \alpha_0 + \alpha_1 IM_{it}^{NB} + \alpha_2 IM_{it}^{B} + \alpha + u_{it} \tag{3c}$$

$$u_{it} = \alpha_i + \varepsilon_{it} \tag{3d}$$

Panel A estimates suggest that internal migration (*IM*) has a statistically significant and quantitatively large positive effect on per capita GDP in both pooled and random effects models. This result aligns with findings of Barro and Sala-i-Martin (1992, 2004) and the Türkiye-specific findings of Özyılmaz and Bayraktar (2021). However, when country-specific unobservable heterogeneity is controlled for in the fixed effects model, the significance of internal migration disappears, suggesting that unobserved, time-invariant characteristics of provinces such as infrastructure or industrial structure rather than migration may account for the effect on per capita GDP.

Table 1
Estimates without consideration for spatial autocorrelation

Variables	Pooled data	Fixed effects (within)	Random effects (GLS)
Dependent variable: ln(GDPpc)			
Panel A: 2016-22[1]			
Intercept	8.8324*** (0.0150)		8.8548*** (0.0310)
IM	0.0491*** (0.0050)	0.0040 (0.0020)	0.0056** (0.0020)
INM^D	-0.0002 (-0.0120)	0.0008 (0.0050)	0.0008 (0.0060)
INM^F	0.2565*** (0.0400)	0.0504 (0.0150)	0.0576*** (0.0160)
R^2 adjusted	0.1907	-0.1407	0.0261
Panel B: 2009-22[2]			
Intercept	8.9508*** (0.0100)		8.9429*** (0.0270)
IM^{NE}	0.1006*** (0.0080)	-0.0094** (0.0040)	-0.0054 (0.0040)
IM^B	0.1281*** (0.0280)	0.1030*** (0.0160)	0.0999*** (0.0160)
R^2 adjusted	0.1397	-0.0382	0.0297

[1] *For the 2016-22 period, 567 observations.* [2] *For the 2009-22 period, 1,134 observations.*

*Note: *, ** and *** denote 10, 5 and 1% statistical significance levels, respectively. The values in parentheses are standard errors.*

International migration of domestic population (INM^D) does not seem to have any statistically significant effect on per capita GDP of provinces. By contrast, international migration of foreign residents (INM^F) is estimated to have a strong and statistically highly significant positive effect in both pooled data and random effects models. But as with internal migration, this effect disappears in the fixed effects specification, indicating that unobserved time-invariant characteristics of provinces are probably more important determinants of per capita GDP than foreign resident immigration.

Panel B introduces information on the educational level of internal migrants, enabling an assessment of the impact of migrants' human capital attainment on regional per capita GDP. The estimates substantiate the role of education: the coefficient on internal migrants with at least a bachelor's degree (IM^B) is positive and statistically highly significant across all three specifications, corroborating the view that better educated internal migrants boost the per capita GDP of host provinces, as documented by Akbari and Haider (2018) and Özyılmaz and Bayraktar (2021).

Internal migrants with only secondary or primary level education (IM^{NB}) seem to contribute to the per capita GDP of host provinces in the pooled data model, lower it in the fixed effects model and are statistically insignificant in the random effects model. This sign instability suggests that the contribution of migrants with lower educational attainments is highly sensitive to, and often outweighed by, fixed province-specific factors. Negative values of adjusted R^2 in the fixed effects model arise from inclusion of many province-specific intercepts (α_i).

To ascertain whether the fixed or random effects specification offers a superior fit for the data, we performed a Hausman specification test (appendix table A3). The test results consistently reject the random effects specification at a 1% significance level for both panels (Panel A: $\chi^2 = 14.89$, $p < 0.01$; Panel B: $\chi^2 = 17.82$, $p < 0.01$). In what follows, our spatial panel models therefore use the fixed effects specification.

The Lagrange multiplier (LM) test statistics for spatial autoregressive (SAR) and spatial error (SEM) models have significant values for all specifications (appendix table A3). The more precise robust LM statistics for spatial lag and spatial error further confirm the presence of spatial dependence in the data across specifications, indicating that the omission of spatial effects would lead to biased and inefficient estimates.

We also conducted Moran's I spatial autocorrelation analysis for per capita GDP to assess directly the degree of spatial dependence (appendix table A4). The test shows significant positive spatial autocorrelation with extremely low p-values, indicating a strong tendency for similar per capita GDP values to cluster spatially.

The SEM approach is often preferred over the SAR approach in method selection (Salima, Julie and Lionel, 2018). However, given that the null hypothesis of no spatial dependence is rejected by robust LM spatial lag and spatial error tests (table A3), as well as Moran's I spatial autocorrelation test (table A4), we proceeded to estimate the spatial autoregressive lag and error (SARAR) model that incorporates both types of spatial interaction.

We estimate the pooled data and fixed effects models with spatial autocorrelation of errors and spatial autoregressive coefficients using several spatial panel econometrics techniques: SEM models that account for spatial correlation in the error term; generalised method of moments (GMM) models that control for fixed effects; and the SARAR models, which provide the most comprehensive framework for interpretation of estimates, as they simultaneously address interactions among provinces and unobserved spatial linkages. Specifically, the model based on Lee and Yu (2010a) incorporates both spatial lag and ensures consistent estimators and centred distributions.[2]

Table 2 presents estimates of the pooled data and fixed effects models with spatial parameters. A key result across both panels is the highly significant and large values of the spatial error ρ and the spatial autoregressive coefficient λ. In SARAR models, estimated coefficients on these variables are statistically significant at the 1% level, suggesting that per capita GDP and unobserved shocks in neighbouring regions strongly and significantly affect the per capita GDP of a given province.

In Panel A, the estimated coefficient on internal migration (*IM*) is statistically significant in the pooled data model, but not in the fixed effects models. The coefficient on international migration of domestic population (INM^D) is statistically insignificant in all models, suggesting there is no substantial effect on provinces' per capita GDP from residents leaving the country – those who migrate abroad from Türkiye are mostly looking for better economic opportunities and their departure does not affect the per capita income of those who remain. But immigration of foreign residents is positively and statistically significantly correlated with per capita GDP in all specifications. For example, a 1% increase in net foreign immigration is associated with about 1.7% increase in per capita GDP in the fixed effects, and 1.6% in SARAR models. The reason for this positive link is most likely the increase in personal consumption of additional immigrant population (Gür, 2017). The receiving province also experiences this increase through spillovers from neighbouring provinces.

[2] Lee and Yu (2010a) show that the transformation using the orthonormal eigenvector matrix of $J_T = \left(I_T - \frac{1}{T} \iota\iota' \right)$ increases the accuracy of variance estimation by compensating for the degrees of freedom. This correction is particularly important when T is small. The variance parameter representing the correction is simply $\frac{T}{T-1}\hat{\sigma}^2_{nT}$ where $\hat{\sigma}^2_{nT}$ is the residuals variance (Piras, 2014).

Table 2

Estimates of the pooled data and fixed effects models with spatial autocorrelation of errors and spatial autoregressive coefficients

	Pooled data	Fixed effects (MV) Baltagi error	Fixed effects (MV) KKP error	Fixed effects (GMM)	Fixed effects (SARAR)	Fixed effects (SARAR) Lee-Yu
Dependent variable: ln(GDPpc)						
Panel A: 2016-22[1]						
IM	0.0191***	0.0019	0.0019	0.0019	0.0018	0.0018
	(0.0040)	(0.0010)	(0.0010)	(0.0010)	(0.0010)	(0.0010)
INM^D	-0.0086	-0.0018	0.0018	-0.0018	-0.0015	-0.0015
	(0.0130)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0020)
INM^F	0.0669**	0.0170**	0.0170**	0.0174**	0.0156*	0.0156*
	(0.0310)	(0.0080)	(0.0080)	(0.0080)	(0.0080)	(0.0080)
Constant	8.8414***					
	(0.0450)					
ρ		0.8838***	0.8838***	0.8475***	0.9417***	0.9417***
		(0.0200)	(0.0200)	(0.0020)	(0.0150)	(0.0160)
λ					-0.5811***	-0.5811***
					(0.1550)	(0.1670)
Panel B: 2009-22[2]						
IM^{NB}	0.0340***	0.0038	0.0038	0.0035	-0.0011	-0.0011
	(0.0060)	(0.0020)	(0.0020)	(0.0020)	(0.0010)	(0.0010)
IM^B	0.1120***	0.0227***	0.0227***	0.0243***	0.0264***	0.0264***
	(0.0210)	(0.0070)	(0.0070)	(0.0080)	(0.0050)	(0.0050)
Constant	8.9260***					
	(0.0360)					
ρ		0.9023***	0.9023***	0.8510***	-0.7883***	-0.7883***
		(0.0120)	(0.0120)	(0.0400)	(0.1050)	(0.1090)
λ					0.9533***	0.9533***
					(0.0070)	(0.0080)

[1] For the 2016-22 period, 567 observations. [2] For the 2009-22 period, 1,134 observations.

Note: ρ = spatial error parameter; λ = spatial autoregressive coefficient. MV and GMM denote Maximum Likelihood and Generalised Method of Moments (Kelejian and Prucha, 1999) estimators. Baltagi error targets spatial dependence in the remainder term (Baltagi, Song and Koh, 2003), while KKP error captures dependence in all error components (Kapoor, Kelejian and Prucha, 2007). SARAR simultaneously models spatial dependence in both dependent variable and error term (Kelejian and Prucha, 2010). SARAR (Lee-Yu) refers to the bias-corrected estimator specifically designed for fixed effects models (Lee and Yu, 2010b).

, ** and * denote 10, 5 and 1% statistical significance levels, respectively. The values in parentheses are standard errors.*

Estimates in Panel B suggest a positive effect of skilled internal migration on per capita GDP: coefficient values of 0.0227 on IM^B in fixed effects, and 0.0264 in SARAR models suggest, e.g. that a 1% (net) increase in skilled domestic migrant population in a province increases its per capita GDP by 2.3-2.6%. This supports the view that better allocation of highly skilled labour improves per capita GDP of all provinces, even though for some it represents a "brain drain" (Pouliakas et al., 2009). Better-educated migrants bring valuable skills that contribute to overall growth, not least because the sectors that employ them tend to generate higher value added.

By contrast, the effect of unskilled internal migration is estimated to be positive and statistically significant only in the pooled data model that does not consider spatial effects. When these effects are taken into account, coefficients on IM^{NB} are not statistically significant, i.e. integration of unskilled domestic migrants into local labour market neither increases nor decreases per capita GDP of the province.

Parameter estimates on spatial autoregressive coefficients r and λ confirm spillovers from per capita GDP of neighbouring provinces and unmodelled spatial shocks on per capita GDP in a given province. For instance, the spatial autoregressive coefficient $\lambda = -0.58$ in Panel A indicates that a 1% increase in per capita GDP in neighbouring provinces leads, other things being equal, to a 0.58% decrease in the per capita GDP of a given province. By contrast, the spatial autoregressive coefficient $\lambda = 0.94$ in Panel B suggests that a 1% increase in the per capita GDP of neighbouring provinces leads to an almost equivalent increase in the host province's per capita GDP, holding other factors constant. This striking difference suggests that spatial interaction effects depend significantly on the time period and type of migration.

Positive estimates of the spatial error parameter ρ indicate that unobserved shocks such as regional economic distress or regional policy changes tend to propagate across provinces. In the SARAR specification for internal migration, for example, the negative value of the spatial error parameter ($\rho = -0.78$) alongside a positive spatial autoregressive coefficient ($\lambda = 0.95$) suggests that unobserved shocks can partly offset spatial spillovers. Technically, this negative coefficient serves as a correction mechanism, balancing the dominant positive effect of spatial dependence.

5 ROBUSTNESS CHECKS

To verify the reliability of spatial panel data models estimated in table 2, we performed three robustness checks. First, we used a lagged dependent variable $ln(GDPpc)_{t-1}$; second, we added financial control variables, i.e. the number of bank branches per capita and the amount of bank deposits per capita; and third, we added a time trend (table 3).

A comparison of table 3 with the baseline results in table 2 indicates that the size of migration coefficients in table 3 is larger, and previously insignificant estimates become statistically significant. This pattern implies that the baseline model provided a conservative estimate, with the omission of financial variables masking the full impact of internal migration. By accounting for this heterogeneity, the expanded model reduces residual variance, thereby enhancing estimator precision and revealing a more pronounced migration-regional income nexus.

Table 3

Robustness test results: spatial panel data estimations with lagged dependent variable and financial control variables

	Pooled data	Fixed effects (MV) Baltagi error	Fixed effects (MV) KKP error	Fixed effects (GMM)	Fixed effects (SARAR)	Fixed effects (SARAR) Lee-Yu
Dependent variable: ln(GDPpc)						
Panel A[1]						
$\ln(GDPpc)_{t-1}$	1.0095*** (0.0101)	0.8582*** (0.0532)	0.3952** (0.0440)	-0.1353** (0.0543)	0.3357*** (0.0453)	-0.1514*** (0.0584)
IM	-0.0022 (0.0014)	-0.0033 (0.0020)	-0.0001 (0.0010)	-0.0005 (0.0014)	-0.0004 (0.0013)	-0.0004 (0.0015)
INM^D	-0.0072* (0.0044)	-0.0079* (0.0047)	-0.0060** (0.0027)	-0.0041 (0.0025)	-0.0037 (0.0023)	-0.0037 (0.0026)
INM^F	0.0245** (0.0096)	0.0304** (0.0122)	0.0165** (0.0079)	0.0218** (0.0088)	0.0202** (0.0082)	0.0202** (0.0092)
Population per bank branch	-0.0237 (0.1901)	-0.1471 (0.2203)	-0.0305 (0.1364)	-0.1176 (0.1459)	-0.1072 (0.1354)	-0.1072 (0.1514)
Per capita bank deposit	0.5621*** (0.0881)	0.5548*** (0.0972)	0.1674*** (0.0614)	0.2205** (0.0958)	0.2130** (0.0894)	0.2130** (0.1000)
Constant	33.3542*** (6.4863)					
Trend	0.0167*** (0.0032)	0.0111*** (0.0038)	0.0034 (0.0082)	0.0048 (0.0049)	0.0051** (0.0024)	0.0051** (0.0024)
ρ		0.7508*** (0.0407)	0.8581*** (0.0262)	0.7215*** (0.0028)	-0.2171*** (0.1620)	-0.2171*** (0.1620)
λ					0.7863*** (0.0356)	0.7863*** (0.0356)
Panel B[2]						
$\ln(GDPpc)_{t-1}$	-0.9789*** (0.0007)	0.6566*** (0.0203)	0.6315*** (0.0243)	-0.6195*** (0.0235)	-0.6251*** (0.0256)	-0.6251*** (0.0256)
IM^{NB}	-0.0006 (0.0019)	-0.0036 (0.0022)	-0.0021 (0.0017)	-0.0023 (0.0017)	-0.0022 (0.0017)	-0.0022 (0.0017)
IM^B	0.0045 (0.0062)	0.0120 (0.0089)	0.0111* (0.0062)	-0.0115* (0.0065)	-0.0114* (0.0063)	-0.0114* (0.0063)
Population per bank branch	-0.5210 (0.1113)	-0.2951*** (0.1055)	-0.1648** (0.0777)	-0.1738** (0.0807)	-0.1535* (0.0785)	-0.1535* (0.0785)
Per capita bank deposit	1.1888** (0.0931)	0.7543** (0.0531)	0.1646** (0.0432)	0.2240** (0.0445)	0.1466** (0.0438)	0.1466** (0.0438)
Constant	-37.7120*** (1.7747)					
Trend	-0.0189*** (0.0009)	-0.0180*** (0.0008)	-0.0131*** (0.0022)	-0.0138*** (0.0014)	-0.0213*** (0.0041)	-0.0213*** (0.0041)
ρ		0.8539*** (0.0017)	0.8283*** (0.0207)	0.7053*** (0.0250)	0.8926*** (0.0198)	0.8926*** (0.0198)
λ					-0.5455*** (0.1781)	-0.5455*** (0.1991)

[1] For the 2016-22 period, 486 observations. [2] For the 2009-22 period, 1,053 observations.

Note: ρ = spatial error parameter; λ = spatial autoregressive coefficient. MV and GMM denote Maximum Likelihood and Generalised Method of Moments (Kelejian and Prucha, 1999) estimators. Baltagi error targets spatial dependence in the remainder term (Baltagi et al., 2003), while KKP error captures dependence in all error components (Kapoor et al., 2007). SARAR simultaneously models spatial dependence in both dependent variable and error term (Kelejian and Prucha, 2010). SARAR (Lee-Yu) refers to the bias-corrected estimator specifically designed for fixed effects models (Lee and Yu, 2010b).

, ** and * denote 10, 5 and 1% statistical significance levels, respectively. The values in parentheses are standard errors.*

It is instructive to view the dynamics of the lagged dependent variable through the lens of econometric theory. While the positive coefficients in pooled OLS models likely stem from upward bias due to unobserved heterogeneity, the negative estimates in standard fixed effects models reflect the so-called Nickell (1981) bias. Consequently, the persistence of negative coefficients in our bias-corrected Lee-Yu and GMM specifications should be interpreted not as an artifact, but as robust evidence of conditional convergence.

Regarding spatial dynamics, Panel A estimates for the SARAR and Lee-Yu models show a positive spatial autoregressive coefficient: $\lambda = 0.78$ indicates strong economic integration among provinces. The model balances the excessively positive income interaction with a negative spatial error parameter ($\rho = -0.21$). On the other hand, in Panel B, a negative spatial autoregressive coefficient ($\lambda = -0.54$) suggests the presence of inter-regional competition and backwash effects, while the diffusion of shocks is confirmed by a positive spatial error parameter ($\rho = 0.89$). Furthermore, positive values of the spatial error parameter ρ across the Baltagi, KKP, and GMM specifications confirm the spatial diffusion of unobserved regional shocks across different model specifications.

Regarding additional controls, the size of bank deposits per capita is positively and significantly associated with provincial per capita GDP. This result aligns with expectations: for instance, in SARAR models, 1% higher bank deposits are associated with a 0.15 – 0.20% higher per capita income. This supports findings by Kim et al. (2018) and Afonso and Blanco-Arana (2024) that retail deposit growth stimulated economic activity. Surprisingly, population per bank branch is negatively correlated with regional per capita GDP. The time trend coefficient varies by period: it is positive in 2016-22 (Panel A), reflecting a steady upward trend, but negative in the longer 2009-22 period (Panel B), likely capturing the lingering effects of the Global Financial Crisis.

6 CONCLUSION

This study used spatial panel models to analyse the impact of international and internal migration on per capita GDP of provinces in Türkiye. Our findings suggest that migration from other countries into Türkiye is associated with a small positive effect on the per capita GDP of provinces. Regarding internal migration, the level of migrants' human capital is important: educated internal migrants have a statistically significant positive effect on per capita GDP of host provinces, whereas less educated migrants have no significant effect, i.e. their migration neither decreases nor increases the per capita GDP of the province they move into.

Our estimates control for spatial dependencies, i.e. spillovers from nearby provinces on the link between net migration and per capita GDP in a given province. They are also robust to alternative specifications that include a lagged per capita GDP, a time trend, and financial development variables.

PUBLIC SECTOR ECONOMICS
50 (2) 259-286 (2026)

The weak but positive effect of net international migration on per capita GDP of Türkiye's provinces is consistent with the view that foreign workers contribute positively to the domestic economy. However, the small magnitude of this effect suggests that additional factors, such as the type of employment and integration policies for incoming migrants, might not fully exploit the potential contribution of this pool of workforce to the Turkish economy.

The significant positive effect of educated native migrants on per capita GDP is in line with the literature suggesting that reallocation of a highly educated workforce benefits domestic economic growth (Akbari and Haider, 2018; Özyılmaz and Bayraktar, 2021). The finding that less-educated domestic migrants do not depress per capita GDP in the provinces they move into is also interesting. We believe that this interpretation is more important than the view that low-skilled labour merely has a "limited impact" on per capita GDP, as argued e.g. by Marto, Lourenço Marques and Madaleno (2022).

One contribution of our paper to the existing literature is that we consider spatial dependency of per capita GDP in a province on per capita GDP in neighbouring provinces. Another is that we account for errors caused by unobservable common factors (e.g. regional policies) that affect neighbouring provinces. The inclusion of these spatial interaction parameters did not change the sign of relationships between migration variables and per capita GDP of provinces. The varying statistical significance of spatial parameters suggests that income inequalities across regions may affect per capita GDP. This opens a new avenue for future research. The main limitations of this study are that indirect effects of migration on per capita GDP of provinces are not analysed, and the time period under study is limited.

The main policy implication of our findings is that one needs to create favourable conditions for the integration of international and as well as of internal migrants into provincial labour markets. Policies that guide migrants to sectors aligned with their skills and human capital could increase the positive impact of migration on the per capita GDP of Türkiye's provinces. Without such guidance, migrants may end up in the informal sector, lowering their potential contribution to economic growth and public finances. The significant role of human capital suggests that attracting and retaining skilled migrants could be a key regional policy priority. Having said that, public policy coordination to address spatial dependencies and ensure balanced economic development across provinces is essential.

Disclosure statement
The authors have no conflicts of interest to declare.

REFERENCES

1. Afonso, A. and Blanco-Arana, M. C., 2024. Does Financial Inclusion Enhance per capita Income in the Least Developed Countries? *International Economics,* 177, 100479. https://doi.org/10.1016/j.inteco.2024.100479
2. Akbari, A. H. and Haider, A., 2018. Impact of Immigration on Economic Growth in Canada and in its Smaller Provinces. *Journal of International Migration and Integration*, 19, pp. 129-142. https://doi.org/10.1007/s12134-017-0530-4
3. Aleksynska, M. and Tritah, A., 2015. The Heterogeneity of Immigrants, Host Countries' Income and Productivity: A Channel Accounting Approach. *Economic Inquiry*, 53(1), pp. 150-172. https://doi.org/10.1111/ecin.12141
4. Altunç, Ö. F., Uçan, O. and Akyıldız, A., 2017. Dış Göçlerin Türkiye Ekonomisinde İşsizlik Enflasyon ve Ekonomik Büyüme Üzerine Etkileri: Ekonometrik Bir Analiz (1985-2015). *Social Science Studies*, 5(1), pp. 197-211.
5. Anselin, L., Gallo, J. L. and Jayet, H. 2008. Spatial panel econometrics. In: L. Mátyás and P. Sevestre, eds. *The econometrics of panel data: Fundamentals and recent developments in theory and practice*. Berlin; Heidelberg: Springer, pp. 625-660. https://doi.org/10.1007/978-3-540-75892-1_19
6. Baltagi, B. H., Song, S. H. and Koh, W., 2003. Testing panel data regression models with spatial error correlation. *Journal of Econometrics*, 117(1), pp. 123-150. https://doi.org/10.1016/S0304-4076(03)00120-9
7. Barro, R. and Sala-i-Martin, X., 1992. Regional Growth and Migration: A Japan-United States Comparison. *Journal of the Japanese and International Economies*, 6(4), pp. 312-346. https://doi.org/10.1016/0889-1583(92)90002-L
8. Barro, R. and Sala-i-Martin, X., 2004. *Economic Growth*. Cambridge MA: The MIT Press.
9. Bashier, A. A. and Siam, A. J., 2014. Immigration and Economic Growth in Jordan: FMOLS Approach. *International Journal of Humanities Social Sciences and Education (IJHSSE)*, 1(9), pp.85-92.
10. Bauer, T. and Zimmermann, K. F., 1999. Assessment of Possible Migration Pressure and its Labour Market Impact Following EU Enlargement to Central and Eastern Europe. *IZA Research Report*, No. 3.
11. Bayraktar, Y. and Özyılmaz, A., 2019. Türkiye'de İç Göç ve Ekonomik Büyüme. *İş ve Hayat*, 5(9), pp. 100-111.
12. Borjas, G. J., 1999. The Economic Analysis of Immigration. In: O. C. Ashenfelter and D. Card, eds. *Handbook of Labor Economics*, 3, pp. 1697-1760. https://doi.org/10.1016/S1573-4463(99)03009-6
13. Borozan, D., 2017. Internal Migration, Regional Economic Convergence, and Growth in Croatia. *International Regional Science Review*, 40(2), pp. 141-163. https://doi.org/10.1177/0160017615572889
14. Boubtane, E., Coulibaly, D. and Rault, C., 2013a. Immigration, Unemployment and GDP in the Host Country: Bootstrap Panel Granger Causality Analysis on OECD Countries. *Economic Modelling*, 33, pp. 261-269. https://doi.org/10.1016/j.econmod.2013.04.017

15. Boubtane, E., Coulibaly, D. and Rault, C., 2013b. Immigration, Growth, and Unemployment: Panel VAR Evidence from OECD Countries. *LABOUR*, 27(4), pp. 399-420. https://doi.org/10.1111/labr.12017
16. Boubtane, E., Dumont, J.-C. and Rault, C., 2016. Immigration and Economic Growth in the OECD Countries 1986-2006. *Oxford Economic Papers*, 68(2), pp. 340-360. https://doi.org/10.1093/oep/gpw001
17. Brunow, S., Nijkamp, P. and Poot, J., 2015. The Impact of International Migration on Economic Growth in the Global Economy. In: *Handbook of the Economics of International Migration*, 1, pp. 1027-1075. https://doi.org/10.1016/B978-0-444-53768-3.00019-9
18. Bunea, D., 2012. Modern Gravity Models of Internal Migration. The case of Romania. *Theoretical and Applied Economics*, 19(4), pp. 127-144.
19. Chen, H.-J., 2006. International Migration and Economic Growth: A Source Country Perspective. *Journal of Population Economics*, 19, pp. 725-748. https://doi.org/10.1007/s00148-005-0023-1
20. Di Maria, C. and Lazarova, E. A., 2012. Migration, Human Capital Formation, and Growth: An Empirical Investigation. *World Development*, 40(5), pp. 938-955. https://doi.org/10.1016/j.worlddev.2011.11.011
21. Dickey, D. A. and Fuller, W. A., 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association,* 74(366), pp. 427-431. https://doi.org/10.2307/2286348
22. Docquier, F., Peri, G. and Ruyssen, I., 2014. The Cross-country Determinants of Potential and Actual Migration. *International Migration Review*, 48(1 suppl), pp. 37-99. https://doi.org/10.1111/imre.12137
23. Dolado, J., Goria, A. and Ichino, A., 1994. Immigration, Human Capital and Growth in the Host Country: Evidence from Pooled Country Data. *Journal of Population Economics*, 7(2), pp. 193-215. https://doi.org/10.1007/BF00173619
24. Draženović, I., Kunovac, M. and Pripužić, D., 2018. Dynamics and Determinants of Emigration: The Case of Croatia and The Experience of New EU Member States. *Public Sector Economics*, 42(4), pp. 415-447. https://doi.org/10.3326/pse.42.4.3
25. Etzo, I., 2008. Internal Migration: A Review of the Literature. *MPRA Paper,* No. 8783.
26. Felbermayr, G. J., Hiller, S. and Sala, D., 2010. Does Immigration Boost Per Capita Income? *Economics Letters*, 107(2), pp. 177-179. https://doi.org/10.1016/j.econlet.2010.01.017
27. Foldvari, P., Van Leeuwen, B. and Van Zanden, J. L., 2013. The Contribution of Migration to Economic Development in Holland 1570-1800. *De Economist*, 161, pp. 1-18. https://doi.org/10.1007/s10645-012-9197-6
28. Ghatak, S., Mulhern, A. and Watson, J., 2008. Inter-Regional Migration in Transition Economies: The Case of Poland. *Review of Development Economics*, 12(1), pp. 209-222. https://doi.org/10.1111/j.1467-9361.2008.00435.x

29. Göv, A. and Dürrü, Z., 2017. Göç ve Ekonomik Büyüme İlişkisi: Seçilmiş OECD Ülkeleri Üzerine Ekonometrik Bir Analiz. *International Journal of Economic Studies*, 3(4), pp. 491-502.
30. Greenwood, M. J., 1975. Research on Internal Migration in the United States: A Survey. *Journal of Economic Literature*, 13(2), pp. 397-433. https://www.jstor.org/stable/2722115
31. Greenwood, M. J., 1997. Internal Migration in Developed Countries. In: *Handbook of Population and Family Economics*, 1, pp. 647-720. https://doi.org/10.1016/S1574-003X(97)80004-9
32. Güner, U. and Yalınız, M., 2013. Immigration and Economic Growth in Europe and their Spatial Allocation. *Актуальні Проблеми Економіки*, 12, pp. 373-380.
33. Gür, N., 2017. Ülke Deneyimleri Işığında Uluslararası Göç Ekonomisi. *SETA,* (224), pp. 1-22.
34. Huber, P. and Tondl, G., 2012. Migration and Regional Convergence in the European Union. *Empirica*, 39, pp. 439-460. https://doi.org/10.1007/s10663-012-9199-2
35. Kang, W., 2018. *Issues in the Distribution Dynamics Approach to the Analysis of Regional Economic Growth and Convergence: Spatial Effects and Small Samples*. Arizona State University.
36. Kang, Y. and Kim, B.-Y., 2018. Immigration and Economic Growth: Do Origin and Destination Matter? *Applied Economics*, 50(46), pp. 4968-4984. https://doi.org/10.1080/00036846.2018.1466987
37. Kanga, D. [et al.], 2022. The Diffusion of Fintech, Financial Inclusion and Income per capita. *The European Journal of Finance*, 28(1), pp. 108-136. https://doi.org/10.1080/1351847X.2021.1945646
38. Kapoor, M., Kelejian, H. H. and Prucha, I. R., 2007. Panel data models with spatially correlated error components. *Journal of Econometrics*, 140(1), pp. 97-130. https://doi.org/10.1016/j.jeconom.2006.09.004
39. Kelejian, H. H. and Prucha, I. R., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40(2), pp. 509-533. https://doi.org/10.1111/1468-2354.00027
40. Kelejian, H. H. and Prucha, I. R., 2010. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157(1), pp. 53-67. https://doi.org/10.1016/j.jeconom.2009.10.025
41. Kırdar, M. G. and Saracoğlu, D. Ş., 2012. İç Göç, Bölgesel Yakınsama Sorunu ve Ekonomik Büyüme: Türkiye Örneği. *Discussion Paper*, No. 2012/75. Ankara: Turkish Economic Association.
42. Kim, D. W., Yu, J. S. and Hassan, M. K., 2018. Financial Inclusion and Economic Growth in OIC Countries. *Research in International Business and Finance,* 43, pp. 1-14. https://doi.org/10.1016/j.ribaf.2017.07.178
43. Lee, L. and Yu, J., 2010a. Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects. *Journal of Econometrics*, 154(2), pp. 165-185. https://doi.org/10.1016/j.jeconom.2009.08.001

44. Lee, L. F. and Yu, J., 2010b. Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, 40(5), pp. 255-271. https://doi.org/10.1016/j.regsciurbeco.2009.09.002
45. LeSage, J. P. and Fischer, M. M., 2009. Spatial econometric methods for modeling origin-destination flows. In: *Handbook of applied spatial analysis: Software tools, methods and applications*. Berlin; Heidelberg: Springer, pp. 409-433. https://doi.org/10.1007/978-3-642-03647-7_20
46. Lucas, R. E., 1997. Internal Migration in Developing Countries. In: *Handbook of Population and Family Economics*, 1, pp. 721-798. https://doi.org/10.1016/S1574-003X(97)80005-0
47. Manavgat, G. and Saygılı, R. F., 2016. Türkiye'de İçgöçü Etkileyen Faktörler Üzerine bir Uygulama: Mekânsal Panel Veri Analizi. *2nd International Conference on Applied Economics and Finance (ICOAEF)*, pp. 1-26
48. Mariya, A. and Tritah, A., 2009. *Immigration, Income and Productivity of Host Countries: A Channel Accounting Approach. CEPII Working Papers,* No. 2009-23.
49. Marto, M., Lourenço Marques, J. and Madaleno, M., 2022. An Evaluation of the Efficiency of Tertiary Education in the Explanation of the Performance of GDP per Capita Applying Data Envelopment Analysis (DEA). *Sustainability*, 14(23), pp. 1-20. https://doi.org/10.3390/su142315524
50. Morley, B., 2006. Causality between Economic Growth and Immigration: An ARDL Bounds Testing Approach. *Economics Letters*, 90(1), pp. 72-76. https://doi.org/10.1016/j.econlet.2005.07.008
51. Nickell, S., 1981. Biases in dynamic models with fixed effects. *Econometrica*, 49(6), pp. 1417-1426. https://www.jstor.org/stable/1911408
52. Orefice, G., 2010. Skilled Migration and Economic Performances: Evidence from OECD Countries. *Swiss Journal of Economics and Statistics*, 146(4), pp. 781-820. https://doi.org/10.1007/BF03399337
53. Ortega, F., 2008. *The Short-run Effects of a Large Immigration Wave: Spain 1998-2008*. Manuscript.Universitat Pompeu Fabra.
54. Ortega, F. and Peri, G., 2009. The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005. *NBER Working Paper,* No. 14833. https://doi.org/10.3386/w14833
55. Özgen, C., Nijkamp, P. and Poot, J., 2010. The Effect of Migration on Income Growth and Convergence: Metaanalytic Evidence. *Papers in Regional Science*, 89(3), pp. 537-562. https://doi.org/10.1111/j.1435-5957.2010.00313.x
56. Öztop, A. O., Köse, T. and Ünlü, Z., 2024. Türkiye'de il düzeyinde GSYİH ve göç arasındaki ilişkinin mekânsal veri analizi ile araştırılması. *Academy 2nd International Conference on Migration Studies*, pp. 488-296. Rome.
57. Özyılmaz, A. and Bayraktar, Y., 2021. Internal Migrations as a Driving Force of Regional Disintegration: An Empirical Analysis of NUTS-2 Regions in Turkey. *The Journal of Humanity and Society*, 11(3), pp. 197-214. https://doi.org/10.12658/M0632

58. Peri, G., 2010. *The Impact of Immigrants in Recession and Economic Expansion*. Washington: Migration Policy Institute.
59. Pesaran, M. H., 2007. A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics,* 22(2), pp. 265-312. https://doi.org/10.1002/jae.951
60. Piras, G., 2014. Impact Estimates for Static Spatial Panel Data Models in R. *Letters in Spatial and Resource Sciences*, 7(3), pp. 213-223. https://doi.org/10.1007/s12076-013-0113-8
61. Pouliakas, K. [et al.], 2009. Modelling the Effects of Immigration on Regional Economic Performance and the Wage Distribution: A CGE Analysis of Three EU Regions. *IZA Discussion Paper*, No. 4648. http://dx.doi.org/10.2139/ssrn.1526078
62. Presidency of Migration Management, 2024. https://en.goc.gov.tr/
63. Pu, Y. [et al.], 2019. A spatial dynamic panel approach to modelling the space-time dynamics of interprovincial migration flows in China. *Demographic Research*, 41, pp. 913-948. http://doi.org/10.4054/DemRes.2019.41.31
64. Ravenstein, E. G., 1885. The Laws of Migration. *Journal of the Royal Statistical Society*, 52(2), pp. 241-305. https://doi.org/10.2307/2979333
65. Salima, B. A., Julie, L. G. and Lionel, V., 2018. Spatial Econometrics on Panel Data. In: *Handbook of Spatial Analysis Theory and Practical Application with R*, pp. 179-203.
66. Sonkar, S. and Sarkar, A. K., 2020. Exploring the Direct Relationship between GDP per-capita and Financial Inclusion. *Annals of Management and Organization Research,* 1(3), pp. 187-202. https://doi.org/10.35912/amor.v1i3.415
67. Şerban, A. C. [et al.], 2020. The Impact of EU Immigration on Economic Growth through the Skill Composition Channel. *Technological and Economic Development of Economy*, 26(2), pp. 479-503. https://doi.org/10.3846/tede.2020.11954
68. Sevinç, H. [et al.], 2016. Ekonomik Büyüme ve Göç İlişkisi: Gelişmekte Olan Ülkelere Dayalı Bir Analiz. *International Conference on Eurasian Economies*, pp. 398-403. https://doi.org/10.36880/C07.01715
69. Ulucan, H., 2022. Türkiye'de Şehirler Arası İç Göç Akımlarını Belirleyen Faktörler: Panel Veri Analizi. *Ekonomik Yaklasim*, 33(122), pp. 45-65. https://doi.org/10.5455/ey.21006
70. Vakulenko, E., 2014. Does Migration Lead to Regional Convergence in Russia? *Higher School of Economics Research Paper,* No. WP BRP, 53. https://doi.org/10.2139/ssrn.2395835
71. Yürük, B. and Batmaz, T., 2023. İç Göç ve Ekonomik Büyüme Arasındaki İlişkinin Mekansal Analizi: Türkiye Örneği (2008-2020). *Academic Review of Economics and Administrative Sciences*, 16(3), pp. 763-785. https://doi.org/10.25287/ohuiibf.1248210
72. Zhang, K. H. and Song, S., 2003. Rural-urban Migration and Urbanization in China: Evidence from Time-series and Cross-section Analyses. *China Economic Review*, 14(4), pp. 386-400. https://doi.org/10.1016/j.chieco.2003.09.018

APPENDIX

Table A1
Overview of the empirical literature on migration and economic growth

Author	Country or region	Time span	Method	Result
International migration and economic growth				
Dolado et al. (1994)	23 OECD countries	1960-1985	Panel NLS, NLS2SLS	M*→G (+)
Chen (2006)	USA, Philippines	1985-1994	Stochastic dynamic model method	M→G (+) M*→G (-)
Morley (2006)	Australia, Canada, USA	1930-2002	ARDL, Granger non-causality tests	G=> M
Mariya and Tritah (2009)	20 OECD countries	1960-2005	Panel regressions analysis	M→ G (✓)
Ortega and Peri (2009)	74 countries	1980-2005	Panel OLS, 2SLS	M→ G (+)
Orefice (2010)	24 OECD countries	1998-2007	Panel OLS, 2SLS	M*→G (+)
Boubtane, Coulibaly and Rault (2013a)	22 OECD countries	1980-2005	Bootstrap panel Granger causality analysis	G=> M
Boubtane, Coulibaly and Rault (2013b)	22 OECD countries	1987-2009	Panel VAR	M↔G (+)
Güner and Yalınız (2013)	EU 14 countries	1984-2008	Panel data fixed effect model	M→ G (-)
Bashier and Siam (2014)	Jordan	1980-2012	Fully modified ordinary least squares method (FMOLS)	M→ G (+)
Docquier et al. (2014)	138 sending countries and 30 major destination countries	2000-2010	OLS regression analysis	G→M (+)
Aleksynska and Tritah (2015)	20 OECD countries	1960-2005	Panel 2SLS	M→ G (+/-)
Brunow et al. (2015)	149 countries	1950-2010	Panel OLS	M→ G (+)
Boubtane, Dumont and Rault (2016)	22 OECD countries	1986-2006	SYS-GMM	M*→G (+)
Sevinç et al. (2016)	18 developing countries	1962-2012	Panel data methods	M→ G (-)

Author	Country or region	Time span	Method	Result
International migration and economic growth				
Altunç et al. (2017)	Türkiye	1985-2015	Johansen cointegration test, Granger causality tests	M→ G (X) M<=> G
Göv and Dürrü (2017)	7 OECD countries	2000-2016	Panel causality test	M=> G
Kang and Kim (2018)	90 countries	1960-2000	Panel GMM	M*→G (+)
Öztop et al. (2024)	Türkiye	2022	Spatial regression analysis	M→ G (✓)

*Note: (→) and (↔) signify one-way and bi-directional relationships; (=>) and (<=>) signify one-way and a bi-directional causal nexus; (✓) and (X) indicate that a relationship exists or does not exist; (+) and (-) indicate positive and negative relationship; * refers to human capital.*

Overview of the empirical literature on migration and economic growth (continued)

Author	Country or region	Time span	Method	Result
Internal migration and economic growth				
Barro and Sala-i-Martin (1992)	Japan and United States	1900-1987	Regression analysis	M→ G (+)
Zhang and Song (2003)	China	1978-1999	Time series and cross-section analyses	G→ M (+)
Ghatak et al. (2008)	Poland	1995-2001	SURE, FGLS SURE, ML	G→ M (+)
Ortega (2008)	Spanish regions	1998-2008	Panel OLS	M→ G (+)
Pouliakas et al. (2009)	Scotland, Greece, Latvia	2004/2005	CGE analysis	M*→G (+/-)
Peri (2010)	USA	1960-2006	Panel 2SLS	M→ G (+)
Bunea (2012)	Romania	2004-2008	Fixed effects and generalized method of moments	G→ M (+)
Huber and Tondl (2012)	EU27 NUTS2 regions	2000-2007	Panel system GMM	M→ G (+/-)
Kırdar and Saraçoğlu (2012)	Türkiye	1975-2000	OLS	M→ G (-)
Foldvari et al. (2013)	Holland	1570-1800	VAR analysis	M→ G (+)
Vakulenko (2014)	Russian	1995-2010	Spatial dynamic panel data analysis	M→ G (-)
Manavgat and Saygılı (2016)	Türkiye	2008-2011	Spatial danel data analysis	G→ M (+/-)
Borozan (2017)	Croatia	2000-2011	Panel data analysis	M→ G (-)
Akbari and Haider (2018)	Canada	2006-2013	Feasible generalized least squares (FGLS)	M*→G (+)
Bayraktar and Özyılmaz (2019)	Türkiye's NUTS-2 Regions	2008-2017	POLS, fixed effects and random effects models	M→ G (+)
Özyılmaz and Bayraktar (2021)	Türkiye	2008-2019	Bootstrap quantile regression method	M→ G (+) M*→G (+)
Ulucan (2022)	Türkiye	2008-2019	Fixed effect panel method	G→ M (-)
Yürük and Batmaz (2023)	Türkiye	2008-2020	Spatial panel data analysis	M→ G (-)

*Note: (à) and (↔) signify one-way and bi-directional relationships; (+) and (-) indicate positive and negative relationship; * refers to human capital.*

Table A2
Descriptive statistics

Variables	Mean	St. dev.	Min.	Max.
Panel A				
Per capita GDP (USD)	7,5100	2,7270	2,9070	18,2690
IM (% of population)	-0.2600	2.3400	-21.0600	14.9700
INM^D (% of population)	0.0300	0.6900	-10.0800	6.6300
INM^F (% of population)	0.1700	0.3300	-0.9000	2.9400
Population per bank branch	0.0100	0.0200	-0.0600	0.0800
Per capita bank deposit	0.1200	0.0500	-0.0200	0.2600
Panel B				
Per capita GDP (USD)	8,0460	2,9950	2,7670	20,8830
IM^{NB} (% of population)	-0.0400	1.2400	-9.2900	11.8900
IM^B (% of population)	-0.1500	0.3500	-3.3800	2.4800
Population per bank branch	0.0000	0.0200	-0.1200	0.0800
Per capita bank deposit	0.0900	0.0500	-0.0800	0.2600

Note: N for Panel A = 567; N for Panel B = 1,134. Panel A examines the effects of net internal migration (IM) and net international migration separated into domestic (INM^D) and foreign (INM^F) components, on per capita GDP in USD. Panel B disaggregates internal migration by high (IM^B) and low education (IM^{NB}).

These statistics underscore the pronounced regional disparities that underpin our spatial approach. The mean per capita GDP in fixed US dollars during the core period was $7,500, but the standard deviation of the data exceeded $2,700, indicating a high degree of dispersion. In the economic laggard eastern provinces, the median per capita income was below $3,000, whereas in the more prosperous urban west it approached $18,000.

These differences are further exacerbated by net internal migration losses (negative averages across panels) and extreme outliers (e.g. 21% outflows in rural areas versus 15% inflows in urban areas). Such fluctuations are indicative of the risk of brain drain in peripheral regions and may have a negative impact on the growth of neighbouring regions unless balanced by foreign capital inflows or skilled migration.

Population per bank branch indicators highlight persistent inclusivity gaps: near-zero log ratios for branches imply sparse coverage (typically 1 branch per 10,000+ people), while per capita bank deposit shares are below 10 – 12% of per capita GDP and show negative values in underserved areas. This situation has the effect of limiting productivity growth in remittance revenues and increasing spatial inequalities.

PUBLIC SECTOR
ECONOMICS
50 (2) 259-286 (2026)

Table A3

Spatial dependence test

	Panel A	Panel B
	Chi-squared statistic	**Chi-squared statistic**
Hausman test	14.8900***	17.8190***
Hausman test for spatial error models	0.2062	0.2991
Hausman test for spatial lag models	0.4085	0.3393
	LM statistic	**LM statistic**
LM spatial lag	1622.4000***	3814.8000***
LM spatial error	1620.7000***	3768.0000***
Robust LM spatial lag	15.5500***	68.9320***
Robust LM spatial error	13.8380***	22.1580***

*Note: *, ** and *** denote 10%, 5% and 1% statistical significance levels, respectively.*

Table A4

Moran's I spatial autocorrelation test for per capita GDP

Year	Observed Moran's I value	Z-score
2009	0.4997	11.6213***
2010	0.4942	11.4987***
2011	0.4966	11.5725***
2012	0.4975	11.5890***
2013	0.4803	11.2300***
2014	0.5305	12.3582***
2015	0.5343	12.4742***
2016	0.5348	12.4831***
2017	0.5251	12.2924***
2018	0.5274	12.3391***
2019	0.4972	11.6083***
2020	0.4885	11.4015***
2021	0.5092	11.9086***
2022	0.5429	12.6434***

*Note: *, ** and *** denote 10%, 5% and 1% statistical significance levels, respectively.*

Are taxes too high? A machine-learning approach to Laffer curve estimation

HERMES MORGAVI, Ph.D.*

Article**
JEL: H21, C51, C54
https://doi.org/10.3326/pse.50.2.5

* The author is grateful to two anonymous referees who have contributed to the quality of the final version of the paper.

** Received: September 24, 2025
Accepted: March 23, 2026

Hermes MORGAVI
OECD Economics Department, 2 Rue André Pascal, 75016 Paris, France
e-mail: hermes.morgavi@oecd.org
ORCiD: 0000-0001-8427-3919

PUBLIC SECTOR ECONOMICS
50 (2) 287-319 (2026)

Abstract

This paper estimates Laffer curves for personal income tax, corporate income tax, and value-added tax across OECD countries. While the Laffer curve is widely used for assessing the revenue effects of taxation, existing empirical estimates typically rely on restrictive functional forms and are vulnerable to misspecification, when the true relationship between tax rates and revenues is unknown. In response to this limitation, this paper develops a model that allows data-driven flexibility and enforces the defining properties of the Laffer curve. The parameters governing the curvature and turning points of the curve depend on a rich set of structural and institutional characteristics while LASSO regularisation mitigates overfitting. The results reveal substantial cross-country heterogeneity in revenue-maximising tax rates among OECD countries and suggest there is limited scope for further revenue mobilisation through higher income tax rates in several countries, while highlighting a comparatively greater fiscal space in consumption taxation.

Keywords: optimal taxation, Laffer curve, macroeconomic modelling, LASSO

1 INTRODUCTION

Rising fiscal pressures associated with population ageing, climate mitigation and adaptation, defence spending, and digital investment have renewed policy interest in understanding how changes in tax rates affect public revenue. In this context, the Laffer curve provides a unifying framework to analyse the trade-off between statutory tax rates and the responsiveness of tax bases. While higher tax rates mechanically increase revenues, they may also trigger behavioural, compositional, administrative, and macroeconomic responses that will erode the tax base and ultimately reduce revenue. Identifying the shape of the Laffer curve and the location of its turning point therefore remains central to the design of sustainable and efficient tax systems.

Despite a substantial empirical literature, most reduced-form estimates of the Laffer curve rely on restrictive functional forms – typically quadratic, polynomial, or log-linear specifications – and concentrate on identifying revenue-maximising tax rates in the neighbourhood of observed policy settings. These approaches often neglect key theoretical properties of the Laffer relationship – most notably the requirement that revenues approach zero at tax rates of 0% and 100% – and may suffer from misspecification bias when extrapolated beyond the observed data range.

The true functional relationship between tax rates and aggregate revenues is inherently non-linear and difficult to specify *ex ante*, particularly in cross-country settings characterised by heterogeneity in economic structures, institutional frameworks, and tax systems. The present paper contributes to the empirical literature by proposing a flexible model, grounded in the constant elasticity model of the taxable income literature, that incorporates the core theoretical properties of the Laffer curve while allowing for a broad range of shapes reflecting country heterogeneity. Rather than unduly restricting the model, these constraints incorporate prior economic structure into the estimation, thereby improving statistical identification, excluding economically implausible functional forms, and enhancing the credibility of out-of-sample extrapolation beyond the observed tax rates range.

The shape of the Laffer curve is modelled as a function of a rich set of structural, institutional, and macroeconomic characteristics. The specification improves estimation precision by pooling information across countries and over time. The model is estimated using a LASSO technique, whose regularisation properties reduce the risk of overfitting, improve out-of-sample predictive performance, and mitigate multicollinearity among structural covariates.

The model is applied to personal income tax (PIT), corporate income tax (CIT), and value-added tax (VAT) data for a panel of OECD countries from 2000 onwards. The analysis yields estimates of revenue-maximising tax rates and highlights substantial heterogeneity across countries, reflecting differences in economic structure, institutions, and tax design. By integrating regularisation techniques into a theoretically grounded Laffer curve framework, this paper advances the empirical literature and provides policy-relevant insights into the scope and limits of revenue mobilisation through tax rate changes.

The remainder of the paper is organised as follows. Section 2 reviews the related literature. Section 3 describes the data. Section 4 presents the model; section 5 presents the estimation strategy. Section 6 discusses the empirical results, and section 7 concludes.

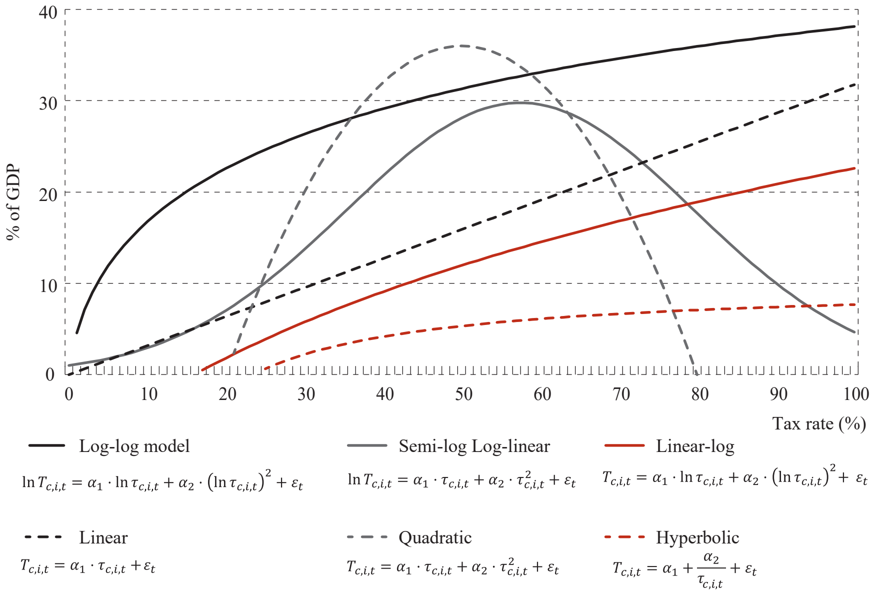
2 LITERATURE REVIEW

The empirical estimation of the Laffer curve has produced a substantial body of reduced-form studies that seek to characterise the macroeconomic effects of changes in tax rates on tax revenues, often applying polynomial, hyperbolic, or log-linear functional forms to explore the relationship between average tax rates and tax revenues. Hsing (1996) estimated four alternative specifications of the US Laffer curve and found a peak tax rate near 33%. Heijman and van Ophem (2005) applied a quadratic optimisation model to estimate country-specific optimal tax rates for 12 OECD countries, while Liapis et al. (2020) compared multiple functional forms for average personal income tax (PIT) and corporate income tax (CIT) rates in the EU and concluded that quadratic models offer the best fit. Other studies focused on consumption taxes. Matthews (2003) and de Oliveira and Costa (2015) estimated value added tax (VAT), Laffer curves and confirmed a hump-shaped relationship, with peak rates between 20% and 25%. Akgun, Bartolini, and Cournède (2017) extended this approach to a broader set of consumption taxes across OECD countries and found that VAT revenues tend to peak around 19%. Ferreira-Lopes, Martins, and Espanhol (2019) used implicit tax rates and concluded that several countries operate well below their potential revenue levels. Liapis et al. (2020) and Agersnap and Zidar (2020) applied this methodology to capital taxation, showing that capital mobility in open economies constrains the achievable revenue-maximising rate.

Much of the empirical literature has remained narrowly focused on estimating this optimal point and fitting a polynomial or similar smooth function in its vicinity. This emphasis on local approximation often led to neglect of the key theoretical restrictions of the Laffer curve, such as revenue approaching zero at 0% and 100% tax rates, the increase in revenue away from these endpoints, and the requirement that revenue-to-GDP ratios cannot exceed the corresponding tax rate. Traditional reduced-form specifications assume functional forms – typically quadratic, cubic, or other low-order polynomials – that do not guarantee consistency with the theoretical curvature of the Laffer relationship (figure 1). The resulting misspecification bias may lead to implausible marginal effects or optimal points.

Figure 1

Common used function forms for the Laffer curve



Note: The graph shows the functions estimated in the quoted empirical papers. Hsing (1996) estimates the log-log model, the semi-log log-linear model and the linear-log models using real income tax revenue per capita as a dependent variable and total tax collected on taxable income as a measure of the average tax rate. Liapis et al. (2020) estimate the linear, the quadratic and the hyperbolic models using the average tax rate as independent variable and the share of income tax revenue as a share of GDP as dependent variable. Akgun, Bartolini and Cournède (2017) estimate the quadratic model using the effective marginal tax rate as independent variable and the share of income tax revenue as a share of GDP as dependent variable; Ferreira-Lopes, Martins and Espanhol (2019) estimate the quadratic model using the implicit tax rate as independent variable and tax revenue as dependent variable; Agersnap and Zidar (2020) estimate the log-log model.

Source: Author's calculation.

The present paper proposes a reduced-form model inspired by the underlying intuition of the elasticity of taxable income (ETI) framework originally developed using microdata. Early studies by Lindsey (1987) and Feldstein (1995) showed that reductions in marginal tax rates led to substantial increases in reported income, especially among top earners, implying that high-income individuals exhibit large behavioural responses. Building on these findings, Saez (2001) developed a theoretical formula linking optimal top marginal tax rates to ETI.

The present paper builds on the elasticity-based intuition underlying much of the empirical Laffer curve literature, while adapting it to the explicit objective of modelling aggregate tax revenue as a share of GDP. To this end, the proposed functional form departs from approaches that focus on behavioural responses to marginal tax rates and instead models the aggregate relationship between average tax rates and total revenues. This modelling choice is also consistent with the practice adopted in general equilibrium approaches that analyse the revenue effects of taxation at the macroeconomic level (see, for example, Mankiw and Weinzierl, 2006; Agell and Persson, 2001; Tsuchiya, 2016; Sanz-Sanz, 2016; and Holter, Krueger and Stepanchuk, 2019). When the object of interest is total revenue, the relevant elasticity is not a marginal elasticity defined at a specific income threshold, but an aggregate elasticity that combines behavioural responses, compositional effects across taxpayers, administrative features of the tax system, and broader macroeconomic adjustments. This distinction is particularly important in the context of personal income tax (PIT) and value added tax (VAT) systems, which feature multiple schedules and multiple statutory or effective rates. In such settings, the use of a single marginal tax rate – common in parts of the empirical literature – is not well suited to capture the effective tax burden faced by the economy as a whole. Estimating aggregate revenue effects instead requires a synthetic measure of average taxation that reflects the diversity of tax rates and taxpayer characteristics. Accordingly, for PIT in the present paper, the simple mean of the average tax wedges for eight representative household types that reflect different income levels and family compositions from the OECD Tax Database was used. Similarly, for VAT the simple mean of the standard, reduced, and higher rates was used. Although simplified, these synthetic indicators provide a parsimonious measure capturing the structural relationship between several tax policy instruments and aggregate revenue.

The model proposed in this paper provides a flexible yet disciplined model that incorporates these theoretical properties while allowing the parameters of the curve to depend on structural economic variables. This improves robustness to misspecification and allows for cross-country heterogeneity in a way that traditional functional forms cannot. The use of a LASSO estimator provides a regularisation mechanism, thus reducing the risk of overfitting.

While the application of LASSO-type regularisation to non-linear econometric models is well established in the broader economics and econometrics literature, to the best of the author's knowledge, no existing study has applied machine learning techniques to the empirical estimation of the Laffer curve. This gap is notable, as the Laffer relationship is characterised by strong non-linearities and binding theoretical constraints, while its empirical functional form remains uncertain.

3 THE DATA

3.1 DEPENDENT VARIABLES

The empirical analysis is conducted on an unbalanced panel of 26 OECD countries covering the period from 2000-21. The resulting sample comprises a total of 529 country-year observations, reflecting heterogeneous temporal coverage rather than deliberate sample selection. Following the methodology in Bloch et al. (2016), the following three public revenue aggregates were chosen: personal income taxes, corporate income taxes revenue, and consumption taxes revenue. To assure comparability among countries the three aggregates were expressed as a share of GDP. A summary description of the three revenue aggregates is shown in table 1.

Table 1

Summary table of the dependent variables

	Mean	SD	Median	Min	Max
Personal income taxes revenue as a % of GDP (revitem1_GDP)	8.91	4.77	7.72	2.77	29.64
Corporate income taxes as a % of GDP (revitem3_GDP)	2.86	1.55	2.60	0.16	12.54
Environmental and Consumption taxes as a % of GDP (revitem4_5_GDP)	12.15	2.47	11.97	5.49	17.95

Source: Author's calculations, based on Bloch et al. (2016).

3.2 TAX RATES

For the estimation of the Laffer curve the choice of the taxation rate of reference is capital for the analysis. Generally, tax systems have a variety of tax rates, and a synthetic indicator is necessary for both the parametric and the non-parametric approach.

In the case of personal income tax, the tax rates may vary based on family composition, income, and region or state of residence. Moreover, due to the progressivity of the taxation systems, taxpayers' incomes are subject to increasing marginal tax rates, according to the income brackets. In the present work, the average tax wedge was chosen from the OECD Tax database[1], in agreement with the reasoning that it is the average tax rate that determines the total tax revenue. This indicator corresponds to the ratio between the amount of taxes paid by the taxpayers and the corresponding total labour cost for the employer, measured in percentage of labour cost. The present work uses a simple mean of the average tax wedges of the following household types: (1) a single person earning 67% of the average wage; (2) a single person earning the average wage; (3) a single person earning 167% of

[1] OECD (2024), Tax wedge (indicator). doi: 10.1787/cea9eba3-en.

the average wage; (4) a single person with 2 children earning 67% of the average wage; (5) a couple without children earning 100% and 67% of the average wage; (6) a couple with 2 children earning 100% and 67% of the average wage; (7) a couple with 2 children both earning the average wage; and (8) a one-earner married couple with two children earning the average wage.

For corporate income tax, the combined corporate income tax rate from the OECD Tax database was used, which combines central and sub-central (statutory) corporate income tax rates.

For VAT some authors, like Matthews (2003) and Akgun, Bartolini and Cournède (2017) used the standard VAT rate because most of the goods are taxed at the standard rate and, for the available data, the correlation between the standard rate of VAT and the effective tax rate is high. For the present work, however a simple mean of the standard, reduced, and higher VAT rates was used from the OECD Tax database. The correlation between this indicator and the standard VAT rate is around 89% and in addiction it captures the balance between reduced and higher VAT rates.

Table 2
Summary table of the tax rates

Variable	Mean	SD	Median	Min	Max
Average tax wedge personal income tax rate (avPIT)	35.45	7.56	36.35	16.11	50.79
Corporate income tax on distributed profit (CIT)	24.42	6.47	25.00	9.00	46.10
Average value added tax rate (avVAT)	13.67	3.05	14.00	4.76	20.56

Source: Author's calculation on OECD Tax database.

3.3 INDEPENDENT VARIABLES

The set of dependent variables includes a set of indicators of quality of the institutions, of economic activity composition, and demography. The quality of institutions such as government effectiveness, control of corruption, regulatory quality, and rule of law, can improve competition among economic actors, increase the cost of tax avoidance and evasion, and the compliance of taxpayers, affecting in this way the scale effect, the revenue maximizing tax rate and the tax base effect. Moreover, the quality of institutions affects competition, which may affect the curvature of the Laffer curve[2].

[2] See, for example, Papp and Takáts (2008), Busato and Chiarini (2013) who analysed the effects of lowering tax rates on tax compliance; Feld and Frey (2002), Berenson (2018), who studied the effects of institutions and trust on implementation of tax policy; Sanyal, Gang and Goswami (2000), who modelled the effects of corruption on tax collections; Miravete, Seim and Thurk (2018), who studied how lack of market competition may flatten the Laffer curve.

The economic activity composition variables such as the decomposition of the GDP using the expenditure and the output approaches could capture the differences in the economic structures of the countries that may affect the structure of the Laffer curve. Some sectors, for example, may have a stronger presence of informality or unregistered activities and therefore be more likely to evade taxation[3]; a country whose GDP is mostly driven by labour intensive sectors may offer less possibility for the evasion or avoidance of personal income taxation; an economy with a strong presence of capital intensive sectors may show strong tax base effects due to an increase in the corporate income tax rate[4]; an economy with strong internal consumption may show smaller tax base effects due to an increase in the VAT rate but its GDP may be affected more. The demographic composition of a country can affect the shape of the Laffer curve due to age and gender wage gaps[5]. Inflation can alter the tax burden through collection lags, absence of indexation, deductibility of expenditures, interest payments on debt[6]. The progressivity indicator of the personal income tax is a measure of the dispersion of the marginal tax rates. It is computed as

$$progPIT = 1 - \frac{1 - margPIT_{single-167}}{1 - margPIT_{single-67}} \tag{1}$$

Higher progressivity of taxation is generally associated with lower tax revenues.[7]

To reduce the omitted variable bias, the set of dependent variables also includes a set of country fixed effects, which capture the effect of non-observable country-specific characteristics that do not vary substantially over the sample period.

Table 3

Summary table of the independent variables

Variable	Mean	SD	Median	Min	Max
Average tax wedge personal income tax rate (avPIT)	35.45	7.56	36.35	16.11	50.79
Corporate income tax on distributed profit (CIT)	24.42	6.47	25.00	9.00	46.10
Average value added tax rate (avVAT)	13.67	3.05	14.00	4.76	20.56
Private final consumption as a % of GDP (CP_GDP)	53.26	7.92	53.38	23.65	71.06

3 See, for example, Heijman and van Ophem (2005).
4 See, for example, Trabandt and Uhlig (2011).
5 See, for example, Lozachmeur (2006), Cataldi, Kampelmann and Rycx (2011), Prammer (2019), who analysed the productivity-wage gap by age and how the effective tax rate changes with age; Creedy et al. (2010), Crowe et al. (2022), who quantified the effects of aging population on public revenue from income and consumption taxation.
6 See, for example, Akgun, Bartolini and Cournède (2017), who found a positive effect of inflation on public revenue from PIT and CIT; Kimbrough (2006), who estimated the revenue-maximising inflation rate.
7 See, for example, Holter, Krueger and Stepanchuk (2019).

Variable	Mean	SD	Median	Min	Max
Government final consumption as a % of GDP (CG_GDP)	20.20	3.38	19.87	10.42	27.86
Gross capital formation as a % of GDP (ITISK_GDP)	23.14	4.78	22.60	11.89	54.43
Exports of goods and services as a % of GDP (XGS_GDP)	57.68	32.10	51.20	18.73	215.35
Imports of goods and services as a % of GDP (MGS_GDP)	54.28	26.65	48.08	22.77	182.15
Trade openness indicator (XMGS_GDP)	111.97	58.41	99.32	45.14	397.50
VA Agriculture, forestry and fishing, as a % of GDP (GDP_agriculture_sh)	2.28	1.37	2.06	0.21	8.36
VA Industry non-manufacturing, as a % of GDP (GDP_industry_nmfc_sh)	4.51	4.46	3.64	1.27	31.38
VA Manufacturing, as a % of GDP (GDP_manufacturing_sh)	16.35	5.36	15.93	5.08	37.31
VA Construction, as a % of GDP (GDP_construction_sh)	6.00	1.72	5.78	1.25	12.42
VA Distributive trade, repairs; transport; accommod., food serv., as a % of GDP (GDP_trade_sh)	20.02	3.94	19.61	8.73	32.03
VA Information and communication, as a % of GDP (GDP_ITC_sh)	4.96	1.60	4.63	2.79	18.83
VA Financial and insurance activities, as a % of GDP (GDP_financial_sh)	6.05	4.65	4.68	1.91	29.82
VA Real estate activities, as a % of GDP (GDP_real_estate_sh)	9.69	2.72	9.35	5.09	18.34
VA Prof., scientific, techn.; admin., support serv. activities, as a % of GDP (GDP_professions_sh)	9.20	2.50	8.84	4.80	16.74
VA Public admin.; compulsory s.s.; education; human health, as a % of GDP (GDP_publ_admin_sh)	18.18	2.74	18.07	9.40	25.13
Government effectiveness (gov_effectiveness)	1.31	0.53	1.36	0.13	2.35
Control of corruption (ctrl_corruption)	1.27	0.75	1.36	-0.19	2.46

Variable	Mean	SD	Median	Min	Max
Political stability and absence of violence/ terrorism (pol_stability)	0.82	0.46	0.90	-1.12	1.76
Regulatory quality (reg_quality)	1.31	0.39	1.29	0.14	2.04
Rule of law (rule_law)	1.30	0.53	1.37	0.05	2.12
Voice and accountability (voice_accountability)	1.23	0.30	1.28	0.34	1.80
Average latitude (av_latitude)	50.86	7.45	50.83	31.50	65.00
Average longitude (av_longitude)	11.00	11.32	12.83	-18.00	34.75
Capital latitude (cap_latitude)	50.51	6.94	50.84	31.78	64.15
Capital longitude (cap_longitude)	11.34	11.47	13.38	-21.94	35.22
Share of population aged 15-24 (%) (pop_sh_1524)	14.58	2.36	14.39	10.37	21.10
Share of population aged 25-64 (%) (pop_sh_2564)	66.42	3.02	65.82	60.31	74.88
Share of population aged 65-69 (%) (pop_sh_6569)	6.15	0.91	6.14	4.15	8.88
Share of population aged 70-74 (%) (pop_sh_7074)	5.20	0.83	5.24	3.07	8.53
Growth of population aged 15-24, y-o-y % change (pop_g_1524)	-0.67	2.36	-0.41	-10.73	16.24
Growth of population aged 25-64, y-o-y % change (pop_g_2564)	0.34	0.98	0.30	-3.52	4.89
Growth of population aged 65-69, y-o-y % change (pop_g_6569)	1.55	3.46	1.48	-15.17	12.98
Growth of population aged 70-74, y-o-y % change (pop_g_7074)	1.50	3.85	1.18	-13.97	27.02
Consumer price index, y-o-y % change (CPI_YTYPCT)	2.07	1.97	1.85	-4.45	15.40
Progressivity indicator of the personal income tax (prog_PIT)	0.07	0.18	0.10	-0.51	0.45
Unemployment rate (%) (UNR)	7.97	4.21	7.04	1.96	27.47
General government gross financial liabilities, as a % of GDP (GGFLQ)	71.28	38.32	63.15	6.80	239.14
Output gap (GAP)	-0.20	3.52	-0.16	-11.79	18.45

Variable	**Mean**	**SD**	**Median**	**Min**	**Max**
Employment gap rate (%) (GAPER)	-0.10	1.33	-0.05	-5.27	4.94
Current account balance, as a % of GDP (CBGDPR)	0.43	6.09	0.78	-22.54	16.48
Employment rate (%) (ER)	45.30	4.54	45.44	32.04	55.34
GDP per capita, constant PPP USD (GDPCAP)	51,900	22,042	49,418	17,088	138,154
Education as a % of GDP (expitem1_GDP)	4.44	0.82	4.41	2.36	6.48
Health as a % of GDP (expitem2_GDP)	5.75	1.65	6.15	0.14	9.38
Other Wages and intermediate consumption as a % of GDP (expitem3_GDP)	9.74	1.89	9.50	3.79	15.41
Old-age and survivor pensions as a % of GDP (expitem4_GDP)	8.88	3.30	8.11	1.51	17.62
Sickness and disability as a % of GDP (expitem5_GDP)	2.41	0.91	2.18	0.85	5.77
Unemployment benefits as a % of GDP (expitem6_GDP)	0.98	0.75	0.75	0.05	3.72
Family and children as a % of GDP (expitem7_GDP)	1.43	0.60	1.46	0.10	3.09
Subsidies as a % of GDP (expitem8_GDP)	1.49	0.87	1.33	0.03	5.74
Investment as a % of GDP (expitem9_GDP)	3.75	1.08	3.74	1.54	7.67
Other primary expenditure as a % of GDP (expitem10_GDP)	4.49	1.90	4.20	1.21	24.55
Property income paid (including interest payments) as a % of GDP (expitem11_GDP)	2.16	1.60	1.88	0.03	12.02
Recurrent taxes on immovable property as a % of GDP (revitem6_GDP)	0.85	0.72	0.59	0.05	3.22
Other property taxes as a % of GDP (revitem7_GDP)	0.81	0.72	0.63	0.00	3.84
Sales of goods and services as a % of GDP (revitem8_GDP)	3.30	1.01	3.24	0.90	7.13
Other primary revenue as a % of GDP (revitem9_GDP)	2.58	1.52	2.22	0.45	10.47
Property income received as a % of GDP (revitem10_GDP)	0.68	1.20	0.35	0.01	12.78

Source: OECD national accounts database; OECD economic outlook database, World Bank governance indicators database; Labour force statistics database; OECD tax database.

PUBLIC SECTOR ECONOMICS
50 (2) 287-319 (2026)

4 THE MODEL

The parametric model proposed in this paper is inspired by the elasticity-based framework first introduced by Lindsey (1987) and later formalised by Saez (2001) and Diamond and Saez (2011) to derive optimal tax rates. In their approach, the elasticity of taxable income is used to identify revenue-maximising and welfare-maximising tax rates, particularly for high-income earners. The proposed model Laffer curve is derived from the ETI approach and adapted to express revenues as a share of GDP instead of local currency units, as in most of the empirical cross-country literature[8]:

$$T_{c,i,t} = \alpha_{c,it} \cdot \left(\frac{\delta_{c,i,t} + \beta_{c,i,t}}{\delta_{c,i,t}} \right)^{\delta_{c,i,t} - 1} \cdot \left(\frac{\delta_{c,i,t} + \beta_{c,i,t}}{\beta_{c,i,t}} \right)^{\beta_{c,i,t}} \cdot \tau_{c,i,t}^{\delta_{c,i,t}} \cdot \left(1 - \tau_{c,i,t} \right)^{\beta_{c,i,t}} + \varepsilon_t \quad (2)$$

where $T_{c,i,t}$ represents revenue item i, as a percentage of GDP.

To generate a subset of curves that accurately reflect the Laffer curve, the following parameter restrictions were imposed: $\alpha \in [0,1]$, $\delta \geq 1$, $\beta > 0$.

Under these restrictions the proposed model has the following properties:

1) Tax revenue remains zero when the tax rate is either 0% or 100%.
   $T\left(\tau_{c,i,t} = 0\%\right) = 0;\ T\left(\tau_{c,i,t} = 100\%\right) = 0$
2) The curve reaches its maximum at $\tau^*_{c,i,t} = \dfrac{\delta_{c,i,t}}{\delta_{c,i,t} + \beta_{c,i,t}}$, with maximum revenue as a share of GDP given by $T\left(\tau^*\right) = \alpha \cdot \tau^*_{c,i,t}$.
3) At any point on the curve, the estimated tax-to-GDP ratio does not exceed the corresponding average tax rate $\left(\dfrac{T\left(\tau_{c,it}\right)}{GDP} \leq \tau_{c,it} \right)$.

The proposed model allows a decomposition of the effect of tax rate changes into three components: a rate effect, a base effect, and a GDP effect[9]:

$$\ln T_1 - \ln T_0 = \underbrace{\ln \tau_1 - \ln \tau_0}_{\text{Rate effect}} + \underbrace{(\delta - 1) \cdot (\ln \tau_1 - \ln \tau_0)}_{\text{GDP effect}} + \underbrace{\beta \cdot \left[\ln(1 - \tau_1) - \ln(1 - \tau_0) \right]}_{\text{Tax base effect}} \quad (3)$$

where:

– The tax rate effect measures the increase in tax revenues due a rise in the tax rate, assuming the tax base and the economy remains unaffected. It depends solely on the initial tax rate, and it is independent of other parameters.

– The tax-base effect captures the responsiveness of the tax base to changes in the tax rate, reflecting tax avoidance, income shifting, deferred income, and the disincentive to generate additional income due to a higher tax burden. β can be interpreted as the ETI, and given the restriction that $\beta > 0$, this effect is always negative, with larger values of β indicating greater sensitivity as found by Keen (1996), Agha and Haughton (1996) and Orsi, Raggi and Turino (2014).

[8] See appendix for the derivation of the model from the ETI approach.

[9] See appendix for the derivation of this decomposition.

– The GDP effect reflects the impact of increased taxation on overall economic activity, capturing general equilibrium effects. δ can be interpreted as the adverse effect on economic efficiency following an increase in the tax level. Given the restriction that $\delta > 1$, this effect is always positive, implying that the model predicts a negative impact of taxation on GDP, with the impact decreasing with tax rate.
– The scale parameter α by contrast measures the relative importance of the tax base in total income generation. Its value is determined by the composition of national income, the progressivity of the tax system, and the capacity of the tax administration to define and enforce the tax base. When α is equal to 1, the tax-revenue-to-GDP ratio is equal to the corresponding tax rate.

Assuming as above that $\beta = \zeta$ and given the constraints on the parameters ($\delta > 1$ and $\beta > 0$), the tax rate that maximises the total revenues expressed in value is below the maximum when the revenues are expressed as a share of GDP.

$$\hat{\tau} = \frac{1}{1+\beta_{c,i,t}} < \frac{\delta_{c,i,t}}{\delta_{c,i,t} + \beta_{c,i,t}} = \tau^{*}_{c,i,t} \tag{4}$$

The difference is made by the contractionary effect of taxation on GDP.

5 THE ESTIMATION METHOD

In principle, the model parameters could be estimated separately for each country. However, due to the limited variation in tax rates over time, country-level estimates have low explanatory power. To better reflect general equilibrium effects and how a country's economic structure shapes its Laffer curve, the parameters $\alpha_{c,i,t}$, $\delta_{c,i,t}$ and $\beta_{c,i,t}$ are modelled as functions of macroeconomic and policy variables.

To ensure that $\alpha_{c,i,t}$ remains within the [0,1] interval, and thus that the tax to GDP ratio is always less than or equal to the tax rate, it is expressed using a logistic transformation:

$$\alpha_{c,i,t} = \frac{e^{\lambda \cdot X_{c,i,t}}}{1+e^{\lambda \cdot X_{c,i,t}}} \tag{5}$$

where λ is a vector of parameters and $X_{c,i,t}$ is the set of variables, discussed above.

The parameters $\beta_{c,i,t}$ and $\delta_{c,i,t}$ are similarly defined as exponential functions of the same set of explanatory variables:

$$\delta_{c,i,t} = 1 + exp\left(\phi \cdot X_{c,i,t}\right),\ \beta_{c,i,t} = exp\left(\theta \cdot X_{c,i,t}\right) \tag{6}$$

where θ and ϕ are vectors of parameters. This ensures that both the elasticity of the tax base and GDP to tax rate is negative.

The model is estimated in levels using a LASSO (Least Absolute Shrinkage and Selection Operator) technique.

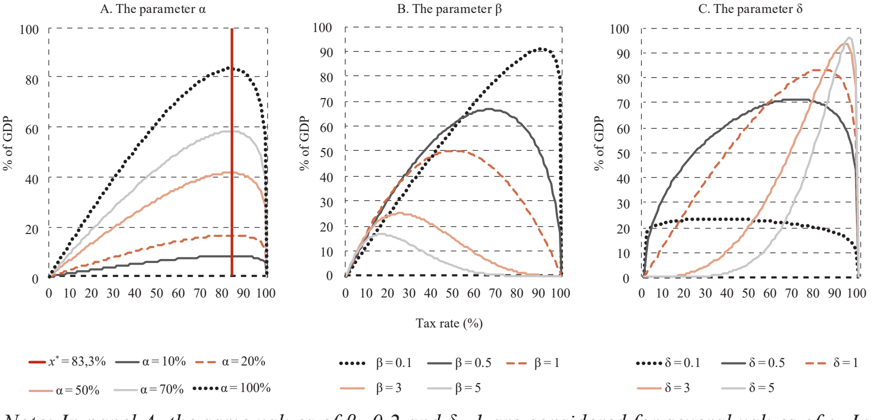
$$\hat{\xi}_{LASSO} = \underset{\xi}{argmin}\left\{\sum\left[T_{c,i,t} - \alpha_{c,it}\cdot\left(\frac{\delta_{c,i,t}+\beta_{c,i,t}}{\delta_{c,i,t}}\right)^{\delta_{c,i,t}-1}\cdot\left(\frac{\delta_{c,i,t}+\beta_{c,i,t}}{\beta_{c,i,t}}\right)^{\beta_{c,i,t}} \cdot\tau_{c,i,t}^{\delta_{c,i,t}}\cdot\left(1-\tau_{c,i,t}\right)^{\beta_{c,i,t}}\right]^2\right\}, \ subject\ to \left\|\xi\right\|_1 < \overline{\xi} \tag{7}$$

where $\alpha_{c,it} = \alpha_{c,i,t} = \frac{e^{\lambda\cdot X_{c,i,t}}}{1+e^{\lambda\cdot X_{c,i,t}}}$, $\delta_{c,i,t} = 1+exp\left(\phi\cdot X_{c,i,t}\right)$, $\beta_{c,i,t} = exp\left(\theta\cdot X_{c,i,t}\right)$, and $\xi = \left[\lambda,\phi,\theta\right]^T$ and all explanatory variables $X_{c,i,t}$ are standardised prior to estimation.

The underlying functional relationship between tax rates and revenues is known and correctly specified, traditional estimators such as ordinary least squares or maximum likelihood are well suited and deliver efficient estimates. In the empirical estimation of Laffer curves, however, the functional form linking tax rates, tax bases, and revenues is inherently uncertain and non-linear. As shown in figure 1, the proposed model is flexible enough to incorporate several functional forms and yet sufficiently disciplined to incorporate the main theoretical properties of the Laffer curve[10].

FIGURE 2

Examples of functional forms for different values of the parameters α, β, δ.



Note: In panel A. the same values of β=0.2 and δ=1 are considered for several values of α. In panel B. the same values of α=100% and δ=1 are used for several values of β. In panel C. the same values of α=100% and β=0.2 are used, for several values of δ.

Source: Author's calculation.

[10] For example, the quadratic form can be obtained from the proposed model with $\delta = 1$, $\beta = 1$.

Moreover, allowing the parameters of the curve to depend on structural economic variables, improves robustness to misspecification and allows for cross-country heterogeneity in a way that traditional functional forms cannot. The use of a LASSO estimator helps prevent overfitting in complex, non-linear models by introducing a penalty term that reduces the variance of the estimates – at the cost of a small bias and reduces model misspecification bias. LASSO is particularly appropriate for forecasting and for identifying parsimonious representations of complex relationships when the true data-generating process is unknown.

Choosing the appropriate regularisation parameter is essential for the performance of the LASSO estimator. In this analysis, the optimal penalty value is selected through k-fold cross-validation. The full dataset contains 529 observations. In each cross-validation fold, 476 observations are used for training and 53 for validation. Since the model includes country fixed effects, the folds are stratified to ensure that each test set contains at least one observation per country. The optimal regularisation value is the one that minimises the Root Mean Squared Error (RMSE) across all test folds. The estimation method does not explicitly take into account the potentially distorting effects of the COVID-19 pandemic. The purpose of the long-term model is to estimate average structural relationships between tax rates and public revenues across a broad range of macroeconomic conditions. Excluding specific years would reduce the model's ability to capture these long-term patterns and weaken its statistical precision. Moreover, removing or adjusting the pandemic years could introduce bias by arbitrarily excluding relevant observations. Any correction would rely on strong counterfactual assumptions, potentially distorting the estimated elasticities more than including the raw data. The use of regularised estimation techniques better mitigates the influence of extreme values by penalising overfitting. This ensures that unusually volatile observations in 2020-2021 do not unduly influence the estimated long-run parameters. The following quantifications are however made for the year 2019 to show the results in the last available year not affected by major economic shocks.

6 ESTIMATION RESULTS

6.1 PERSONAL INCOME TAX

The estimated parameters λ, ϕ and θ allow us to draw the full Laffer curve for each country for all years, and to analyse which variables contribute more to the determination of each parameter. In 2019, according to the model's estimates, 11 countries surpassed their revenue-maximising tax rate, $\tau^{*}_{c,i,t}$, see table 4.

Table 4

PIT, estimated effect of a 1 percentage point increase in the average tax wedge[11]

	PIT revenue as a % of GDP	Average tax wedge (%)	Estimated effect	$\tau^*_{c,i,t}$ (%)	$T(\tau^*_{c,i,t})$ (%)
NOR	10.4	33.0	-0.16	29.0	10.5
ESP	7.6	36.3	-0.13	28.2	8.2
PRT	6.3	35.0	-0.13	27.0	6.9
GRC	6.2	38.5	-0.08	30.2	6.6
ITA	11.2	42.5	-0.08	36.4	10.9
DEU	10.3	43.2	-0.08	39.5	9.9
FRA	9.8	41.7	-0.08	36.6	9.6
GBR	9.1	27.2	-0.04	26.0	9.0
LTU	7.1	32.5	-0.04	30.9	6.2
BEL	11.2	45.8	-0.02	45.1	11.7
LVA	6.5	36.9	-0.02	36.0	6.7
AUT	10.9	40.5	0.04	42.9	10.8
SVK	3.9	37.9	0.07	50.7	4.5
NLD	8.1	30.3	0.08	34.7	8.2
HUN	5.3	38.0	0.10	52.7	6.1
CZE	4.5	37.4	0.13	57.4	6.1
POL	5.3	25.8	0.16	41.3	6.5
EST	5.3	32.0	0.22	54.0	8.4
FIN	12.1	38.4	0.24	51.3	13.9
ISR	6.9	18.4	0.24	28.3	8.6
CHE	10.2	17.6	0.35	26.4	11.4
IRL	6.8	28.0	0.37	66.7	16.9
LUX	9.4	29.0	0.51	72.6	26.6
DNK	27.5	31.3	0.51	41.4	29.8

Note: The table reports the observed average tax wedges (calculated as the simple average among the eight average tax wedges) and tax revenues as a share of GDP for 2019. The third column shows the estimated change in revenue-to-GDP from a one percentage point increase in the average tax wedge. The last two columns present the revenue-maximising tax rate and the corresponding maximum revenue-to-GDP ratio.

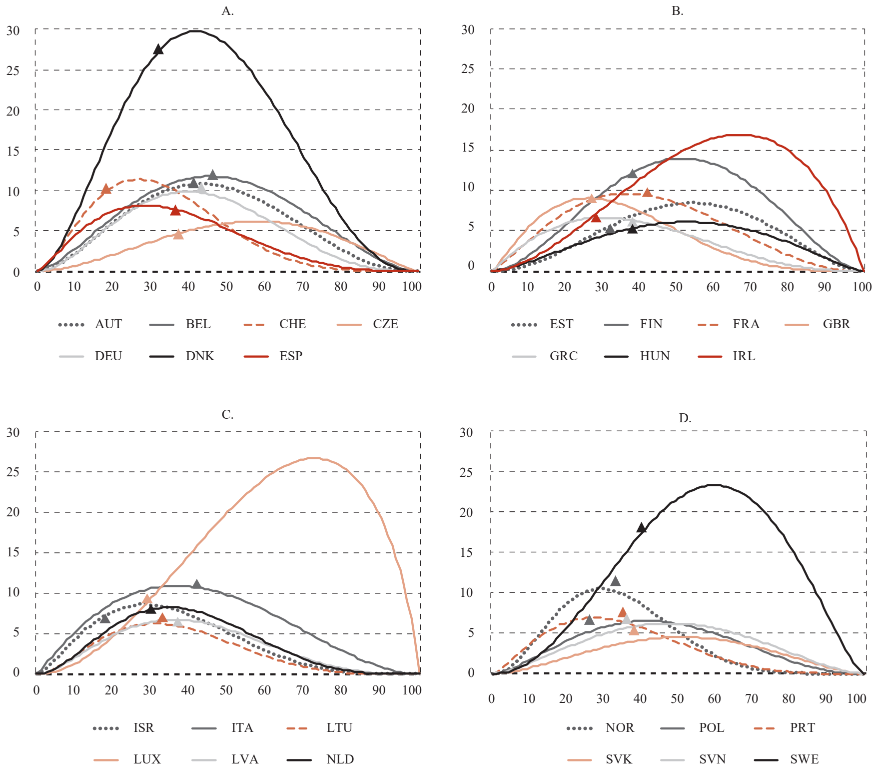
Source: Author's calculations.

The estimated Laffer curves for 2019 together with the observed value of revenues from personal income tax as a share of GDP for the same year are shown in figure 3.

[11] The estimated coefficients are not shown for reason of space, but the author would be happy to share regression results on request.

Figure 3

Personal income tax, the estimated Laffer curves by country, 2019 (% of GDP)



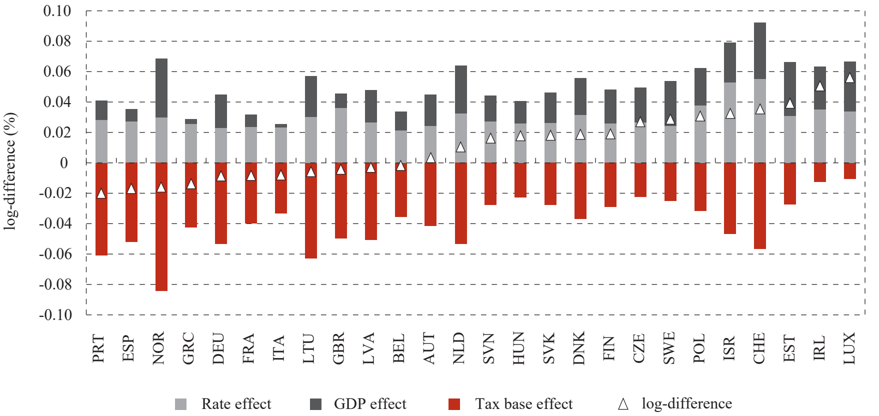
Note: The graphs show the estimated Laffer curves for year 2019 based on the estimated parameters α, β and δ together with the observed value of revenues from personal income tax as a share of GDP for the same year.

Source: Author's calculation.

A decomposition of the estimated effects in term of log differences is shown in figure 4. Countries such as Switzerland, Israel and Poland show the highest tax rate effects due to their relatively low rates, whereas Belgium, Germany and Italy exhibit the highest tax burden on personal income. Countries like Norway, Switzerland and Estonia experience the most significant GDP effects, indicating that an increase in the personal income tax rate would have a more severe negative impact on GDP in these countries. In contrast, Italy, Greece and France are least affected in this regard. Luxembourg, Ireland and Estonia display the largest tax base effects and Norway, Lithuania and Portugal the smallest.

Figure 4

Personal income tax: decomposition of the effect of raising the average tax wedge by 1 percentage point on the tax revenue on GDP ratio



Note: The decomposition is made based on equation (3).

Source: Author's calculations.

6.2 CORPORATE INCOME TAX

The estimated parameters λ, ϕ and θ allow us to draw the full Laffer curve for each country for all years, and to analyse which variables contribute more to the determination of each parameter. In 2019, according to the model's estimates, six countries surpassed their revenue-maximising tax rate, $\tau^*_{c,i,t}$, see table 5.

Table 5

CIT, estimated effect of a 1 percentage point increase in the tax rate[12]

	CIT revenue as a % of GDP	Tax rate (%)	Estimated effect (%)	$\tau^*_{c,i,t}$ (%)	$T(\tau^*_{c,i,t})$ (%)
LUX	6.0	24.9	-0.35	18.2	7.6
DEU	2.0	29.9	-0.03	24.9	2.1
NLD	3.5	25.0	-0.03	23.3	3.2
NOR	6.0	22.0	-0.03	21.5	5.4
FRA	2.3	34.4	-0.02	31.9	2.6
BEL	3.7	29.6	-0.01	29.4	4.0
DNK	3.2	22.0	0.00	22.3	3.3
AUT	3.0	25.0	0.03	28.5	2.7
ITA	1.9	24.0	0.06	36.5	2.7
PRT	3.1	31.5	0.07	41.6	3.8
GRC	2.2	24.0	0.09	49.3	3.2
ESP	2.0	25.0	0.10	40.9	3.5
FIN	2.5	20.0	0.11	29.6	3.3
GBR	2.4	19.0	0.12	33.2	3.6
ISR	3.1	23.0	0.16	40.2	4.7
LVA	0.2	20.0	0.16	54.8	5.9
SVK	3.0	21.0	0.16	34.6	4.7

[12] The estimated coefficients are not shown for reason of space, but the author would be happy to share regression results on request.

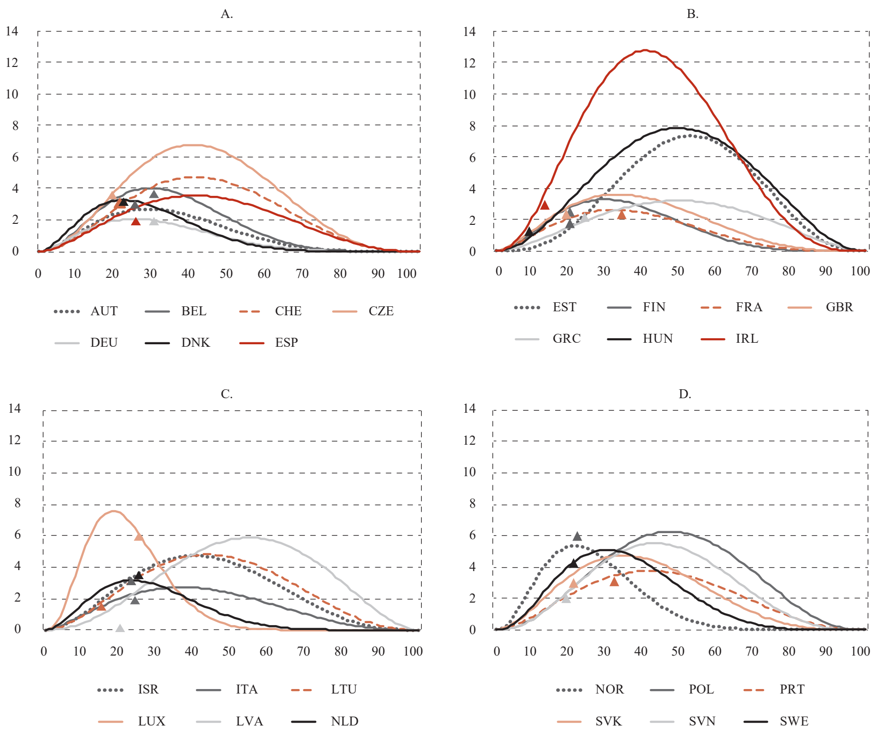
	CIT revenue as a % of GDP	Tax rate (%)	Estimated effect (%)	$\tau^*_{c,i,t}$ (%)	$T(\tau^*_{c,i,t})$ (%)
SWE	4.3	21.4	0.16	30.0	5.1
CHE	3.1	21.1	0.17	41.2	4.7
LTU	1.5	15.0	0.17	43.5	4.8
HUN	1.3	9.0	0.17	49.3	7.8
EST	1.8	20.0	0.19	52.8	7.3
SVN	2.0	19.0	0.21	43.2	5.5
POL	2.2	19.0	0.21	47.3	6.3
CZE	3.6	19.0	0.26	40.5	6.8
IRL	3.0	12.5	0.47	40.5	12.8

Note: The table reports the observed tax rates and tax revenues as a share of GDP for 2019. The third column shows the estimated change in revenue-to-GDP from a one percentage point increase in the tax rate. The last two columns present the revenue-maximising tax rate and the corresponding maximum revenue-to-GDP ratio.

These effects though do not include the effects of tax competition among countries, exposed by the EU Tax Observatory (2004). The estimated Laffer curves for 2019 together with the observed value of revenues from corporate income tax as a share of GDP for the same year are shown in figure 5.

Figure 5

Corporate income tax, the estimated Laffer curves by country (% of GDP)



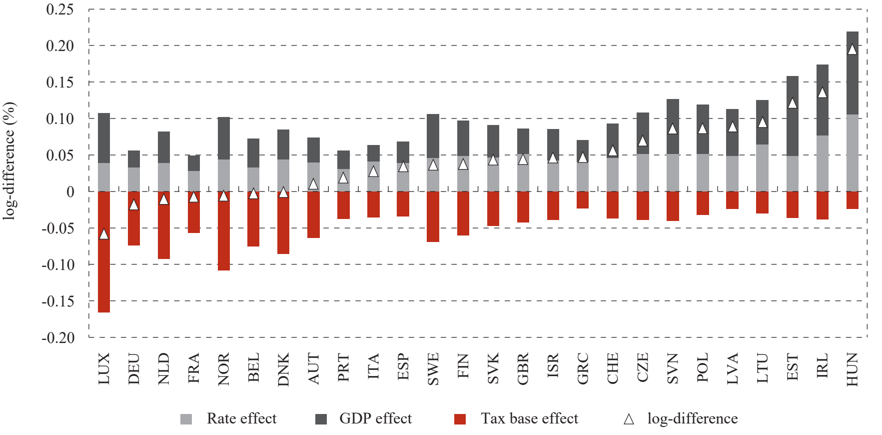
Note: The graphs show the estimated Laffer curves for year 2019 based on the estimated parameters α, β and δ together with the observed value of revenues from corporate income tax as a share of GDP for the same year.

Source: Author's calculation.

A decomposition of the estimated effects in terms of log differences is shown in figure 6. Countries such as Hungary, Ireland, and Lithuania show the highest tax rate effects due to their relatively low rates, whereas France, Portugal, and Germany exhibit the highest tax burden on corporate income. Countries like Hungary, Estonia, and Ireland experience the most significant GDP effects, indicating that an increase in the corporate income tax rate would have a more severe negative impact on GDP in these countries. In contrast, France, Italy, and Germany are least affected in this regard. Hungary, Latvia, and Lithuania display the largest tax base effects, while Luxembourg, Norway, and the Netherlands show the smallest effects.

Figure 6

Corporate income tax, decomposition of the effect of raising the average tax wedge by 1 percentage point on the tax revenue on GDP ratio



Note: The decomposition is made based on equation (3).

Source: Author's calculations.

6.3 VALUE ADDED TAX

The estimated parameters λ, ϕ and θ allow us to draw the full Laffer curve for each country for all years, and to analyse which variables contribute more to the determination of each parameter. According to the estimated model only Norway and Hungary have already surpassed their revenue-maximising VAT rates, $\tau^*_{c,i,t}$, in 2019 (see table 6).

Table 6

VAT, estimated effect of a 1 percentage point increase in the average tax rate[13]

	VAT revenue as a % of GDP	Tax rate (%)	Estimated effect (%)	$\tau^*_{c,i,t}$ (%)	$T(\tau^*_{c,i,t})$ (%)
NOR	11.7	17.4	-0.26	14.6	12.4
HUN	16.8	20.6	-0.11	19.7	16.6
ESP	9.7	14.2	0.08	16.0	10.1
GRC	15.4	17.1	0.10	19.0	15.5
LVA	14.0	15.2	0.15	17.3	14.2
SVK	11.9	14.4	0.18	17.2	11.9
CZE	11.4	14.9	0.25	18.6	12.8
PRT	13.9	16.7	0.29	22.3	14.4
LTU	11.3	13.9	0.37	19.3	13.0
ITA	13.1	14.1	0.40	19.5	13.8
POL	12.8	14.3	0.42	21.0	14.6
SVN	14.1	15.1	0.45	20.9	15.5
NLD	11.6	14.3	0.51	21.5	13.8
AUT	13.2	15.8	0.63	24.9	16.4
FRA	13.6	13.1	0.64	21.8	16.1
EST	13.7	12.9	0.72	22.0	16.8
IRL	6.6	14.8	0.81	35.5	17.6
SWE	17.2	15.8	0.86	23.6	21.0
GBR	11.0	11.1	0.88	26.6	17.8
LUX	9.1	13.4	0.89	31.6	18.0
DEU	10.0	12.3	0.89	24.7	16.2
ISR	11.1	7.6	1.08	20.3	17.9
DNK	14.4	11.1	1.11	27.5	23.6
BEL	10.7	14.0	1.23	29.1	21.5
CHE	5.6	4.9	1.31	26.1	20.4
FIN	14.1	18.0	1.48	35.8	30.7

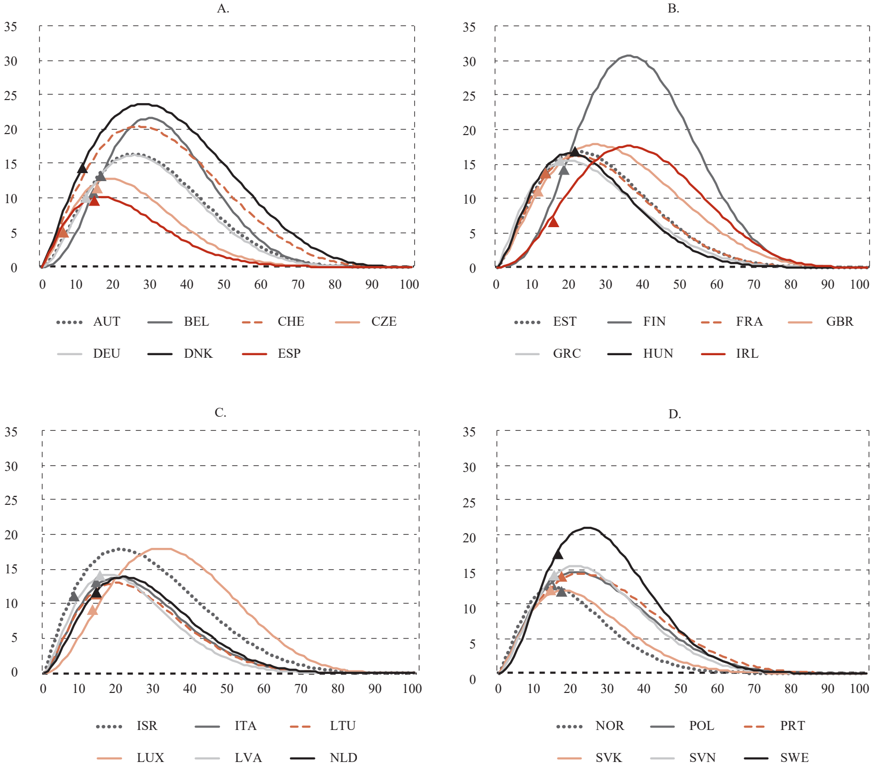
Note: The table reports the observed average tax rates and tax revenues as a share of GDP for 2019. The third column shows the estimated change in revenue-to-GDP from a one percentage point increase in the average tax rate. The last two columns present the revenue-maximising average tax rate and the corresponding maximum revenue-to-GDP ratio.

The estimated Laffer curves for 2019 together with the observed value of revenues from value added tax as a share of GDP for the same year are shown in figure 7.

[13] The estimated coefficients are not shown for reason of space, but the author would be happy to share regression results at the request.

FIGURE 7

Value added tax, the estimated Laffer curves by country (% of GDP)



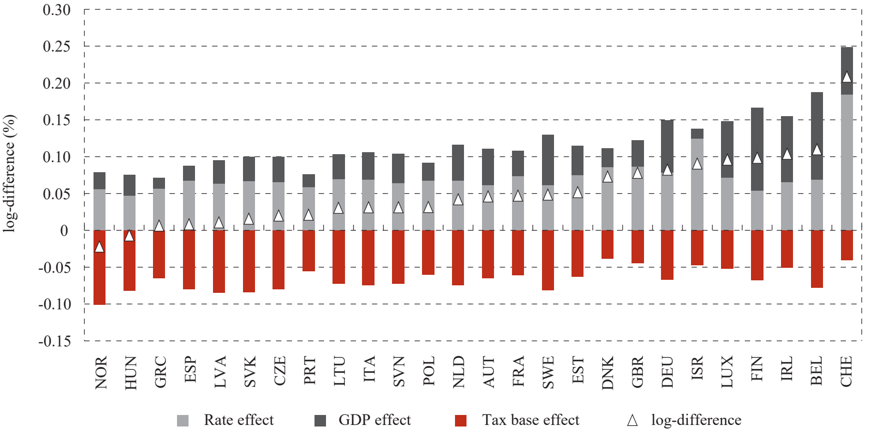
Note: The graphs show the estimated Laffer curves for year 2019 based on the estimated parameters α, β and δ together with the observed value of revenues from value added tax as a share of GDP for the same year.

Source: Author's calculation.

A decomposition of the estimated effects is shown in figure 8. Switzerland, Israel and United Kingdom exhibit the highest tax rate effects due to their low average VAT rates. On the other hand, Hungary, Finland, and Norway bear the highest VAT burdens. Belgium, Finland, and Ireland demonstrate the largest GDP effects, indicating that VAT rate increases would have a more pronounced negative impact on GDP in these countries. By contrast, Israel, Greece, and Portugal show the smallest GDP effects. The countries exhibiting the strongest tax base effects are Denmark, Switzerland, and the United Kingdom while Norway, Latvia, and Slovakia register the lowest tax base effect.

Figure 8

Value added tax, decomposition of the effect of raising the average tax rate by 1 percentage point on the tax revenue on GDP ratio



Note: The decomposition is made based on equation (3).

Source: Author's calculations.

7 COMPARISON WITH ALTERNATIVE MODELS

To evaluate the empirical performance of the proposed framework, the proposed model was estimated alongside several benchmark specifications commonly used in the Laffer-curve literature.

The proposed model systematically outperforms the alternative specifications in terms of out-of-sample fit (table 7).

Moreover, several benchmark specifications generate revenue schedules in which the tax-to-GDP ratio increases monotonically or continues rising beyond theoretically credible ranges (figure 9), whereas the proposed model, by construction, excludes such implausible configurations[14]. The consequences are visible in table 7, which reports the estimated effect of a one-percentage-point increase in the corporate income tax rate for each model and country in the sample. The alternative specifications only predict positive increases in the revenue-to-GDP ratio, with relatively similar magnitudes across countries. In contrast, the proposed model yields more heterogeneous estimated effects, with several countries appearing to operate above their revenue-maximising tax rate[15].

[14] Only 4 countries are shown for reason of space, but similar results are obtained for the other countries.

[15] The table compares the results only for the CIT, but similar results are obtained for the PIT, and VAT.

Table 7

CIT, Estimated effect of a 1 percentage point tax rate increase across models[16]

	Personal income taxes revenue as a % of GDP	Average tax wedge (%)	Proposed model	Log-log model	Semi log-log model	Linear-log model	Linear model	Quadratic model	Hyperbolic model
LUX	6.0	24.9	-0.35	0.07	0.07	0.03	0.04	0.04	0.02
DEU	2.0	29.9	-0.03	0.02	0.02	0.03	0.04	0.03	0.02
NLD	3.5	25.0	-0.03	0.04	0.04	0.03	0.04	0.04	0.02
NOR	6.0	22.0	-0.03	0.06	0.07	0.04	0.04	0.04	0.03
FRA	2.3	34.4	-0.02	0.02	0.03	0.02	0.04	0.02	0.01
BEL	3.7	29.6	-0.01	0.04	0.04	0.03	0.04	0.03	0.02
DNK	3.2	22.0	0.00	0.04	0.05	0.04	0.04	0.04	0.03
AUT	3.0	25.0	0.03	0.03	0.04	0.03	0.04	0.04	0.02
ITA	1.9	24.0	0.06	0.03	0.03	0.03	0.04	0.04	0.03
PRT	3.1	31.5	0.07	0.03	0.04	0.03	0.04	0.03	0.02
GRC	2.2	24.0	0.09	0.02	0.03	0.03	0.04	0.04	0.03
ESP	2.0	25.0	0.10	0.03	0.03	0.03	0.04	0.04	0.02
FIN	2.5	20.0	0.11	0.04	0.04	0.04	0.04	0.04	0.04
GBR	2.4	19.0	0.12	0.04	0.04	0.04	0.04	0.05	0.04
ISR	3.1	23.0	0.16	0.05	0.05	0.04	0.04	0.04	0.03
LVA	0.2	20.0	0.16	0.02	0.02	0.04	0.04	0.04	0.04
SVK	3.0	21.0	0.16	0.04	0.05	0.04	0.04	0.04	0.03
SWE	4.3	21.4	0.16	0.06	0.06	0.04	0.04	0.04	0.03
CHE	3.1	21.1	0.17	0.04	0.04	0.04	0.04	0.04	0.03

[16] The estimated coefficients are not shown for reason of space, but the author would be happy to share regression results on request.

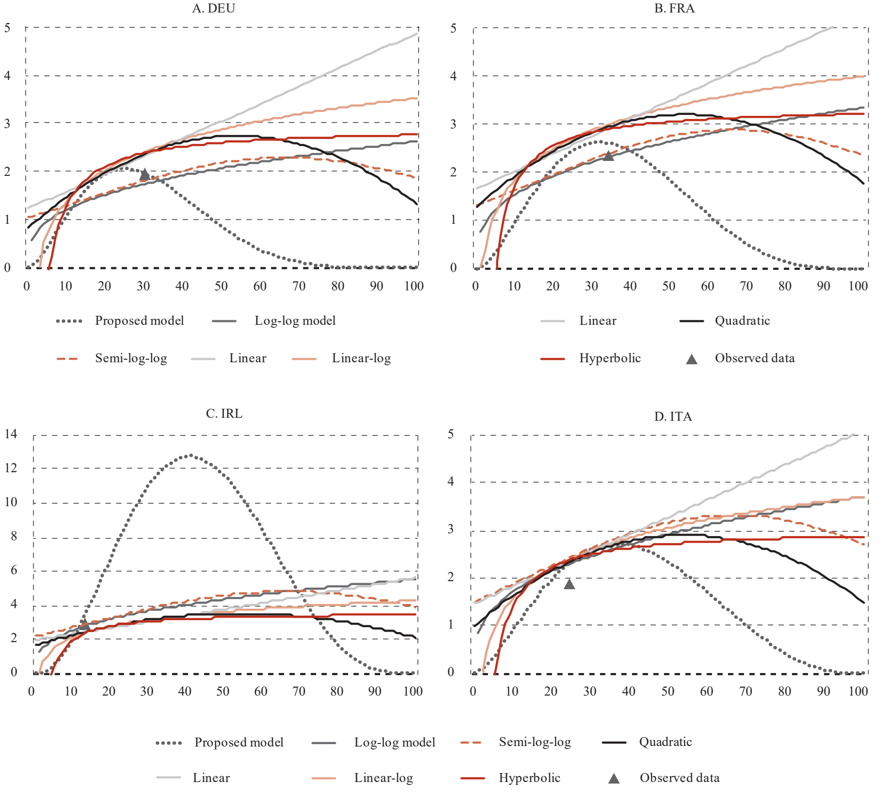
	Personal income taxes revenue as a % of GDP	Average tax wedge (%)	Proposed model	Log-log model	Semi log-log model	Linear-log model	Linear model	Quadratic model	Hyperbolic model
LTU	1.5	15.0	0.17	0.03	0.03	0.06	0.04	0.05	0.07
HUN	1.3	9.0	0.17	0.06	0.04	0.09	0.04	0.06	0.18
EST	1.8	20.0	0.19	0.02	0.02	0.04	0.04	0.04	0.04
SVN	2.0	19.0	0.21	0.03	0.03	0.04	0.04	0.05	0.04
POL	2.2	19.0	0.21	0.04	0.04	0.04	0.04	0.05	0.04
CZE	3.6	19.0	0.26	0.06	0.06	0.04	0.04	0.05	0.04
IRL	3.0	12.5	0.47	0.08	0.06	0.07	0.04	0.05	0.10
RMSE		0.37	0.42	0.42	0.42	0.42	0.42	0.41	0.41

Note: All models are estimated using the same dataset. The table reports out-of-sample root mean squared error (RMSE).

Source: Author's calculations.

FIGURE 9

CIT, the estimated Laffer curves across models for selected countries (% of GDP)



Note: All models are estimated using the same dataset. The graphs display the estimated Laffer curves for 2019, constructed using the estimated parameters, together with the observed corporate income tax revenue as a share of GDP in that year.

Source: Author's calculations.

8 CONCLUSIONS

This paper proposes a parametric framework for estimating Laffer curves that combines theoretical discipline with the flexibility of machine learning techniques. The model is estimated using a LASSO technique, which enables the incorporation of a rich set of structural covariates while mitigating overfitting and reducing sensitivity to variable selection misspecification.

Consistently with the economic literature, the coefficients associated with the determinants of β – which captures the responsiveness of the tax base to tax rates – show that stronger governance and regulatory effectiveness are associated with lower estimated base elasticities, consistent with the interpretation that improved enforcement and compliance reduce the scope for avoidance and evasion. Conversely, variables capturing economic openness or sectoral composition are associated with higher estimated β in the case of corporate taxation, reflecting greater capital mobility or profit-shifting opportunities.

Similarly, the coefficients entering the specification of δ – which captures the responsiveness of GDP to tax changes – show how macroeconomic structure shapes the overall distortionary impact of taxation. The estimated effects of variables such as trade openness, sectoral shares, and labour market indicators suggest that economies with greater exposure to external competition or higher structural flexibility exhibit stronger output responses to tax rate changes.

The empirical results suggest that several OECD countries operate personal and corporate income tax rates above their estimated revenue-maximising levels. By contrast, most countries appear to remain below the revenue-maximising point for value-added tax, suggesting comparatively greater fiscal space in consumption taxation. The findings suggest that in several countries additional revenue mobilisation through higher income tax rates may be limited, whereas moderate adjustments in consumption taxation may yield revenue gains. More importantly, because the curvature of the Laffer relationship depends on structural characteristics, reforms aimed at strengthening institutions, reducing informality, broadening tax bases, and improving compliance may expand revenue capacity without increasing statutory rates.

Overall, the estimated effects of changes in tax rates are broadly consistent with estimates reported in previous empirical and quantitative studies. The present model, though, shows more heterogeneous results in the estimated revenue maximising tax rates. This is because the present framework derives them from the parameters β and δ, which capture the elasticities of the tax base and of the GDP based on economic, institutional and structural variables differently for each country.

In some cases, the estimated revenue-maximising rates are relatively high and substantially greater than the observed tax rates. This is the case for countries such as Czechia, Ireland, and Luxembourg for PIT, of Estonia, Hungary, and Ireland for CIT, and Finland, Ireland, and Luxembourg for VAT. This reflects a relatively high estimated responsiveness of GDP to the average tax rate (δ) together with a relatively low estimated responsiveness of the tax base (β). These results should be interpreted cautiously, as they are derived holding constant the independent variables that determine the estimated parameters α, β and δ. It is implausible to assume that substantial changes in tax rates would not affect these variables and, consequently, change both the shape of the curve and $\tau^*_{c,i,t}$. Estimating the joint effects of tax increases on the dependent variables is beyond the scope of this paper. Moreover, the model does not account for international tax competition effects, which may become substantial in the case of large tax rate increases. However, for small increases in tax rates these effects can reasonably be considered negligible, and the proposed model provides a reliable estimate of the corresponding revenue impact. Similarly, the estimated value of $\tau^*_{c,i,t}$ should not be interpreted as a precise measure of how much tax rates can be increased, but rather as an indication of the proximity of current tax rates to their revenue-maximising level. At the same time, a different value of $\tau^*_{c,i,t}$ would imply a different degree of sensitivity of either the tax base or GDP to changes in the tax rate. These sensitivities are captured by the parameters β and δ, which in the model are estimated as functions of observable institutional, macroeconomic, and structural characteristics. Therefore, different values of $\tau^*_{c,i,t}$ should be mirrored by changes in the underlying institutional, macroeconomic, and structural conditions that shape these parameters. Accordingly, the framework is best interpreted as identifying structural revenue capacity conditional on prevailing institutional and macroeconomic conditions, rather than as prescribing immediate tax policy adjustments.

By enforcing the core theoretical properties of the Laffer curve and allowing its parameters to depend on observable structural characteristics, the proposed framework provides a flexible yet disciplined alternative to traditional reduced-form approaches. In settings where the true functional relationship between tax rates and revenues is unknown, regularisation plays a crucial role in stabilising estimation and reducing specification bias.

Relative to specifications commonly employed in empirical literature, the present framework differs in several respects. It imposes global theoretical consistency by construction. The estimated revenue effects of tax policy changes are more heterogeneous and can be decomposed into a mechanical rate effect, a tax-base effect, and a GDP effect. This heterogeneity is further explained by a set of indicators capturing institutional quality, economic structure, and demographic characteristics.

The framework developed in this paper provides policymakers a tool with which to assess revenue capacity in a manner that is both empirically robust and grounded in economic theory. In the context of rising fiscal pressures related to ageing, climate transition, and public investment needs, these estimates provide a structured benchmark for assessing feasible revenue mobilisation strategies. Future research could extend the framework to incorporate dynamic adjustment processes, international tax competition, or interactions between tax policy and tax administration reforms.

Disclosure statement
The author has no conflicts of interest to declare.

REFERENCES

1. Agell, J. and Persson, M., 2001. On the analytics of the dynamic Laffer curve. *Journal of Monetary Economics,* 48(2), pp. 397-414. https://doi.org/10.1016/S0304-3932(01)00074-5
2. Agersnap, O. and Zidar, O., 2020. The Tax Elasticity of Capital Gains and Revenue-Maximizing Rates. *NBER Working Paper,* No. 27705. https://doi.org/10.3386/w27705
3. Agha, A. and Haughton, J., 1996. Designing Vat Systems: Some Efficiency Considerations. *The Review of Economics and Statistics*, 78(2), pp. 303-308. https://doi.org/10.2307/2109932
4. Akgun, O., Bartolini, D. and Cournède, B., 2017. The capacity of governments to raise taxes. *OECD Economics Department Working Papers,* No. 1407. https://doi.org/10.1787/6bee2df9-en
5. Berenson, M., 2018. Trust and Post-communist Policy Implementation. In: *Taxes and Trust: From Coercion to Compliance in Poland, Russia and Ukraine.* Cambridge University Press, pp. 12-55. https://doi.org/10.1017/9781108333580.002
6. Bloch, D. [et al.], 2016. Trends in Public Finance: Insights from a New Detailed Dataset. *OECD Economics Department Working Papers*, No. 1345. https://doi.org/10.1787/4d3d8b25-en
7. Busato, F. and Chiarini, B., 2013. Steady State Laffer Curve with the Underground Economy. *Public Finance Review,* 41(5), pp. 608-632. https://doi.org 10.1177/1091142113487006
8. Cataldi, A., Kampelmann, S. and Rycx, F., 2011. Productivity-Wage Gaps Among Age Groups: Does the ICT Environment Matter? *De Economist,* 159(2), pp. 193-221. https://doi.org/10.1007/s10645-011-9162-9
9. Creedy, J. [et al.], 2010. Population ageing and taxation in New Zealand. *New Zealand Economic Papers,* 44(2), pp. 137-158. https://doi.org/10.1080/00779954.2010.492574
10. Crowe, D. [et al.], 2022. Population ageing and government revenue: Expected trends and policy considerations to boost revenue. *OECD Economics Department Working Papers*, No. 1737. https://doi.org/10.1787/9ce9e8e3-en
11. De Oliveira, F. G. and Costa, L., 2015. The VAT Laffer curve and the business cycle in the EU27: an empirical approach. *Economic Issues,* 20, pp. 29-44.
12. Diamond, P. and Saez, E., 2011. The Case for a Progressive Tax: From Basic Research to Policy Recommendation. *Journal of Economic Perspectives,* 25(4), pp. 165-190. https://doi.org/10.1257/jep.25.4.165
13. EU Tax Observatory, 2004. *Global Tax Evasion Report*.
14. Feld, L. P. and Frey, B. S., 2002. Trust breeds trust: How taxpayers are treated. *Economics of Governance,* 3(2), pp. 87-99. https://doi.org/10.1007/s101010100032
15. Feldstein, M., 1995. The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act. *Journal of Political Economy,* 103(3). https://doi.org/10.1086/261994

16. Ferreira-Lopes, A., Martins, L. F. and Espanhol, R., 2019. The relationship between tax rates and tax revenues in eurozone member countries – exploring the Laffer curve. *Bulletin of Economic Research,* 72(2), pp. 121-145. https://doi.org/10.1111/boer.12211
17. Heijman, W. J. M. and van Ophem, J. A. C., 2005. Willingness to pay tax: The Laffer curve revisited for 12 OECD countries. *The Journal of Socio-Economics,* 34(5), pp. 714-723. https://doi.org/10.1016/j.socec.2005.07.013
18. Holter, H. A., Krueger, D. and Stepanchuk, S., 2019. How do tax progressivity and household heterogeneity affect Laffer curves? *Quantitative Economics,* 10(4), pp. 1317-1356. https://doi.org/10.3982/QE653
19. Hsing, Y., 1996. Estimating the laffer curve and policy implications. *Journal of Socio-Economics,* 25(3), pp. 395-401. https://doi.org/10.1016/S1053-5357(96)90013-X
20. Keen, M. [et al.], 1996. The Future of Value Added Tax in the European Union. *Economic Policy,* 11(23), pp. 373-420. https://doi.org/10.2307/1344708
21. Kimbrough, K. P., 2006. Revenue maximizing inflation. *Journal of Monetary Economics,* 53(8), 1967-1978. https://doi.org/10.1016/j.jmoneco.2005.07.023
22. Liapis, K. J. [et al.], 2020. Investigating the relationship between tax revenues and tax ratios: an empirical research for selected OECD countries. *International Journal of Economics and Business Administration,* 8(1), pp. 215-229. https://doi.org/10.35808/ijeba/420
23. Lindsey, L. B., 1987. Individual taxpayer response to tax cuts: 1982-1984: with implications for the revenue maximizing tax rate. *Journal of Public Economics*. https://doi.org/10.3386/w2069
24. Lozachmeur, J. M., 2006. Optimal Age-Specific Income Taxation. *Journal of Public Economic Theory,* 8(4), pp. 697-711. https://doi.org/10.1111/j.1467-9779.2006.00284.x
25. Mankiw, N. G. and Weinzierl, M., 2006. Dynamic scoring: A back-of-the-envelope guide. *Journal of Public Economics*, 90(8-9), pp. 1415-1433. https://doi.org/10.1016/j.jpubeco.2005.11.006
26. Matthews, K., 2003. VAT evasion and VAT avoidance: Is there a European Laffer curve for VAT? *International Review of Applied Economics,* 17(1), pp. 105-114. https://doi.org/10.1080/713673162
27. Miravete, E., Seim, K. and Thurk, J., 2018. Market Power and the Laffer Curve. *Econometrica,* 86(5), pp. 1651-1687. https://doi.org/10.3982/ECTA12307
28. Orsi, R., Raggi, D. and Turino, F., 2014. Size, trend, and policy implications of the underground economy. *Review of Economic Dynamics,* 17(3), pp. 417-436. https://doi.org/10.1016/j.red.2013.11.001
29. Papp, T. K. and Takáts, E., 2008. Tax Rate Cuts and Tax Compliance: The Laffer Curve Revisited. *IMF Working Papers,* WP/08/7. https://doi.org/10.5089/9781451868692.001
30. Prammer, D., 2019. How does population ageing impact on personal income taxes and social security contributions? *The Journal of the Economics of Ageing,* 14, 100186. https://doi.org/10.1016/j.jeoa.2018.12.005

31. Saez, E., 2001. Using Elasticity to Derive Optimal Income Tax Rates. *Review of Economic Studies,* 68, pp. 205-229.
32. Sanyal, A., Gang, I. N. and Goswami, O., 2000. Corruption, Tax Evasion and the Laffer Curve. *Public Choice,* 105(1/2), pp. 61-78. https://doi.org/10.1023/A:1005105822911
33. Sanz-Sanz, J. F., 2016. The Laffer curve in schedular multi-rate income taxes with non-genuine allowances: An application to Spain. *Economic Modelling,* 55, pp. 42-56. https://doi.org/10.1016/j.econmod.2016.01.024
34. Trabandt, M. and Harald, U., 2011. The Laffer curve revisited. *Journal of Monetary Economics*, 58(4), pp. 305-327. https://doi.org/10.1016/j.jmoneco.2011.07.003
35. Tsuchiya, Y., 2016. Dynamic Laffer curves, population growth and public debt overhangs. *International Review of Economics & Finance*, 41, pp. 40-52. https://doi.org/10.1016/j.iref.2015.10.001

APPENDIX

DERIVATION OF THE MODEL

A similar approach can be adopted to identify the optimal average tax rate. Given an initial tax rate τ_0 and total taxable income $B(\tau_0)$, an increase in the tax rate from τ_0 to τ_1 has two effects on tax revenues: a tax rate effect given by the increase in tax revenues at constant taxable income and a tax-base effect. As in the ETI model the latter can be assumed to be:

$$B(\tau_1) = B(\tau_0) \cdot \left(\frac{1-\tau_1}{1-\tau_0}\right)^{\zeta} \tag{A1}$$

where ζ is the elasticity of taxable income. Therefore, the total change in tax revenues can be written as:

$$\frac{R_1}{R_0} = \frac{\tau_1 \cdot B(\tau_1)}{\tau_0 \cdot B(\tau_0)} = \frac{\tau_1}{\tau_0} \cdot \left(\frac{1-\tau_1}{1-\tau_0}\right)^{\zeta} \tag{A2}$$

Which in logarithms becomes:

$$\ln(R_1) - \ln(R_0) = \ln(\tau_1) - \ln(\tau_0) + \zeta \cdot \left[\ln(1-\tau_1) - \ln(1-\tau_0)\right] \tag{A3}$$

This can be expressed in differential terms as:

$$\frac{d}{d(\tau)} \ln(R(\tau)) = \frac{d}{d(\tau)} \cdot \left[\ln(\tau) + \zeta \cdot \ln(1-\tau)\right] \tag{A4}$$

Solving this equation one obtains:

$$\ln(R(\tau)) = \left[\ln(\tau) + \zeta \cdot \ln(1-\tau)\right] + C \Rightarrow R(\tau) = A \cdot \tau \cdot (1-\tau)^{\zeta} \tag{A5}$$

Where A is a scaling parameter and whose maximum is at:

$$\hat{\tau} = \frac{1}{1+\zeta} \tag{A6}$$

The model proposed is similar to equation (A5), where the scaling parameter A is replaced by a function of the average tax rate, which enforces the core properties of the Laffer curve and allows revenues to be expressed as a share of GDP instead of local currency units, as in most of the empirical cross-country literature:

$$T_{c,i,t} = \alpha_{c,it} \cdot \left(\frac{\delta_{c,i,t} + \beta_{c,i,t}}{\delta_{c,i,t}}\right)^{\delta_{c,i,t}-1} \cdot \left(\frac{\delta_{c,i,t} + \beta_{c,i,t}}{\beta_{c,i,t}}\right)^{\beta_{c,i,t}} \cdot \tau_{c,i,t}^{\delta_{c,i,t}} \cdot \left(1-\tau_{c,i,t}\right)^{\beta_{c,i,t}} + \varepsilon_t \tag{A7}$$

where $\tau_{c,i,t}$ is represents revenue item i, as a percentage of GDP.

From equation A2, the change in tax revenues as a share of GDP can be written in the form:

$$\frac{T_1}{T_0}=\frac{\dfrac{R_1}{GDP(\tau_1)}}{\dfrac{R_0}{GDP(\tau_0)}}=\frac{GDP(\tau_0)}{GDP(\tau_1)}\cdot\frac{\tau_1}{\tau_0}\cdot\left(\frac{1-\tau_1}{1-\tau_0}\right)^{\zeta} \tag{A8}$$

While from equation (2), omitting the country-time indices, it can be written:

$$\frac{T_1}{T_0}=\left(\frac{\tau_1}{\tau_0}\right)^{\delta}\cdot\left(\frac{1-\tau_1}{1-\tau_0}\right)^{\beta} \tag{A9}$$

Equating the two previous equations:

$$\frac{GDP(\tau_0)}{GDP(\tau_1)}\cdot\frac{\tau_1}{\tau_0}\cdot\left(\frac{1-\tau_1}{1-\tau_0}\right)^{\zeta}=\left(\frac{\tau_1}{\tau_0}\right)^{\delta}\cdot\left(\frac{1-\tau_1}{1-\tau_0}\right)^{\beta} \tag{A10}$$

If one assumes $\beta=\zeta$. Then

$$\frac{B(\tau_1)}{B(\tau_0)}=\left(\frac{1-\tau_1}{1-\tau_0}\right)^{\beta} \tag{A11}$$

and

$$\frac{GDP(\tau_0)}{GDP(\tau_1)}=\left(\frac{\tau_1}{\tau_0}\right)^{\delta-1} \tag{A12}$$

Passing equation to logarithmic form and using equations A10 and A11, it is possible to decompose the change in tax revenues on GDP between two points in time τ_0 and τ_1 into the rate effect the tax-base effect and, the GDP effect in the following way.

$$lnT_1-lnT_0=\underbrace{ln\tau_1-ln\tau_0}_{Rate\ effect}+\underbrace{(\delta-1)\cdot(ln\tau_1-\ln\tau_0)}_{GDP\ effect}+\underbrace{\beta\cdot\left[ln(1-\tau_1)-ln(1-\tau_0)\right]}_{Tax\ base\ effect} \tag{A13}$$

United Nations World Public Sector Report 2025: Supreme Audit Institutions and the Sustainable Development Goals

United Nations, Division for Public Institutions and Digital Government, Department of Economic and Social Affairs, New York, 2025, pp. 214
https://publicadministration.desa.un.org/publications/world-public-sector-report-2025

Book review by DAGMAR RADIN*
https://doi.org/10.3326/pse.50.2.6

* Received: February 23, 2026
Accepted: February 24, 2026

Dagmar RADIN, Ph.D.
Faculty of Political Science, University of Zagreb, Lepušićeva 6, 10000 Zagreb, Croatia
e-mail: Dagmar.Radin@fpzg.hr
ORCiD: 0000-0002-8636-9516

PUBLIC SECTOR ECONOMICS
50 (2) 321-328 (2026)

The focus of the World Public Sector Report 2025 is on the role that Supreme Audit Institutions (SAI) have on the implementation of the Sustainable Development Agenda 2030. Given the short period left for the attainment of the goals set by the Agenda, the role of SAIs and their contribution towards achieving them is ever more important. With the enforcement of the Agenda, the International Organization of Supreme Audit Institutions (INTOSAI) was launched as an autonomous and non-political umbrella organization that would provide external support to national supreme audit institutions and play a consultative role for the UN Economic and Social Council (ECOSOC) in the monitoring of the implementation of the Agenda. The present report gives a first of its kind overview of the role, work and influence of SAIs on the success of the implementation of the SDGs. It begins with an overview of the four chosen areas where SAIs have conducted work since 2016, followed by four thematic chapters including the development of SAIs, the challenges they faced, the conclusions and recommendations in their audit reports as well as major contributions of the SAIs in different areas, and offering a prospective outlook on their role.

The first chapter, written by David Le Blanc, focuses on the role of SAIs in national governance explaining how they function to ensure transparency, accountability, accuracy of public accounts and compliance with the law, despite national differences in size and capacity. While their primary mission was initially to ensure financial compliance, over time their scope has expanded to include performance audits to ensure efficiency and effectiveness of public spending, including during the COVID 19 pandemic as well as monitoring the implementation of the 2030 UN Agenda, which is the focus of this report. The chapter notes that the experiences and results of the national audit institutions vary considerably because of the countries' differing capacities and sizes, as well as their independence in interactions with the Executive and the resources available to them. This experience – as also shown by the results of research[1] on anticorruption initiatives – is that the level of independence and resource allocation signals whether governments are committed to the cause. This assessment is based on two sets of data, including the INTOSAI 2023 Global Survey as well as the United Nations Department of Economic and Social Affairs (UNDESA) 2024 Survey of the INTOSAI members. The report shows that while two thirds of SAIs work within environments which to a great extent prescribe their independence, least developed countries (LDCs) and small island developing states (SIDSs) lag behind. Furthermore, when the percent of SAI recommendations adopted is observed, results show that fewer than 5% of countries have adopted them in full, while 45% have adopted them in part, revealing that there is still room for improvement. When it comes to the integration of SDG monitoring in their work, 37% of INTOSAI members claim to have broadly integrated the SDGs while 22% have integrated them in a very limited way. When looking at a broader group of SAIs involved in the wider survey, 50% undertook performance audits on the

[1] Maor, M., 2004. Feeling the heat? Anticorruption mechanisms in comparative perspective. *Governance,* 17(1), pp. 1-28; Brown, A. J. and Heinrich, F (2017). National Integrity Systems – An evolving approach to anti-corruption policy evaluation. *Crime, Law and Social Change,* 68(3), pp. 283-292.

preparedness of governments to implement SDGs while over 55% contributed to the implementation of SDG 16, ensuring peace, justice and promoting strong institutions. What the report noted is that in part, the priorities of activities that the SAIs undertook were guided by specific requests from parliaments or that they focus on certain national priorities. The chapter also reports examples of direct and indirect engagement with SDGs on the positive side, while on the opposite side SAIs more critical of and sceptical concerning their integration. Overall, the SDGs most covered by SAIs globally are 3, 6, and 15, with over 75% of SAIs reporting that they have covered them since 2016, while SDG 10 and SDG 17 are the least covered by SAI audits. To conclude, the role and mission of the SAIs have been growing within the limited contexts in which they operate but the benefits of engaging with the monitoring of SDGs has benefitted both national governments and the SAIs themselves in their greater role as evidenced by the changes in the perceptions of their relationships with other government institutions.

The second chapter, written by Julie Powell, explains the process the SAIs undertook in conducting preparedness audits, the preparation steps, the methodology used, as well as the findings of the audits. The SDG preparedness audits were done as performance audits in the period between 2016 and 2019. The data used in the analysis in this chapter come from a 2019 INTOSAI Development Initiative report, from the INTOSAI 2024 survey as well as from the analysis of performance audits and interviews with SAI representatives. The auditing project was started by UNDESA, which also created a guideline document to support the SAIs in the implementation of the audits to streamline the process and make it more uniform. While the focus later shifted to the implementation audits, the role of the performance audits served to shed more light on the SDG integration as well as to help governments determine where attention was needed to successfully implement the 2030 Agenda. In addition to focusing on the auditing of institutional preparedness, on the adequacy of budgeting and resource allocation as well as the coherence of policies with the SDGs, the SAIs took a "whole government" approach, that is, whether governments included an interministerial coordination mechanism, central planning units and other cross-sector coordination mechanisms, since without proper coordination the implementation of the 2030 Agenda would remain fragmented and limited in its success. The chapter goes on to explain the methodology used in the performance audits, including desk review, interviews and focus groups as well as observations and questionnaires and citizen surveys. Some of the tools used by the SAIs included various analyses including of budget, policy, risk, gap, data, and root cause. A separate section describes the challenges of performing audits including absence and quality of data, taking in categories such as missing baseline data, incompatible databases, insufficient disaggregation, weak IT infrastructure, limited statistical capacity, government resistance and political sensitivity, time and resource constraints and difficulties in ensuring multi-stakeholder engagement. SAIs also experienced enabling factors including (but not limited to): commitment by the SAI staff, auditors experience, INTOSAI priorities, and collaboration with other SAIs. The key findings of the SAI performance, based on the analysis of the 62 audits, indicate that the results were remarkably similar across both the developed and developing countries. The most frequent categories of audit observations included focus on whether governments had set in place monitoring systems, coordinating structures, budgets and the alignment of planning within their policies. The positive findings indicate that many governments were formally committed to SDGs and had begun integrating them into national development plans, sector strategies and programs, and many consulted stakeholders during planning. However, they also identified partial or selective alignment with SDGs, and examples where SDGs had not been referenced at all in national strategies. When it comes to policies and regulations, some countries issued presidential regulations and embedded SDGs into legal structures, but struggled with the adaptation of policies to local contexts, and had limited resources for policy reform. A major recurring issue in institutional coordination was that in several countries no central SDG coordination body existed, while those that existed were poorly structured and lacked clear mandates. Another major area of weakness is budgeting and resources: audits showed that many countries had no assessment of required financial resources, no long-term resource mobilization strategy, paid insufficient attention to human resource capacity, and that ministries lacked staff competencies. In summary, while most governments showed political commitment to the SDGs, many lacked operational readiness. The last section of the chapter indicates that the audits had a positive impact on both the national government and the SAIs performing the audit. While the governments became more aware of the steps they needed to take to ensure better integration in the preparation phase, the benefits to the SAIs included better communication and reach with other stakeholders throughout government as well as increasing skill-based resources beyond influencing expertise.

The third chapter, written by Aránzazu Guillán Montero, looks at the effects of performance auditing on particular goals and presents the contribution of SAIs to the creation of stronger national budgets and sounder public financial management. SAIs are crucial in providing fiscal oversight though performing public finance audits and their role in ensuring transparent and responsible budgeting. This role has remained important because the effective and efficient use of public funds is fundamental in the successful inclusion and implementation of the SDGs. The data from the audits show significant divergencies between planned and executed public budgets which can have a significant impact on universalistic programs such as health care and education. The most represented categories in public finance audits are emergency spending related to the COVID pandemic and the collection of taxes and revenues, while the least represented is the performance audit of tax policy. The focus on the COVID pandemic is not surprising and is warranted given the large amounts of public funds that were spent under emergency conditions and the potential for their ineffective use. The author of the chapter also visually presents a timeline infographic with milestones reached by

PUBLIC SECTOR ECONOMICS
50 (2) 321-328 (2026)

INTOSAI on public debt spanning from 1991 until 2025, given the number of activities and roles SAIs have played over time. The mapping of the SAIs' work on public finance shows that audits focused for the most part on four areas: procedures for managing and monitoring financial allocations, generation, capture and management of performance information, conformity with accounting standards, and quality and reliability of forecasts. In the public debt area, the audits looked at debt management, debt servicing activities and organizational arrangements related to debt. Despite their significant work the audits show that direct reference and links to the SDGs in this segment of audits is scarce, except for a few countries, such as Austria. The chapter describes the methods used in the public finance audits, looking at frequency, type and processes. The chapter notes that the scope of the audits varies significantly given the different objectives and levels of analysis. However, what many of them have in common is the focus on central financial institutions, while some extend to other responsible institutions, and implementing agencies. Audits reveal innovative methodologies and data analytics such as in Indonesia which used big data to analyse and compare planning and budgeting, or in Brazil which conducted a study to assess the best way to automate the audit processes. The chapter lists several challenges and opportunities, as well as strengths identified in the financial management and public debt external audits. It also identifies the top ten limitations in public debt management, with rising debt appearing as, the prime concern. Other weaknesses include incomplete and inconsistent records and weak legal frameworks. It concludes with recommendations and highlights the impacts of SAI audit reports on the public sector thought to provide more effective implementation by enabling stakeholders to strengthen their fiscal responsibility using the findings from the audits.

The core principle of the 2030 Agenda is to leave no one behind (LNOB). Thus, it was not important only to work towards achieving the SDGs, but to also reach out and include particularly those who are the furthest behind. With that in mind, the fourth chapter, written by Lisa Ainbinder, focuses on how SAI external performance audits advance equity, quality and inclusion. It is based on 145 SAI reports covering 34 countries and one territory and multi country audits, as well as literature, and is complemented by a series of interviews with experts conducted between November 2024 and March 2025. The chapter begins with the analysis of the approach and operationalization used in the audit reports, indicating that the role of SAIs in integrating the principles in the audit reports was important, among other things, to forge better trust between governments and marginalized groups. While some SAIs did not see the LNOB principle as a key priority and treated it as a cross-cutting issue with minimal attention, in some others it has been a long-standing priority. Only 31 percent of SAIs have carried out a gender audit with only 21 percent of them having integrated gender mainstreaming into their auditing. This shows that there is much work to be done. Most SAIs carried out an audit on the equity, equality and inclusiveness of the institutions responsible for caring for underprivileged groups and over time referencing to quality and inclusion has grown in strategic planning, while 40 percent of strategic plans did not address gender at all. Some SAIs, such as those of Thailand, Canada and New Zealand have committed to aligning audit processes in such a way as to include underprivileged groups. The methodology used in these audits followed the general audit methods, both quantitative and geospatial data analysis. Of interest are also the new tools such as the Gender Equality Audit Topic Selection Screening Tool developed by the SAIs of Great Britain and Ireland. Furthermore, some SAIs developed policies and strategies to promote equality, equity and inclusion more broadly such as those in the Maldives, Brazil and Rwanda.

A further consideration in this chapter is the diversity of stakeholders involved in these audits showing that among stakeholders, 38 percent were composed of communities, 28 percent of NGOs, and 14 percent of representatives from academia and research institutions. The challenges of applying LNOB in auditing included lack of experience in many SAIs, staff capacity and responsiveness, and the lack of understanding of government bodies of the need to perform such audits by the SAIs. Inadequate data disaggregation continues to represent a significant constraint, as granular data are necessary to identify populations excluded from public services and those disproportionately affected by inequality. The absence of well-defined national targets, together with weak or underdeveloped progress indicators, further complicates assessments. In mapping out the topics of 145 audits, 31 addressed employment, 23 social protection and 23 education, while disaster services, the environment and sanitation, each had two audits. Poverty, an ever-present issue, was addressed in only six out of 145 audits. Romania is singled out with an audit on access to water and sanitation services by using geographic information system (GIS) to examine access in urban and rural area. When it comes to persons with disabilities, in the 27 audits social protection and education were the most prominent topics. The audits also identified strengths, and these include planning, implementation and monitoring, reporting and follow up as the top three. These three also fall within the challenges commonly identified by the audits. As for the developed vs developing economies divide, more developed economies face challenges of poor planning and gaps in oversight and monitoring than their developing counterparts, while developing economies face much greater challenges in implementation. Finally, which may be of most interest, is what impact the SAIs' audit reports actually had on equity, equality and inclusiveness. They raised awareness of LNOB issues, strengthened analytical practices, and in some cases influenced government action through policy changes, improved service delivery, stakeholder engagement, and public outreach.

The fifth chapter, written by Aránzazu Guillán Montero, examines how SAIs contribute to strengthening government accountability, effectiveness, and policy coherence in addressing climate change (SDG 13). It argues that climate governance presents unusually complex audit challenges because it is cross-sectoral, long-term,

PUBLIC SECTOR ECONOMICS
50 (2) 321-328 (2026)

data-intensive, and characterized by uncertainty, multi-level implementation, and substantial fiscal risks. The chapter analysis draws from relevant literature, audit reports (176 reports from 61 countries, in the 2010-2024 period), and recent expert interviews. The initial overview of the role of INTOSAI in climate action auditing presents an infographic timeline with the steps taken from 1992 when the INTOSAI Working Group on Environmental Auditing was established, until 2024 when the Assembly Resolution on role of SAIs regarding climate change took place.

The analysis shows that, over time, the scope of the SAI audits has expanded to include national climate strategies, policies on adaptation and resilience, allocation of funds for climate action, the analysis of targeted sectoral policies such as transportation and energy, disaster risk reduction and preparedness. The results show that there is a shift in national priorities and top issues identified when it comes to climate problems. For example, while drinking water quality and supply were the top national priority identified until 2014, they were replaced by “climate change” in the period 2024-2026. Similarly, while the top issue in 2018-2020 was protected areas and natural parks, there was a shift towards climate change adaptation in the recent period. This change in focus may reflect changes and advances in the understanding of climate change worldwide on the one hand, but also national priorities as countries are faced with different climate-related challenges. Developing countries were more focused on climate mitigation and adaptation and issues related to agriculture and forests, protected areas and land use, followed by climate finance and disaster risk management, while developed countries were more focused on climate mitigation, energy and climate finance.

Taking into consideration that different countries face different vulnerabilities and priorities with respect to climate change, many SAI reports revealed weaknesses such as institutional fragmentation with weak coordination mechanisms. Measurable indicators and reporting systems were also often missing, compounded by missing and inconsistent financial data. In fact, a major focus of this chapter is on climate finance oversight where SAI audits focused on the review of whether climate funds reached intended programs, the transparency of green budgeting, tracking of mitigation vs adaptation spending, the effectiveness of subsidies and incentives and the risk of waste or misclassification of “green” expenditures. The chapter stresses that climate spending requires strong public financial management and traceability systems. The chapter also identifies the challenges faced by the SAIs in performing the audits, including the shortage of climate experts, the rapidly evolving regulations and technologies, the political sensitivity of the issue that made interaction with stakeholders challenging, and the lack of a standardized methodology.

The concluding chapter, written by David Le Blanc, highlights the key trends of the SAIs and the SDGs since 2016 in terms of their coordination, scope, impact, engagement with stakeholders, and limitations as well as advances in methodologies. It closes with key takeaways from the first review and four thematic chapters.

PUBLIC SECTOR ECONOMICS
50 (2) 321-328 (2026)

In summary, the World Public Sector Report 2025 highlights the role of SAIs in the advancement of the 2030 Agenda by showing through four key thematic topics how they proceeded, what challenges they faced, what the results of the audits show in terms of progress made and how much of the journey still remains ahead, as well as how the actual processes of the SAI audits and their findings contributed to progress in the 2030 Agenda.

Institute of
Public Finance