Public Sector Economics

2/2023

MAJA MATANIĆ VAUTMANS, MARIJANA OREB and SAŠA DREZGIĆ: Socioeconomic inequality in the use of long-term care for the elderly in Europe

MARIA TERESA MEDEIROS GARCIA and ANDRÉ FERNANDO RODRIGUES ROCHA DA SILVA: Pension expenditure determinants: the case of Portugal

TIHANA ŠKRINJARIĆ: Leading indicators of financial stress in Croatia: a regime switching approach

MARKO CRNOGORAC and SANTIAGO LAGO-PEÑAS: An analysis of COFOG expenditures in former Yugoslavian countries

PEDRO JORGE HOLANDA FIGUEIREDO ALVES, JEVUKS MATHEUS ARAUJO, ANA KAROLINA ACRIS MELO and EDUARDA MASHOSKI: Fiscal decentralization and economic growth: evidence from Brazilian states

Vol. 47, No. 2 | pp. 149-284
June 2023 | Zagreb

ISSN: 2459-8860
https://doi.org/10.3326/pse.47.2

Institute of
Public Finance

Public Sector Economics
2/2023

Vol. 47, No. 2 | pp. 149-284 | June 2023 | Zagreb

TABLE OF CONTENTS

Socioeconomic inequality in the use of long-term care for the elderly in Europe

MAJA MATANIĆ VAUTMANS, Ph.D. student*
MARIJANA OREB, Ph.D. student*
SAŠA DREZGIĆ, Ph.D., Full Professor*

Article**
JEL: I0, H4, R2
https://doi.org/10.3326/pse.47.2.1

* 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 authors are grateful to two anonymous referees who have contributed to the quality of the final version of the paper.

** Received: October 3, 2022
Accepted: January 31, 2023

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

Marijana OREB
University of Rijeka, Faculty of Economics and Business, Ivana Filipovića 4, 51000 Rijeka, Croatia
e-mail: marijanaoreb5@gmail.com
ORCiD: 0000-0002-1987-4387

Saša DREZGIĆ
University of Rijeka, Faculty of Economics and Business, Ivana Filipovića 4, 51000 Rijeka, Croatia
e-mail: sasa.drezgic@efri.hr
ORCiD: 0000-0002-7712-8112

Abstract

The use of formal and informal care for the elderly depends on many factors: income, urban-rural environment, educational attainment, family composition (singles/multi-member family), age and severity of health complications. For this analysis, a pro-rich poverty model is used based on data from the latest (8th) edition of SHARE (Survey of Health, Aging and Retirement in Europe) to examine the impact of socioeconomic inequalities among older people in the use of home care in European Union. The main results indicate that the pro-poor distribution of long-term care prevails in most of the studied countries. At the same time, health variables contribute to pro-poor inequality in the use of long-term care, mainly in informal care. When it comes to formal care, most countries have pro-poor contributions. Formal care inequalities disappeared when adjusted for need factors, while informal care inequalities remained in most countries.

Keywords: formal care, informal care, long-term care, inequity, SHARE data

1 INTRODUCTION

Due to the uncertain future of the social sector because of the increase in the number of the elderly (over 65 years of age) and the late elderly (over 80 years), the reduced birth rate and longer life expectancy, long-term care is a leading topic in national and international policies (European Commission, 2021). Long-term care (LTC) refers to a set of activities that help a person in need to carry out their daily activities over a period of time in accordance with their physical and mental abilities (OECD, 2005). LTC includes all those activities of care provided by another person to a beneficiary in need in formal (professional care) and informal circumstances (family members, friends, volunteers), either in organized long-term care facilities (institutions, assisted living, day centers) or in their own households (home care) for a long period of time during the day (institutions provide 24-hour care) or for several hours during the day (day care centers, home care). Although the implementation of long-term care should be based on three basic parameters proposed by the European Commission (European Commission, 2008): universal (equal) access, high quality and sustainability of long-term care, according to the report of the European Institute for Gender Equality (EIGE, 2020) 52% of European households cannot afford LTC. An especially important role in overcoming the financial barriers faced by users is played by the policies of the national state, that is, the share of public support. The greater the share of public support in the provision of long-term care, the greater the equality in the implementation of LTC services (Carrieri et al., 2017). In addition to personal and income preferences, many other factors influence the choice and use of long-term care, the effect of which has consequences for care costs. Penning et al. (2018) identified the key determinants of long-term care costs: features of social structure (age, gender, urban/rural environment, level of education), social and economic factors (marital status, income) and health status (chronic diseases, functional dependence). These factors are at the same time the most significant factors of the socioeconomic inequality that occurs in the provision of long-term care to the elderly population.

Many authors (Broese Van Groenou and de Boer, 2016; Saito et al., 2018) mention the connection between higher income and smaller household composition groups with the use of professional home care services, while members of smaller socio-economically developed groups are associated with greater commitment to informal care that they provide with greater intensity. Brackley and Penning (2009) state that the use and intensity of informal home care are inversely proportional to income, i.e., wealthier individuals will use informal home care for more hours. This is refuted by Bonsang (2009) who claims that income plays a very important role in choosing a formal long-term care service, and together with a higher level of education, has a positive impact on the number of hours of care needed.

Based on previous research, our research aims to:
1) examine which distribution of LTC (home care) prevails in most of the countries in the European Union (EU-27)
2) determine whether the use of formal care in most countries has pro-rich inequalities
3) establish whether, adjusted for need factors, inequalities in home care in the EU (formal and informal care) will remain at the same level.

The research problem concerns socioeconomic inequality in the use of long-term care, while the research unit comprises the elderly person in need of long-term care. This study seeks to investigate inequality among elderly people in the use of available forms of home care (formal and informal care). The term care refers to the provision of assistance by caregivers in personal care and household chores to people who are dependent on someone else's help (partial or total immobility, mental or physical disabilities), and does not include legal, financial or emotional support. This study will help to consider the current influencing factors, which in different socioeconomic environments affect inequalities in the use of long-term care. In addition, the latest data will provide a clear overview of the current state of the use of formal and informal care in the European Union, which is of crucial importance in planning policies for the care of elderly people. For this analysis, the model of horizontal equality based on income (pro-rich vs. pro-poor) is used, based on data from the last (8th) edition of SHARE (Survey of Health, Ageing and Retirement in Europe) in order to examine the influence of socioeconomic inequalities among the elderly in the use of long-term care. Wagstaff, van Doorslaer and Paci (1991) investigated models for measuring inequalities in health and concluded that the concentration index is the best measure of socioeconomic inequalities in health. The concentration index considers the inequalities of one variable (e.g., health) in relation to another variable (e.g., income, education). The concentration index measures the level of overall health of individuals in different income classes. The main advantage is that it provides a measure of socioeconomic determinants of health inequality, one that takes into account the entire income distribution (not just the richest or the poorest). This methodology has been used in calculating horizontal equity in long-term care (Ilinca et al., 2017; Rodrigues, Ilinca and Schmidt, 2017) with different age groups and when determining horizontal inequality in health care (Bago d'Uva, Jones and van Doorslaer, 2009; Carrieri and Wuebker, 2013; van Doorslaer, Koolman and Jones, 2004). Considering that the prerequisite for the use of formal or informal care is an impaired state of health, the methodology used to calculate horizontal inequality in health care is considered relevant for defining inequality in long-term care as well. Based on research questions the following hypothesis is stated: there is a pro-poor distribution in the use of formal and informal care in European Union among elderly people.

The variables according to which the methodological part will be implemented are urban/rural environment, educational attainment, family composition (singles/multiple family, children), income, sex, age and severity of health complications.

The structure of this study is as follows: the second part presents theoretical and empirical research on socioeconomic differences, as well as the identification of potential criteria that affect unevenly distributed long-term care (income, education, household composition, rural/urban area) among European countries. The third part presents an overview of the selected variables and the model that will be used in the research. The research results are presented in the fourth part. The discussion and conclusion of the research are contained in the fifth part of this study.

2 THEORETICAL AND EMPIRICAL RESEARCH INTO SOCIOECONOMIC DIFFERENCES

Wealthier people have better access to health care and better outcomes in improving their own health by using these services (van Deurzen, van Oorschot and van Ingen, 2014; Rehnberg and Fritzel, 2016). This suggests the choice of using formal care, which was also confirmed by the research of Tenand, Bakx and van Doorslaer (2020), which showed that the tendency to choose between professional care at home or informal care depends on income. Income and education are interrelated variables. Educated groups predominantly use formal care or a combination of formal and informal care because of their income opportunities (Kemper, 1992). Groups with lower socioeconomic status and the least education are most likely to use informal long-term care, even after accounting for differences in health status across socioeconomic groups and other resident characteristics (Bonsang, 2009; Broese van Groenou et al., 2006). The literature finds horizontal inequality in favor of the richer in the use of paid home (professional) care in Southern Europe and horizontal equality in the use of the same services in Northern European countries (Rodrigues et al., 2013; Tenand, Bakx and van Doorslaer, 2020). Numerous organizations have called for the reduction of inequality in the health and care sector. Equal access and equal right to preventive care and the utilization of health services is one of the tasks of the World Health Organization in the "Health for All" agenda (WHO, 1991). Despite the efforts of international organizations, care in most countries of the European Union is determined by income, the most important parameter (Carrieri and Bilger, 2013). Van Doorslaer, Koolman and Jones (2004) found that there are significant disparities in health and long-term care in favor of the richer in all countries, but especially in Portugal, the United Kingdom, and Denmark. Low

PUBLIC SECTOR ECONOMICS
47 (2) 149-176 (2023)

inequalities, on the other hand, were found in the Netherlands, Germany, Italy, Belgium, Spain, Austria, and Ireland. The latest studies indicate an increase in social equality compared to previous years, and the visibility of these inequalities is manifested at the international and national level (Eurofound, 2017). Although equality in the implementation of long-term care is guaranteed by numerous national and international rights, there is still visible heterogeneity in the implementation of these services. Different local development is a potential source of inequality in the use of LTC. On the other hand, the share of informal care in the countries of Southern, Central and Eastern Europe is remarkably similar (Barbieri and Ghibelli, 2018). The reason for this should also be found in established family relationships, where the countries of Southern and Eastern Europe traditionally nurture large families and societal norms that mandate the care of elderly parents. Nordic and Continental European countries represent smaller nuclear families, often dislocated from elderly parents, and according to social understanding, they have no obligation to sacrifice their own career or family to help an elderly father or mother. Better developed environments have better systems and opportunities to use long-term care, as evidenced by the term "postal code lottery" (Rodrigues, Ilinca and Schmidt, 2017), which is the result of infrastructural facilities adopted for long-term care in local municipality and organized domestic assistance programs in the house care. The elderly in more urban areas most often use a formal version of care, because of the greater availability of this form, whether it is the provision of care at home or in an institution (Lera, Pascual-Sáez and Cantarero-Prieto, 2021). The use of different types of services is not the same in all regions, so it can be concluded that better developed environments have developed a greater number of available services for the provision of long-term care (Fernandez and Forder, 2015). Of course, this also depends on the geographical positioning of a local self-government unit and its dependence on the central part of the country. In rural areas, 80% of people with lower incomes and major illnesses (dementia) use informal services (Chuakhamfoo et al., 2020). The reason for this is the lower disposable income of users and limited access to information about the possibilities of using long-term care (Albertini and Pavolini, 2017; Ilinca et al., 2017). Living in a multi-member family provides the opportunity for one of the members to become a provider of informal care and according to Rodrigues et al. (2013) wives and children are the most important source of informal care. Informal care is the most common choice of care among lower income families because their caregivers have lower opportunity costs when losing their job (Sarasa and Billingsley, 2008). Also, family members often act as advocates when choosing long-term care services (Rodrigues, Ilinca and Schmidt, 2017) especially in rural areas and in households with lower incomes that are initially disadvantaged due to limited access to information and types of services.

Depending on the national plan and program, many countries in Europe subsidize the resort to home care as an adequate substitution for the institutional type of accommodation. The Nordic countries have successfully implemented the deinstitutionalization process and emphasized the importance of using home care. On the other hand, in many European countries, the informal form of care takes precedence over the formal due to the insufficient development of the system.

3 DATA AND METHODS

In this study we are exploring socioeconomic inequalities in utilization of formal and informal long-term care (LTC) in the elderly population (over 65 years). To this end we employed the concentration index (CI), a synthetic measure of inequality in healthcare related to socioeconomic status (SES) (Wagstaff, van Doorslaer and Paci, 1991; Wagstaff and van Doorslaer, 2000) which measures income inequality in the use of health services and is written as follows:

$$CI = \frac{2}{\mu} cov(h_i, r_i) \tag{1}$$

where h_i is the health care utilization variable, μ is the average health care utilization and r_i is the individual belonging to the socioeconomic group. This measure takes values between -1 and 1, negative values implying that LTC use is disproportionately distributed among the poor and positive values that it is mostly distributed among the rich. Variables for formal and informal care utilization are binary, so a corrected concentration index (CCI) was used (Erreygers, 2009):

$$CCI = 4 * \mu * CI = 8cov(h_i, r_i) \tag{2}$$

In order to obtain more informative results, a decomposition of CI into contributing factors was implemented. Each one of these factors contributes to the overall CI, with some factors contributing to the pro-rich direction, while others pull the CI in the opposite direction, i.e., pro-poor. The contributions of individual factors are modeled by considering both the effect that the given factor has on the level of utilization, and also the distribution of the said factor in relation to SES. The former is called elasticity, measuring how sensitive the LTC is to variation in the said factor. The latter is called a partial concentration index, measuring how equal the distribution of the said factor is with respect to SES. The factor contributions are proportional to the product of elasticity and partial CI, meaning that factors with large elasticity but very small partial CI will have small contributions, as well as factors with very small elasticity and large partial CI, so only factors in which both elasticity and partial CI are pronounced will significantly contribute to overall CI. Also, the signs of elasticity and partial CI will determine the sign of the contribution, with same signs (++ or – –) giving a positive, i.e., pro-rich contribution, and different signs (+– or –+) producing a negative (pro-poor) contribution.

Another question not answered by previously described analyses is: to what extent is LTC driven by dependency and need for care? To this end we employ a method called analysis of equity, in which horizontal inequity index (HI) is a measure of inequality. Empirical research on equity in health care utilization examines horizontal equity, defined as equal treatment for equal need, independent of characteristics such as SES, for which differences in health care are considered unacceptable. In practice, treatment is measured by health care utilization and need by health indicators and demographic characteristics. Horizontal inequity is determined by comparing the deviation of the actual distribution of health care from that which would result if utilization were determined by need alone. In other words, equity analysis is based on the principle that individuals with the same level of need for care should receive the same level of care regardless of other potential factors, so positive or negative values signal unequal treatment of equal needs. We calculate the HI using the indirect standardization method (van de Poel, van Doorslaer and O'Donnell, 2012), where HI is obtained as the difference between overall CI and the contributions of all the need factors used in the study, leaving the contributions of non-need factors.

The analyses were performed using R: A Language and Environment for Statistical Computing together with a set of open-source R packages. CI calculation and decomposition was performed with the help of rineq package[1] which is based on decomp package[2].

We used specific methodological documentation based on Bergmann and Börsch-Supan (2021). The analysis is based on microdata from the eighth wave of the Survey of Health, Ageing and Retirement in Europe (SHARE), collected during 2019/2020 in 26 European countries. Israel was part of the SHARE survey but was not included in this sample (not an EU member). The sample consists of people aged 65 or older at the time of the survey. Where possible, imputed values prepared by SHARE were used for variables with many missing values. Of the initial 46,733 subjects, due to these restrictions and missing values, 31,340 individuals were left in the study. Of these 15,393 excluded individuals, 12,327 were younger than 65 years, the remaining 824 were individuals from Israel, and 2,241 were due to missing values. Most missing values were recorded for the variable Area (1505), with the Czech Republic having the most missing values in absolute numbers (174 or 7.7%) and Slovakia having the most missing values in relative numbers (48 or 10.6%). The remaining 736 missing values were recorded for the Formal Care variable (520, most for Belgium: 56 or 3.9%), Informal Care (156, most for Italy: 21 or 1.3%), and the three Health variables, all of which had fewer than 10 individuals with missing values. There were also some negative values recorded in the responses for the health variables that were excluded, 50 in total.
Formal LTC use was measured by a synthetic indicator of use in the last 12 months, which captures professional support including personal care, domestic tasks, other activities and meals-on-wheels. Informal LTC use was measured by the synthetic use indicator in the last 12 months, which includes non-professional support from outside the household as well as inside. Socioeconomic status (SES) was proxied by equivalized net household income, using square root scale and adjusted for purchasing power parity of each country[3], according to the formula:

$$\text{equivalized income} = \text{household income} / \text{sqrt}(\#\text{ people in household}) * \text{exchange rate} / \text{ppp rate} \quad (3)$$

[1] Available at: https://github.com/brechtdv/rineq.
[2] Niko Speybroeck Decomposing socioeconomic health inequalities, Int J Public Health (2010) 55:347-351.
[3] Available at: https://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf.

PUBLIC SECTOR ECONOMICS
47 (2) 149-176 (2023)

4 RESULTS

Descriptive statistics of all dependent and independent variables with sample sizes are presented in table 1, broken down for each analyzed country.

Table 1

Descriptive statistic of all dependent and independent variables

	N	Formal care = Yes (%)	Informal care = Yes (%)	SP health (mean (SD))	ADL (mean (SD))	Chronic (mean (SD))	Area = Rural (%)
Austria	1,153	183 (15.9)	449 (38.9)	3.15 (1.02)	0.28 (0.91)	2.07 (1.66)	457 (39.6)
Germany	1,971	301 (15.3)	582 (29.5)	3.31 (0.94)	0.36 (1.07)	2.32 (1.74)	709 (36.0)
Sweden	1,914	196 (10.2)	464 (24.2)	2.88 (1.08)	0.17 (0.67)	1.81 (1.46)	319 (16.7)
Netherlands	1,482	247 (16.7)	355 (24.0)	2.96 (1.03)	0.16 (0.59)	1.60 (1.39)	293 (19.8)
Spain	1,614	301 (18.6)	310 (19.2)	3.43 (0.96)	0.54 (1.43)	2.35 (1.73)	370 (22.9)
Italy	1,457	163 (11.2)	273 (18.7)	3.46 (0.93)	0.33 (1.04)	1.90 (1.50)	511 (35.1)
France	1,790	299 (16.7)	509 (28.4)	3.28 (0.96)	0.29 (0.83)	2.13 (1.62)	870 (48.6)
Denmark	1,442	180 (12.5)	501 (34.7)	2.68 (1.11)	0.17 (0.63)	1.86 (1.50)	337 (23.4)
Greece	2,040	150 (7.4)	491 (24.1)	3.29 (0.95)	0.21 (0.86)	2.36 (1.61)	365 (17.9)
Switzerland	1,396	171 (12.2)	365 (26.1)	2.81 (0.93)	0.11 (0.45)	1.53 (1.41)	752 (53.9)
Belgium	1,319	370 (28.1)	381 (28.9)	3.09 (0.92)	0.32 (0.85)	2.19 (1.60)	255 (19.3)
Czech Republic	2,035	184 (9.0)	827 (40.6)	3.20 (0.82)	0.32 (0.96)	2.41 (1.68)	583 (28.6)
Poland	1,282	46 (3.6)	239 (18.6)	3.72 (0.83)	0.43 (1.18)	2.83 (2.00)	636 (49.6)
Luxembourg	581	70 (12.0)	109 (18.8)	3.25 (0.93)	0.22 (0.81)	2.64 (2.10)	262 (45.1)
Hungary	574	55 (9.6)	145 (25.3)	3.54 (0.94)	0.31 (0.92)	2.27 (1.56)	190 (33.1)
Slovenia	1,785	95 (5.3)	431 (24.1)	3.35 (0.96)	0.36 (1.15)	2.13 (1.61)	685 (38.4)
Estonia	2,121	107 (5.0)	741 (34.9)	3.96 (0.76)	0.40 (1.05)	2.24 (1.72)	705 (33.2)
Croatia	725	44 (6.1)	214 (29.5)	3.58 (0.98)	0.33 (1.02)	2.42 (1.67)	210 (29.0)
Lithuania	837	40 (4.8)	179 (21.4)	3.78 (0.75)	0.51 (1.35)	2.43 (1.78)	252 (30.1)
Bulgaria	586	33 (5.6)	168 (28.7)	3.55 (0.97)	0.36 (1.05)	2.11 (1.44)	290 (49.5)

	N	Formal care = Yes (%)	Informal care = Yes (%)	SP health (mean (SD))	ADL (mean (SD))	Chronic (mean (SD))	Area = Rural (%)
Cyprus	404	97 (24.0)	88 (21.8)	3.37 (1.05)	0.44 (1.31)	2.51 (1.73)	93 (23.0)
Finland	719	61 (8.5)	217 (30.2)	3.31 (0.90)	0.18 (0.64)	2.34 (1.64)	344 (47.8)
Latvia	486	16 (3.3)	101 (20.8)	4.11 (0.70)	0.22 (0.76)	2.04 (1.33)	151 (31.1)
Malta	494	29 (5.9)	53 (10.7)	3.33 (0.86)	0.17 (0.90)	1.68 (1.28)	120 (24.3)
Romania	735	19 (2.6)	214 (29.1)	3.79 (0.98)	0.49 (1.36)	1.94 (1.49)	551 (75.0)
Slovakia	398	30 (7.5)	72 (18.1)	3.20 (0.96)	0.26 (0.84)	1.59 (1.57)	229 (57.5)

	Age (mean (SD))	Gender = Female (%)	Married = Married (%)	Children (mean (SD))	Education (mean (SD))	Income (mean (SD))
Austria	75.90 (7.11)	695 (60.3)	701 (60.8)	2.21 (1.40)	9.26 (4.87)	2,220 (2,566)
Germany	74.82 (6.86)	1,004 (50.9)	1,424 (72.2)	1.99 (1.22)	12.89 (3.62)	2,062 (883)
Sweden	76.07 (7.02)	1,018 (53.2)	1,329 (69.4)	2.28 (1.24)	12.02 (3.98)	1,708 (800)
Netherlands	74.29 (6.64)	764 (51.6)	1,167 (78.7)	2.30 (1.32)	12.04 (3.83)	1,896 (803)
Spain	77.05 (8.11)	896 (55.5)	1,157 (71.7)	2.36 (1.48)	8.29 (5.00)	1,441 (1456)
Italy	75.60 (7.06)	779 (53.5)	1,121 (76.9)	2.08 (1.26)	8.24 (4.36)	1,358 (917)
France	75.30 (7.72)	1,021 (57.0)	1,094 (61.1)	2.22 (1.32)	11.97 (3.84)	2,118 (1700)
Denmark	74.58 (6.95)	771 (53.5)	980 (68.0)	2.29 (1.21)	13.35 (3.45)	2,579 (3493)
Greece	75.17 (7.28)	1079 (52.9)	1,460 (71.6)	1.91 (1.05)	9.01 (4.26)	1,723 (2634)
Switzerland	75.22 (7.43)	749 (53.7)	927 (66.4)	2.13 (1.36)	8.83 (5.41)	3,837 (6361)
Belgium	74.78 (7.44)	710 (53.8)	862 (65.4)	2.12 (1.30)	12.53 (3.72)	1,798 (717)
Czech Republic	74.69 (6.44)	1,200 (59.0)	1286 (63.2)	2.21 (0.99)	12.36 (3.18)	1,138 (881)
Poland	73.81 (7.24)	690 (53.8)	910 (71.0)	2.45 (1.46)	10.18 (3.32)	1,130 (1042)
Luxembourg	73.49 (6.64)	301 (51.8)	436 (75.0)	1.90 (1.24)	11.95 (4.42)	3,548 (4067)
Hungary	73.60 (6.47)	345 (60.1)	335 (58.4)	1.79 (0.99)	10.86 (2.83)	835 (396)

	Age (mean (SD))	Gender = Female (%)	Married = Married (%)	Children (mean (SD))	Education (mean (SD))	Income (mean (SD))
Slovenia	75.10 (7.35)	1,027 (57.5)	1,255 (70.3)	1.99 (0.95)	10.56 (3.29)	1,779 (2469)
Estonia	76.39 (7.35)	1,358 (64.0)	1,142 (53.8)	1.95 (1.21)	11.76 (3.48)	944 (762)
Croatia	73.97 (6.59)	386 (53.2)	535 (73.8)	1.86 (0.98)	10.21 (3.73)	1,062 (1219)
Lithuania	76.03 (7.44)	535 (63.9)	427 (51.0)	2.09 (1.25)	10.94 (4.09)	1,083 (1,420)
Bulgaria	74.34 (6.68)	356 (60.8)	332 (56.7)	1.95 (0.82)	10.23 (3.35)	653 (748)
Cyprus	76.92 (7.33)	241 (59.7)	286 (70.8)	2.55 (1.29)	8.58 (4.37)	6,565 (7901)
Finland	74.67 (6.61)	381 (53.0)	512 (71.2)	2.22 (1.57)	11.18 (3.68)	2,986 (5077)
Latvia	75.28 (6.83)	321 (66.0)	246 (50.6)	1.74 (1.17)	11.51 (3.24)	1,098 (1,689)
Malta	73.55 (6.33)	260 (52.6)	402 (81.4)	2.43 (1.39)	9.23 (3.69)	1,561 (2403)
Romania	73.74 (6.99)	409 (55.6)	479 (65.2)	2.30 (1.51)	8.93 (3.69)	686 (375)
Slovakia	71.68 (5.79)	209 (52.5)	283 (71.1)	1.98 (1.11)	11.60 (2.08)	1,842 (2890)

Dependent variables are formal and informal care use. Formal care use ranges in proportion from the lowest, 2.6% in Romania, to 28.1% in Belgium, while informal care use is lowest in Malta 10.7% and highest in the Czech Republic 40.6%. High use of formal care was also recorded in Cyprus and Spain, followed by France, the Netherlands, Austria, and Germany (15.3%). Interestingly, it was lower in countries such as Denmark, Switzerland and Finland. Formal care was lowest in Eastern European, Southeastern European and Baltic countries. Informal care was very high in Austria (38.9%), Estonia and Denmark, followed by Finland, Croatia, Germany, Romania, Belgium, Bulgaria and France (28.4%). It was below 20% in Malta (10.7%), Slovakia, Poland, Italy, Luxembourg and Spain (19.2%). The remaining countries had scores between 20.8% and 26.1%.

Independent variables (factors) are separated into need factors: SP health, ADL, Chronic, Age, Gender; and non-need factors: Area, Married, Children, Income, Education. SP (self-perceived) health is a variable scored on a scale from 1 – Excellent to 5 – Poor, and because it assumes 5 different values with an approximate symmetrical distribution, it is used in the analyses as a continuous variable with numeric values. Poorest scores were recorded for Latvia (4.11), Estonia, Lithuania, Romania, Poland, Croatia, Bulgaria and Hungary (3.54), while the best scores were seen in Denmark (2.68), Switzerland, Sweden and Netherlands (2.96). All other countries had scores between 3 and 3.5.
Most of the elderly in the analyzed countries declare they are of average health status, with the best self-perceived score in Denmark and the worst in Latvia. ADL

(activities of daily living limitations) is also treated as a continuous variable, along with Chronic (number of chronic diseases). The worst ADL scores were found for Spain (0.54), Lithuania, Romania, Cyprus, Poland, and Estonia (0.40); the best scores were found for Switzerland (0.11), the Netherlands, Malta, Denmark, Sweden, and Finland (0.18); and the remaining countries ranged from 0.21 to 0.36. The most chronic diseases were reported for Poland (2.83), Luxembourg, Cyprus, Lithuania, Croatia, Czech Republic, Greece, Spain, and Finland (2.34), the least for Switzerland (1.53), Slovakia, the Netherlands, Malta, Sweden, Denmark, Italy, and Romania (1.94), while the remaining countries ranged from 2.04 to 2.32.

Age is a continuous variable measuring age of respondents, and the mean values were in the range from 71.7 for Slovakia to 77.1 for Spain. Gender is a dichotomous categorical variable, indicating if an individual is of a female gender. There were the fewest female respondents in Germany (50.9%) and the most in Latvia (66%). The highest average age is detected in Spain and regarding gender, all EU countries have dominant share of women in the population of elderly people. Area is a dichotomous categorical variable with a positive category defining rural area, where an original variable with five categories was grouped into two categories, with first category encompassing all city and town areas, and second category defining only rural areas. The countries with the fewest rural areas were Sweden (16.7%), Greece, Belgium, the Netherlands, Spain, Cyprus, Denmark, and Malta (24.3%), while the countries with the most rural areas were Romania (75%), Slovakia, Switzerland, Poland, Bulgaria, France, Finland, and Luxembourg (45.1%). The variable "Married" is a dichotomous categorical variable with the first category including original categories "Never married", "Divorced", "Widowed", and second including "Married, living with spouse", "Registered partnership", "Married, not living with spouse". The fewest respondents in the married category were in Latvia (50.6%) and the most were recorded in Netherlands (78.7%).

Children is a continuous variable measuring number of children, with the lowest average number of children recorded for Hungary (1.79) and the highest for Cyprus (2.55). Income is a continuous measure of SES. The countries with the lowest income of respondents were Bulgaria (653), Romania, Hungary, Estonia, Croatia, Lithuania and Latvia (1,098), while the highest incomes were recorded for Cyprus (6,565), Switzerland (3,837), Luxembourg, Finland, Denmark, Austria, France and Germany (2,062). It is interesting to note that Cyprus has the highest income, while Switzerland, in second place, has a much lower mean. The income distributions are very skewed, with most values at the lower end of the range. Comparing the income distributions for these two countries, we see that Cyprus has many more values at the high end of the range than Switzerland, implying that there are more wealthy individuals among respondents in Cyprus. Looking at another measure of central tendency, the median, which marks the point at which 50% of respondents are below and 50% are above this point, both countries had almost the same median value around the year 2000. The median income value for Cyprus was so high because many more respondents had high

PUBLIC SECTOR
ECONOMICS
47 (2) 149-176 (2023)

and very high incomes. Also, this variable is equivalized income, so the original income was corrected for PPP. The PPP ratio between Switzerland and Cyprus was about 2, so this variable was also inflated for Cyprus. The likely reason for this phenomenon is that Cyprus is a popular immigration destination for wealthy retirees from other countries, especially from the United Kingdom, so a large proportion of the respondents consists of these immigrants, whose incomes are much higher than those of native-born Cypriots.

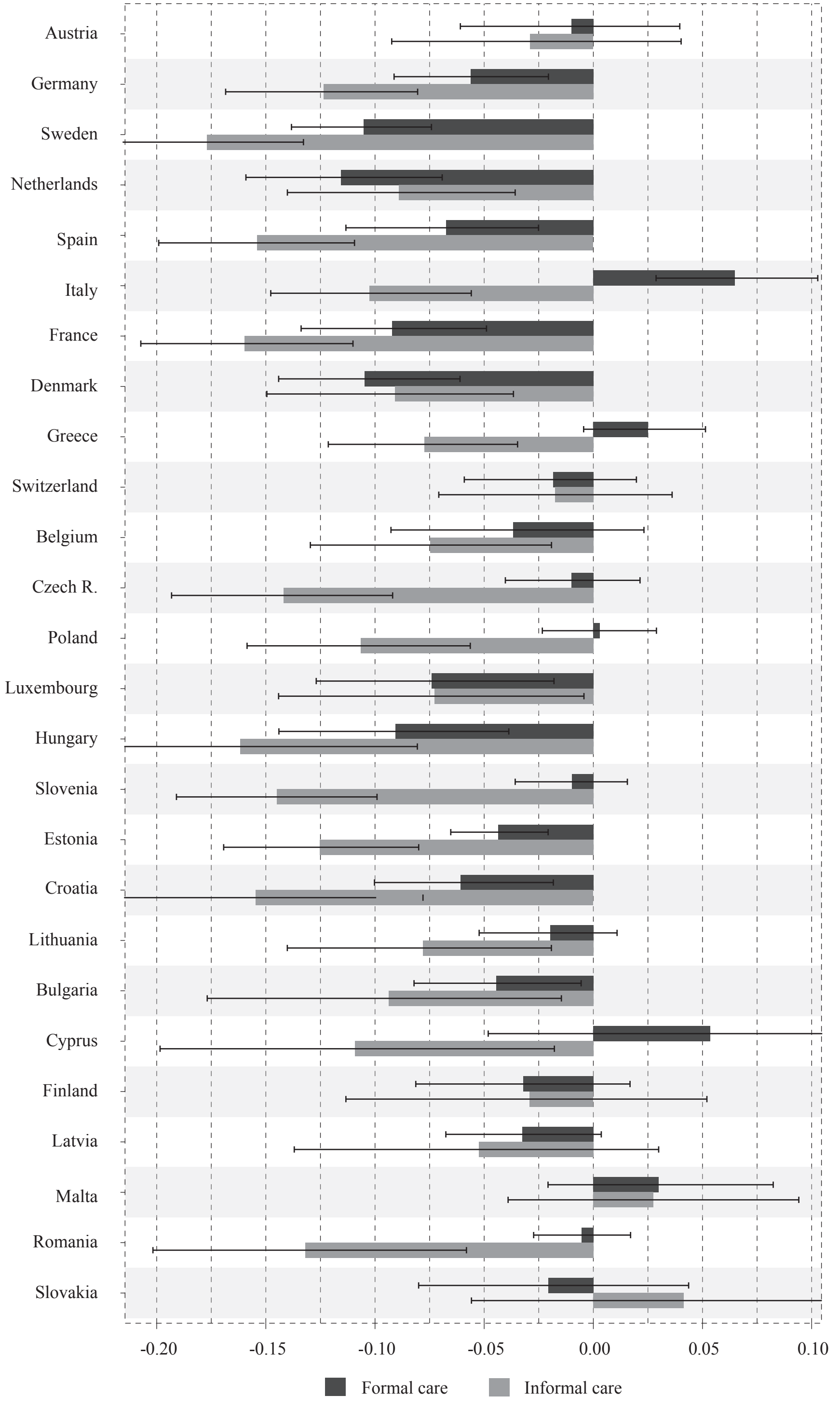
Education measures the overall number of years of the individual's formal education. The countries with the lowest number of years of education were Italy (8.24), Spain, Cyprus, Romania, Greece, Austria, and Malta (9.26); the highest values were recorded for Denmark (13.35), Germany, Belgium, the Czech Republic, the Netherlands, and Sweden (12.02).

It is interesting to compare some countries based on these values. For example, seemingly similar countries such as Spain and Italy show significant differences in the formal use of care, while the informal use of care is comparable. This may be related to differences in health scores, with ADL and chronic disease scores in particular being significantly worse in Spain, leading to a higher need for formal care. Comparing neighbors Belgium and the Netherlands, the difference in formal care use is even greater. However, health scores here also suggest one of the reasons for these differences. Of course, other important reasons may also influence these differences, such as government spending on long-term care and other sociocultural characteristics. According to the recorded data, diversity is observed in use when choosing care (formal vs. informal). In addition, cultural norms (family-oriented societies) and other parameters (age, income, education, life in rural areas) are also different, which indicates the necessity of creating a system in accordance with dissimilarity at the national level.

4.1 ANALYSIS OF INEQUALITY

Figure 1 shows the CI values for each country accompanied by bootstrap 95% confidence intervals. The measures whose intervals do not cross zero are significant at 0.05 significance level. CI for most countries is negative, indicating a pro-poor distribution of LTC.

FIGURE 1
Concentration index values for each country



PUBLIC SECTOR ECONOMICS 47 (2) 149-176 (2023)

The only statistically significant positive CI is recorded for formal care in Italy, where informal care CI is also significantly negative. The only other positive formal care CIs are for Greece, Malta and Cyprus, all not significantly different from zero.

Informal care CIs are almost uniformly smaller than corresponding formal care CIs, making informal care even more pro-poor than formal care in most countries. Austria, Switzerland, Finland, Latvia, Malta and Slovakia have neither measure significantly different from zero. Greece, Belgium, the Czech Republic, Poland, Slovenia, Lithuania, Cyprus and Romania have only informal care CI significantly pro-poor. All other countries have significantly negative measures, with the smallest formal care CIs being in Netherlands, Sweden, Denmark, France and Hungary, and smallest informal care CIs in Sweden, France, Spain, Croatia, the Czech Republic and Romania. From these results, it appears that in most countries both formal and informal care have an unfavorable distribution, i.e., long-term care tends to be used by the poorer strata of the population. However, informal care is even more disproportionately used by the poor in most cases, which may mean that they do not have adequate access to formal care and need to find other informal means of care.

An important consideration in the CI analysis and also in subsequent analyzes is the issue of sample size. To obtain reliable inferences, each country should have sufficient data for the calculations. Since in this study the sample sizes per country are determined a priori by the SHARE sample and could not be increased by additional data collection, we consider here only the sample sizes post-hoc. The basic formula for sample size is:

$$n = z^2\left(p\left(1-p\right)\right)/d^2 \tag{4}$$

where $z = 1.96$ (upper quantile of the normal distribution at a 95% significance level), p is the level of formal/informal care in a given country, and d is the precision of the CI, i.e., the +– edge of the corresponding confidence interval. Setting d to a specific value means that estimates of the CI whose absolute value is greater than d are considered significantly different from zero. With $d = 0.03$ for the CI of formal care, only Cyprus had too small a sample size because an additional 375 subjects would have been required to achieve the desired precision. Applying the same precision value to informal care needs, more countries had sample sizes that were too small: Cyprus (323), Bulgaria (287), Slovakia (234), Hungary (232), Latvia (217), Finland (180), Croatia (163), Romania (146), and Luxembourg (70). The larger sample size needed for informal care is because the proportion of formal care was smaller than that of informal care, especially in the group of countries with smaller sample sizes. This means that the results should be interpreted with caution, especially regarding informal care in the countries mentioned. Figure 1 shows bootstrapped confidence intervals instead of asymptotic intervals, which gives the results more reliability.

4.1.1 FORMAL AND INFORMAL CARE CONCENTRATION INDEX DECOMPOSITION

Contributions of factors in concentration index decomposition are presented in figures 2 and 3, and figures 4 and 5 show plots of corresponding elasticity and partial CIs. Negative contribution values indicate the pro-poor contribution of the relevant factor to the overall CI, while positive values present a pro-rich contribution. Figures 4 and 5 have areas shaded with different colors, white marking an area with pro-rich contributions (same sign) and grey areas marking pro-poor contributions (different sign).

Health variables (SP health, ADL, Chronic) uniformly across countries contribute to pro-poor inequality in use of LTC, with larger relative contributions recorded for informal care. Regarding formal care, SP health shows the greatest contributions among the health measures, with countries Sweden, France, Denmark, Germany, Hungary, Slovenia, Estonia and Slovakia having the largest contributions. Informal care shows a similar pattern, with countries Hungary, Spain, Sweden, Luxembourg, Slovakia, Denmark, Netherlands, France and Poland showing greatest pro-poor contribution.

For informal care, ADL is an important contributor among health measures. It can be seen on figures 4 and 5 that health variables, excluding minor negative values for a few countries, have positive elasticities, meaning that larger values for these factors predict greater utilization of LTC, while partial CIs are in general pro-poor, so people with greater health needs are disproportionately more represented among poor population. Therefore, products of elasticity and CI result in pro-poor contributions. Overall, for formal care, health variables dominate the contributions of other non-health variables in: Germany, France, Netherlands, Hungary and Slovenia. With informal care, other non-health factors add greater contributions, with Spain, Netherlands, Sweden, France, Denmark, Hungary, Poland, Luxembourg having the largest relative contributions of health variables.

Besides health variables, other need factors are also Age and Gender. Age greatly contributes to the pro-poor inequality in several countries, both for formal and informal care: Sweden, Denmark, Spain, the Czech Republic, Netherlands, Belgium, Lithuania, Switzerland and Hungary. Estonia and Greece have larger contributions with informal care. Elasticity and partial CI plots show that elasticity is uniformly positive for age, which is expected, as older people will utilize more LTC. Also, partial CIs are negative as older individuals are more represented in the poor population. Sweden has the highest contribution of age, which is due to having the most pro-poor distribution of age and one of the highest elasticities for formal care, and relatively high elasticity for informal care. This could be explained by strong state support of elderly people who, although having markedly lower SES, receive high levels of LTC. Other countries with older age contribution have a similar pattern of elasticity and partial CI.

Variable Gender did not much contribute to the overall CI. Females in most countries are disproportionately represented in the poorer population, while in most countries females utilize more LTC, with only some countries having men use more LTC (Slovenia, Poland, Bulgaria and Slovakia for formal care; Bulgaria, Romania and Croatia for informal care). Relatively largest contributions of gender are recorded in Lithuania, Finland and Slovenia for formal care, and the Czech Republic, Hungary and Switzerland for informal care.

Figure 2

Formal care concentration index decomposition

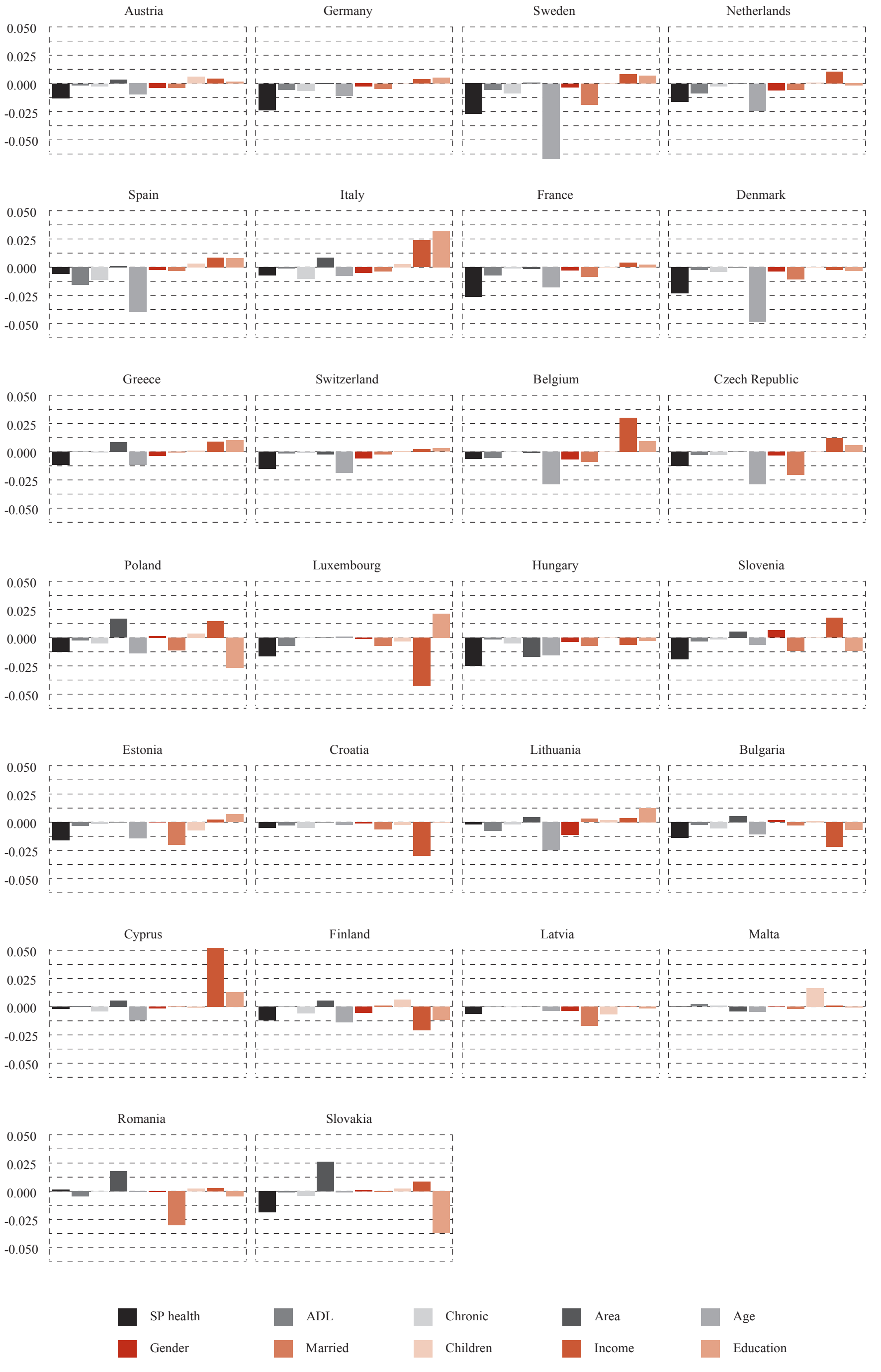
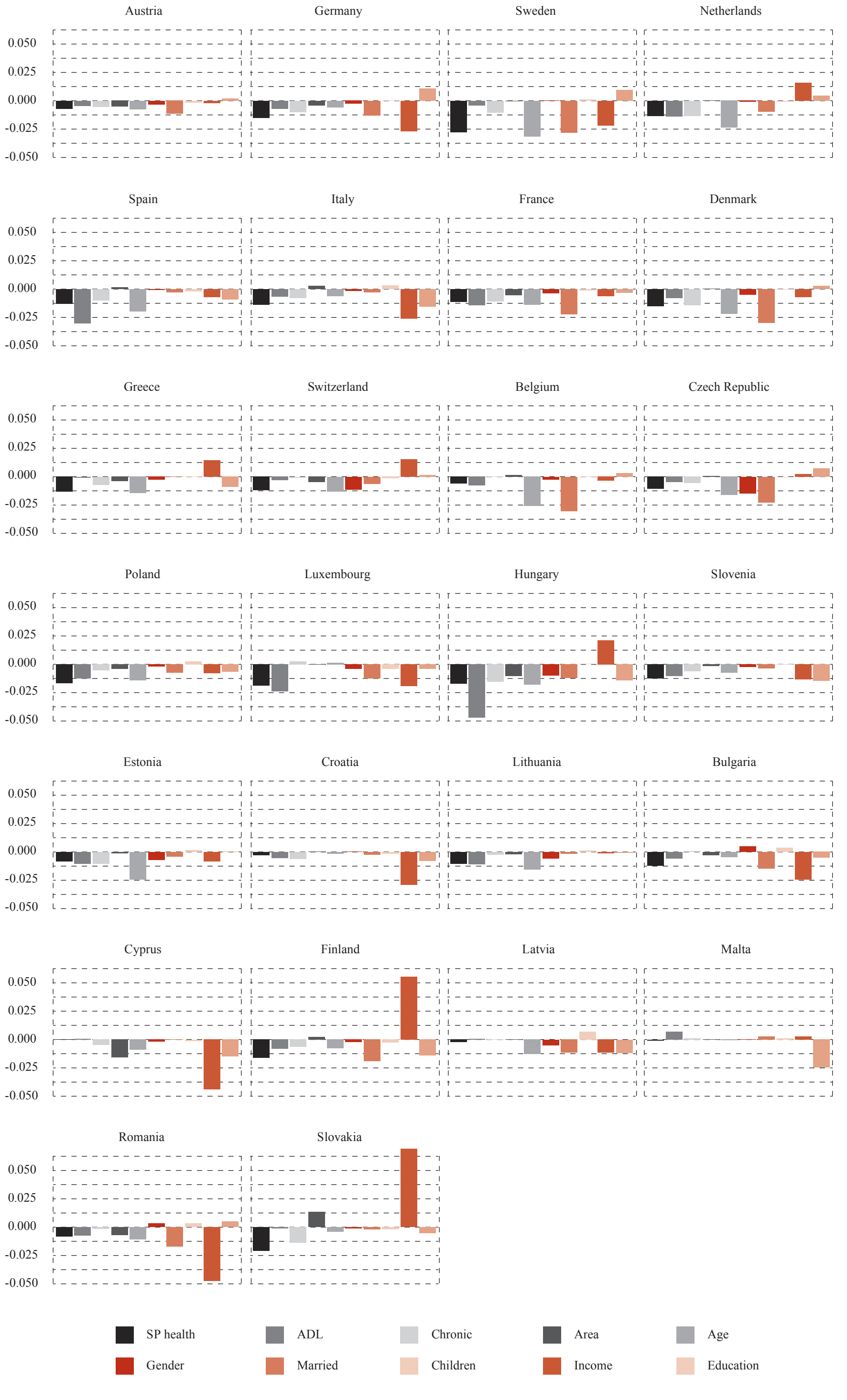


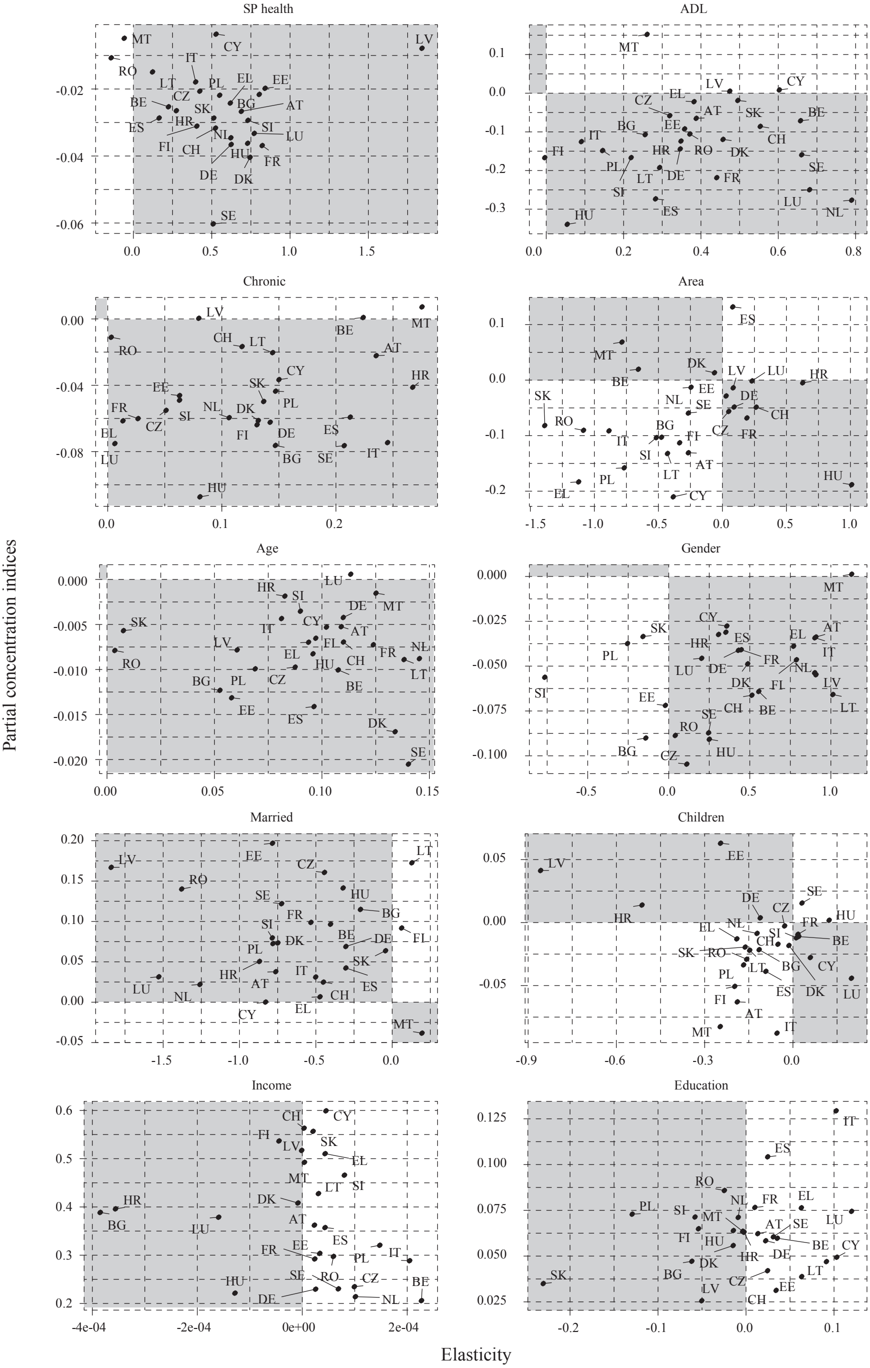
Figure 3
Informal care concentration index decomposition



PUBLIC SECTOR ECONOMICS
47 (2) 149-176 (2023)

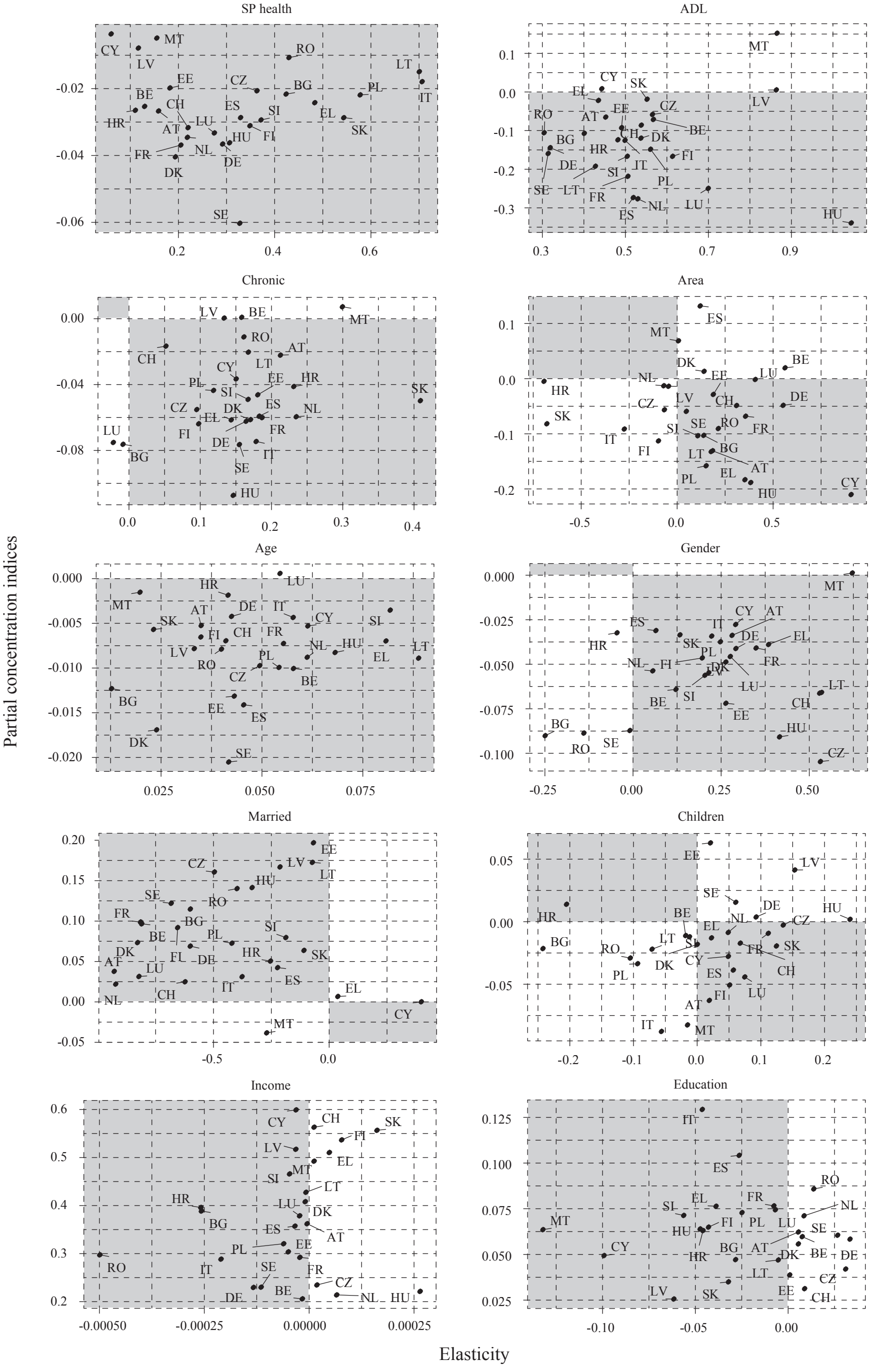
Figure 4

Formal care concentration index decomposition



Note: Belgium – BE, Bulgaria – BG, Denmark – DK, Germany – DE, Estonia – EE, Ireland – IE, Greece – EL, Spain – ES, France – FR, Croatia – HR, Italy – IT, Cyprus – CY, Latvia – LV, Lithuania – LT, Luxembourg – LU, Hungary – HU, Malta – MT, Netherlands – NL, Austria – AT, Poland – PL, Portugal – PT, Romania – RO, Slovenia – SI, Slovakia – SK, Finland – FI, Sweden – SE, Switzerland – CH, Czech Republic – CZ.

Figure 5
Informal care concentration index decomposition



Note: Belgium – BE, Bulgaria – BG, Denmark – DK, Germany – DE, Estonia – EE, Ireland – IE, Greece – EL, Spain – ES, France – FR, Croatia – HR, Italy – IT, Cyprus – CY, Latvia – LV, Lithuania – LT, Luxembourg – LU, Hungary – HU, Malta – MT, Netherlands – NL, Austria – AT, Poland – PL, Portugal – PT, Romania – RO, Slovenia – SI, Slovakia – SK, Finland – FI, Sweden – SE, Switzerland – CH, Czech Republic – CZ.

As for non-need factors, for most countries the greatest contributions were recorded for Income, followed by Education, Married and Area. Income has by definition a positive partial CI (higher income is more concentrated among richer individuals), so the sign of contribution depends on the sign of elasticity.

Regarding formal care, most countries have positive elasticity, giving pro-poor contributions. Largest positive contributions were recorded for Cyprus (although small sample size makes these values less reliable), Belgium, Italy, Slovenia and Poland. This is of course due to positive associations between income levels and formal care use (positive elasticity). Countries that displayed negative contributions of income were Luxembourg, Croatia, Bulgaria, Finland and Hungary. For these countries, higher income is associated with lower levels of formal care use, which could be because lower income individuals are favored in state support, while higher income individuals find other means of meeting their needs. With informal care income contributions, for the most countries the trend is more negative, with many more countries showing a pronounced pro-poor contribution, like Romania, Cyprus, Croatia, Germany, Italy, Bulgaria, Sweden, Luxembourg, Slovenia and Latvia. For these countries informal care is more utilized among the poor, driving pro-poor inequality. Of the countries with pro-rich contributions, most pronounced are Slovakia (also small sample size so less reliable), Finland, Hungary, Netherlands, Switzerland and Greece, as here there is higher utilization of informal care by richer individuals.

Education is similar to income in terms of elasticity and partial CI, with all countries having a pro-rich partial CI, and elasticity defining the sign of contribution. Education shows a smaller contribution than Income for most countries, with largest pro-rich contributions in formal care recorded for Italy, Luxembourg, Lithuania, Cyprus and Belgium, while Slovakia, Poland, Finland and Slovenia showed a pro-poor contribution. Interestingly, Slovakia and Poland have opposite signs of income and education contributions, which could be because here the Area factor is more pronounced, explaining some variation for the rural, less educated population, while the other part is explained by income which correlates with education, so education elasticity has a negative sign. Regarding informal care, education elasticity for most countries is negative, giving pronounced pro-rich contributions only for Germany and Sweden, while most other countries have pro-poor contributions, although relatively unimportant compared to other factors, except for Malta, Slovenia and Italy.

Area was not an important contributor for formal care in most countries, except to some extent in Slovakia, Romania, Poland, Hungary and Greece. Most contributions were pro-rich, as most countries have poorer rural areas and less utilization of LTC in such areas, with Hungary being one of the exceptions with greater utilization recorded for rural areas. Informal care showed somewhat similar pattern, with both positive and negative contributions. Relatively largest contributions were recorded for Cyprus, Slovakia and Hungary.

PUBLIC SECTOR
ECONOMICS
47 (2) 149-176 (2023)

Last group of factors are Married and Children. For most countries, including both formal and informal care, these two factors were not important contributors to inequality, with Children being mostly negligible, except for Malta (formal care). All contributions of being married were pro-poor, owing to partial CI being pro-rich (married people are disproportionately represented among richer population) and elasticity predominantly negative, both for formal and informal care, meaning that married people use less LTC. Largest contributions in formal care were recorded for Romania (where other contributions dominate), the Czech Republic, Sweden, Estonia and Latvia, while informal care contributions were the largest for Sweden, Denmark, Belgium, Czech Republic, France, Finland and Romania. Informal care contributions were more pronounced than formal, and in more countries, implying that being married results in less use of informal care.

4.2 INEQUITY ANALYSIS

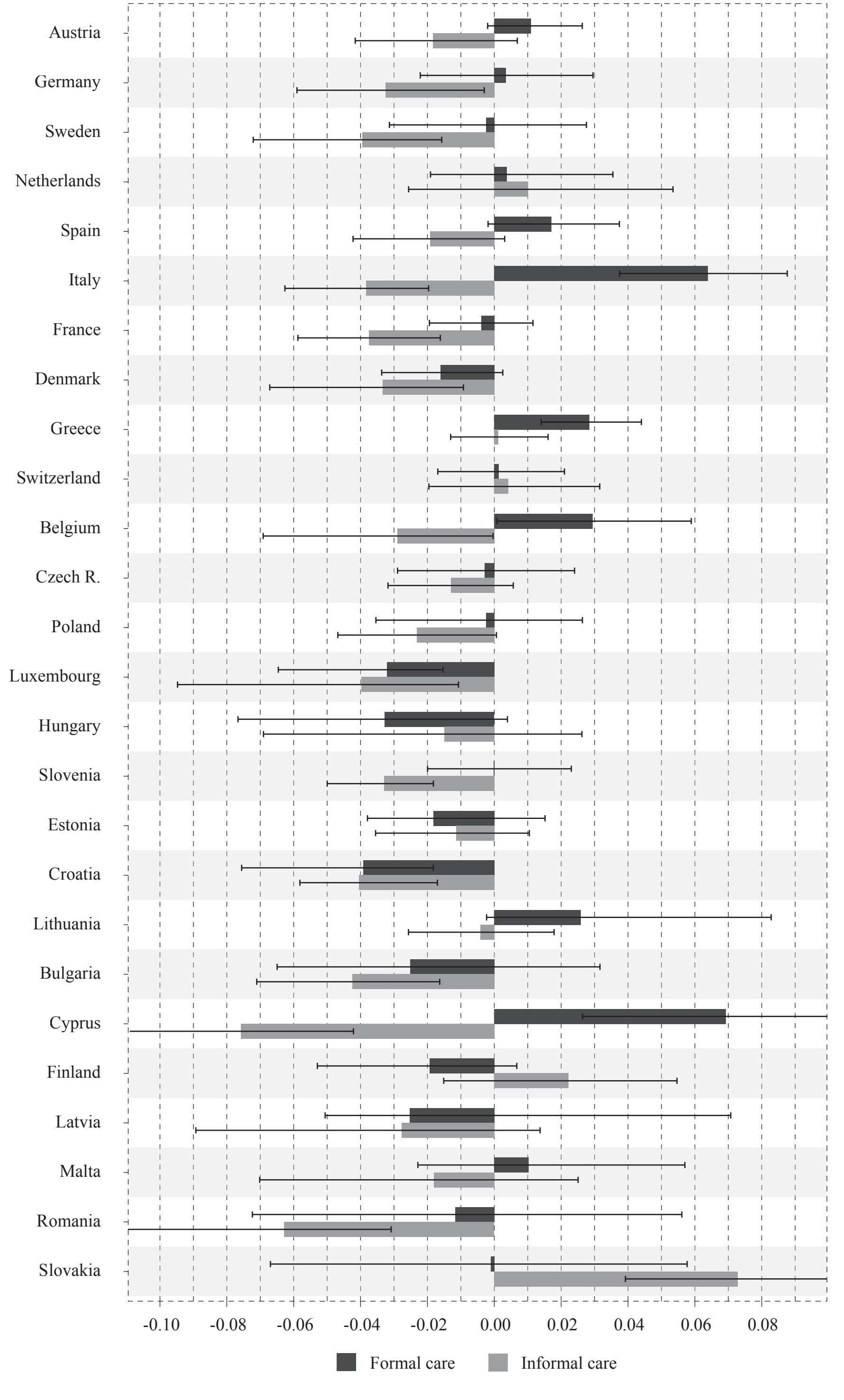
Previous results of decomposition showed that need factors greatly contributed to the pro-poor inequality for most countries. These factors can be defined as legitimate sources of inequality in LTC utilization, so with an analysis of equity we measure inequality after accounting for need factors. Figure 6 shows horizontal inequity indices for all countries with corresponding bootstrapped 95% confidence intervals.

Figure 6 shows that for most countries HI is more positive than CI, and many previously significant pro-poor concentration indices are now not significantly different from zero. A significantly pro-rich HI was recorded for Italy, Greece, Belgium and Cyprus for formal care, and Slovakia for informal care, and a significantly pro-poor HI for Luxembourg and Croatia for formal care, and Germany, Sweden, Italy, France, Denmark, Belgium, Luxembourg, Slovenia, Croatia, Bulgaria, Cyprus and Romania for informal care.

After accounting for need factors, formal care inequalities disappeared for most countries, while informal care inequalities remained in much greater number for most countries, mostly with pro-poor orientation. The inequality analysis shows that formal care use is more evenly distributed after accounting for legitimate sources of inequality than the CI analysis showed, so that inequality between the poor and CIs in many countries is mainly driven by legitimate sources of need. On the other hand, the use of informal care, although less pronounced than the CI distribution, is still pro-poor oriented, which in turn means that poorer people who cannot use formal care resort to informal sources even after accounting for need factors. A notable exception is Slovakia, where HI is even more pro-rich than the corresponding CI.

PUBLIC SECTOR
ECONOMICS
47 (2) 149-176 (2023)

Figure 6
Horizontal inequity



PUBLIC SECTOR
ECONOMICS
47 (2) 149-176 (2023)

5 CONCLUSION AND DISCUSSION

These findings lend considerable importance to the existing scientific approach to the long-term sector in relation to home care. The differences in the use of home care show that there is diversity among the 27 EU member states and that the use of home care is determined by a sum of need and non-need factors in the region that may influence the choice of care used. In addition, differences among income groups within EU-27 members and within countries themselves were found, even when controlling for need factors. Thus, it was possible to determine which factors cause inequality between socioeconomic groups in home care.

This study provides an overview of horizontal income inequality in the availability of home care among 27 European Union countries. In addition to using the most recent SHARE data collected last year, this study included, to our knowledge, the largest number of countries analyzed using traditionally developed methods for inequality. The problem of inequality has been previously addressed by authors Ilinca et al., (2017); Rodrigues, Ilinca and Schmidt (2017) using data from previous editions of SHARE and with a smaller number of countries observed in Europe. Our study confirms the findings of the aforementioned authors that a pro-poor distribution of long-term care prevails in most countries of the EU-27. In addition, due to the large volume of countries included in the research and the lack of relevant literature dealing with horizontal inequality in long-term care for the elderly, we believe that this study will make a great contribution to the consideration of sufficient equal services for the same group of users. Also, the latest data provide a clear overview of the current state of the use of formal and informal care in the European Union, which is of crucial importance in planning policies for the care of elderly people. Although we are aware that due to the market economy, it is not in the interest of countries to have perfect equality, we wanted to investigate which countries are close to this concept, that is, how formal and informal care are distributed within the EU member states in accordance with the level of income of individuals. It is particularly important to emphasize that at the very beginning of this analysis, the member states are not in the same starting position. Take, for example, the countries of the Balkans, where, due to the former communist regimes and stronger family attachment, there is a stronger preference for an informal form of care. The formal form is an alternative form of long-term care for the majority of users due to smaller national generosity (smaller grants) and greater reliance on family financial support.

Based on analysis of inequality, pro-poor distribution of LTC was indicated in most of the observed countries, which is consistent with the previous research (Ilinca et al., 2017). Health variables across countries contribute to pro-poor inequality in use of LTC, mainly for informal care. People with greater health needs are disproportionately more represented among poor population. Age greatly contributes to the pro-poor inequality in several countries, both for formal and informal care. The Gender variable did not much contribute to the overall CI. Females in most countries are disproportionately represented in the poorer population, while in most countries females utilize more LTC, which is consistent with other research on the subject (Forma et al., 2007; 2017). Regarding formal care, most countries have pro-poor contributions but surprisingly for some countries higher income is associated with lower levels of formal care use, which could be the result of lower income individuals being favored in state support, while higher income individuals find other means of meeting their needs. The Education variable shows all countries having a pro-rich partial CI, which can be seen as more-educated people experience fewer inequalities. Education and Income are in positive correlation, which means that higher-educated people enjoy better income. Bearing in mind that a pro-poor distribution is presented in the use of formal and informal care in European Union among elderly people, our initial hypothesis is accepted. Most contributions of the Area variable were pro-rich, as most countries have poorer rural areas and less utilization of LTC in such areas, probably because most countries have centralized systems of care placed in better developed areas. After accounting for need factors, formal care inequalities disappeared for most countries, while informal care inequalities remained in much greater numbers for most countries, mostly with pro-poor orientation.

6 LIMITATIONS

Two criteria were excluded from the analyzed sample: persons younger than 65 years of age and persons who have a permanent residence in long-term care institutions. Institutional care is understood as care for an elderly person in homes for the elderly or in any other facility of a formal nature (assisted living, day care) without including a stay in one's own home. It is characteristic for people who are institutionalized to have a high rate of morbidity and seriously impaired health, which can influence the underestimation of the connection of certain variables to the final outcome. When considering the sample, the mortality rate was not included (people of lower standards who live in rural areas have a shorter life expectancy), which can affect the outcome of the results.

The next type of limitation is the small representative sample, which may affect the study outcome results. Regarding sample sizes in relation to HI, most of the same conclusions given for CI hold here. Bootstrapped confidence intervals for HI are comparable to those for CI, so the problems with sample size pointed out in the CI section should here also be taken into consideration.

In addition, when talking about the health status of users, it should be kept in mind that this is a subjective perception by the participants of the long-term SHARE study. The study used total household income as a measure of total income, as opposed to equivalent income or per capita income, which are used in various scientific studies. Although some authors state net worth as more adequate, income, due to its measurability, variability (transition from working income to pensions) and frequency (the most common measurable data when expressing socioeconomic inequalities) represents a better choice for measuring the availability of long-term care among elderly people.

Disclosure statement

The authors hereby state that they are not aware of any conflict of interest that might influence the results or interpretation of the paper.

Acknowledgements

This paper uses data from SHARE Waves 1, 2, 3, 4, 5, 6, 7, 8 and 9 (DOIs: 10.6103/SHARE.w1.800, 10.6103/SHARE.w2.800, 10.6103/SHARE.w3.800, 10.6103/SHARE.w4.800, 10.6103/SHARE.w5.800, 10.6103/SHARE.w6.800, 10.6103/SHARE.w7.800, 10.6103/SHARE.w8.800, 10.6103/SHARE.w8ca.800, 10.6103/SHARE.w9ca800, 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, SHARELIFE: 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, and VS 2020/0313. 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, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

REFERENCES

1. Albertini, M. and Pavolini, E., 2017. Unequal inequalities: The stratification of the use of formal care among older Europeans. *Journals of Gerontology – Series B Psychological Sciences and Social Sciences*, 72(3), pp. 510-521. https://doi.org/10.1093/geronb/gbv038
2. Bago d'Uva, T., Jones, A. M. and van Doorslaer, E., 2009. Measurement of horizontal inequity in health care utilization using European panel data. *Journal of Health Economics*, 28(2), pp. 280-289. https://doi.org/10.1016/j.jhealeco.2008.09.008
3. Barbieri, D. and Ghibelli, P. (eds.), 2018. *Formal Versus Informal Long-term Care: Economic and Social Impacts*. https://doi.org/10.5281/zenodo.1410379
4. Bergmann, M. and Börsch-Supan, A. (eds.), 2021. *SHARE Wave 8 Methodology: Collecting Cross-National Survey Data in Times of COVID-19*. Munich: MEA, Max Planck Institute for Social Law and Social Policy.
5. Bonsang, E., 2009. Does informal care from children to their elderly parents substitute for formal care in Europe? *Journal of Health Economics*, 28(1), pp. 143-154. https://doi.org/10.1016/j.jhealeco.2008.09.002
6. Börsch-Supan, A. [et al.], 2013. *Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)*. https://doi.org/10.1093/ije/dyt088
7. Börsch-Supan, A., 2022. *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. Release version: 8.0.0. SHARE-ERIC*. https://doi.org/10.6103/SHARE.w8.800
8. Brackley, M. E. and Penning, M. J., 2009. Home-care utilization within the year of death: Trends, predictors and changes in access equity during a period of health policy reform in British Columbia, Canada. *Health and Social Care in the Community*, 17(3), pp. 283-294. https://doi.org/10.1111/j.1365-2524.2008.00830.x
9. Broese van Groenou, M. I. [et al.], 2006. Socio-economic status differences in older people's use of informal and formal help: A comparison of four European countries. *Ageing and Society*, 26(5), pp. 745-766. https://doi.org/10.1017/S0144686X06005241
10. Broese van Groenou, M. I. and de Boer, A., 2016. Providing informal care in a changing society. *European Journal of Ageing*, 13, pp. 271-279. https://doi.org/10.1007/s10433-016-0370-7
11. Carrieri, V. and Bilger, M., 2013. Preventive care: Underused even when free. Is there something else at work? *Applied Economics*, 45(2), pp. 239-253. https://doi.org/10.1080/00036846.2011.597729
12. Carrieri, V. and Wuebker, A., 2013. Assessing inequalities in preventive care use in Europe. *Health Policy*, 113(3), pp. 247-257. https://doi.org/10.1016/j.healthpol.2013.09.014
13. Carrieri, V., di Novi, C. and Orso, C. E., 2017. Home Sweet Home? Public Financing and Inequalities in the Use of Home Care Services in Europe. *Fiscal Studies*, 38(3), pp. 445-468. https://doi.org/10.1111/j.1475-5890.2017.12138

14. Chuakhamfoo, N. N. [et al.], 2020. Health and long-term care of the elderly with dementia in rural Thailand: A cross-sectional survey through their caregivers. *BMJ Open*, 10(3), pp. 1-11. https://doi.org/10.1136/bmjopen-2019-032637
15. EIGE, 2020. *Gender Equality and Long-Term Care at home.* Vilnius: European Institute for Gender Equality.
16. 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
17. Eurofound, 2017. *Income inequalities and employment patterns in Europe before and after the Great Recession*. Luxembourg: Publications Office of the European Union.
18. European Commission, 2008. *Long term care in the European Union. Employment, Social Affairs and Equal Opportunities*.
19. European Commission, 2021. *Long-term care report: trends, challenges and opportunities in an ageing society. Volume II, Country profiles*.
20. Fernandez, J. L. and Forder, J., 2015. Local variability in long-term care services: Local autonomy, exogenous influences and policy spillovers. *Health Economics (United Kingdom)*, 24(S1), pp. 146-157. https://doi.org/10.1002/hec.3151
21. Forma, L. [et al.], 2007. Health and social service use among old people in the last 2 years of life. *European Journal of Ageing*, 4(3), pp. 145-154. https://doi.org/10.1007/s10433-007-0054-4
22. Forma, L. [et al.], 2017. Trends in the use and costs of round-the-clock long-term care in the last two years of life among old people between 2002 and 2013 in Finland. *BMC Health Services Research*, 17(1). https://doi.org/10.1186/s12913-017-2615-3
23. 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). https://doi.org/10.3390/ijerph14101224
24. Kemper, P., 1992. The use of formal and informal home care by the disabled elderly. *Health Services Research,* 27(4), pp. 421-451.
25. Lera, J., Pascual-Sáez, M. and Cantarero-Prieto, D., 2021. Socioeconomic Inequality in the Use of Long-Term Care among European Older Adults: An Empirical Approach Using the SHARE Survey. *International Journal of Environmental Research and Public Health,* 18(1). https://doi.org/10.3390/ijerph18010020
26. OECD, 2005. *Long-term Care for Older People.* Paris: OECD.
27. Penning, M. J. [et al.], 2018. Long-term care trajectories in Canadian context: Patterns and predictors of publicly funded care. *Journals of Gerontology – Series B Psychological Sciences and Social Sciences*, 73(6), pp. 1077-1087. https://doi.org/10.1093/geronb/gbw104
28. Rehnberg, J. and Fritzell, J., 2016. The shape of the association between income and mortality in old age: A longitudinal Swedish national register study. *SSM – Population Health,* 2, pp. 750-756. https://doi.org/10.1016/j.ssmph.2016.10.005

29. Rodrigues, R. [et al.], 2013. The indirect costs of long-term care. *Social Situation Monitor Research Note,* No. 8/2013.
30. Rodrigues, R., Ilinca, S. and Schmidt, A. E., 2017. Inequality and inequity in the use of long-term care services in Europe: is there reason for concern? *Eurohealth,* 23(4), pp. 7-10.
31. Saito, T. [et al.], 2018. Income-based inequalities in caregiving time and depressive symptoms among older family caregivers under the Japanese long-term care insurance system: A cross-sectional analysis. *PLoS ONE*, 13(3). https://doi.org/10.1371/journal.pone.0194919
32. Sarasa, S. and Bilingsley, S., 2008. Personal and household care giving from adult children to parents and social stratification. In: C. Saraceno, ed. *Families, Ageing and Social Policy*. Cheltenham: Edward Elgar, pp. 123-146.
33. Tenand, M., Bakx, P. and van Doorslaer, E., 2020. Equal long-term care for equal needs with universal and comprehensive coverage? An assessment using Dutch administrative data. *Health Economics*, 29(4), pp. 435-451. https://doi.org/10.1002/hec.3994
34. van de Poel, E., van Doorslaer, E. and O'Donnell, O., 2012. Measurement of inequity in health care with heterogeneous response of use to need. *Journal of Health Economics*, 31(4), pp. 676-689. https://doi.org/10.1016/j.jhealeco.2012.05.005
35. van Deurzen, I., van Oorschot, W. and van Ingen, E., 2014. *The Link between Inequality and Population Health in Low and Middle Income Countries: Policy Myth or Social Reality?* https://doi.org/10.1371/journal.pone.0115109
36. van Doorslaer, E., Koolman, X. and Jones, A. M., 2004. Explaining income-related inequalities in doctor utilization in Europe. *Health Economics*, 13(7), pp. 629-647. https://doi.org/10.1002/hec.919
37. Wagstaff, A. and van Doorslaer, E., 2000. The Board of Regents of the University of Wisconsin System Measuring and Testing for Inequity in the Delivery of Health Care Measuring and Testing for Inequity in the Delivery of Health Care. *The Journal of Human Resources,* 35(4), pp. 716-733. https://doi.org/10.2307/146369
38. 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
39. WHO, 1991. *Health for all targets: the health policy for Europe*. Copenhagen: World Health Organization, Regional Office for Europe.

Pension expenditure determinants: the case of Portugal

MARIA TERESA MEDEIROS GARCIA, Ph.D., Associate Professor*
ANDRÉ FERNANDO RODRIGUES ROCHA DA SILVA, MSc*

Article**
JEL: C32, C51, C52, H55
https://doi.org/10.3326/pse.47.2.2

* The authors acknowledge financial support from FCT – Fundação para a Ciência e a Tecnologia (Portugal), national funding through research grant UIDB/05069/2020. The authors are grateful to two anonymous referees who have contributed to the quality of the final version of the paper.

** Received: April 21, 2022
Accepted: January 23, 2023

Maria Teresa MEDEIROS GARCIA
Lisbon School of Economics and Management, Universidade de Lisboa, Rua Miguel Lupi, 20, 1249-078 Lisboa, Portugal
e-mail: mtgarcia@iseg.ulisboa.pt
ORCiD: 0000-0001-8683-2112

André Fernando Rodrigues ROCHA DA SILVA
Lisbon School of Economics and Management, Universidade de Lisboa, Rua Miguel Lupi, 20, 1249-078 Lisboa, Portugal
e-mail: andresilva_gk@hotmail.com
ORCiD: 0000-0001-8123-6557

Abstract

Assessing pension expenditure determinants is crucial for the sustainability of public finances. This study aims to disentangle the impact of demographic and economic variables, such as ageing, productivity, and unemployment, on pension expenditure in Portugal. With the use of time-series data, from 1975 to 2014, statistical evidence was found of co-integration between unemployed people aged between 15 and 64 years old, apparent productivity of labour, the old age dependency ratio and pension expenditure as a share of gross domestic product. The use of a vector error correction model, with impulse-response functions and variance decomposition, showed that ageing has an almost insignificant impact in the long-run, when compared with unemployment and productivity.

Keywords: pension expenditure, ageing, productivity, unemployment, linear regression analysis

1 INTRODUCTION

Worldwide, there is increasing interest in the analysis of the impact of ageing, productivity, and unemployment on pension expenditure. The concern of European social security systems with the rise of pension expenditure has motivated several reforms, which have included adjustments of the age eligibility for a pension benefit and of the size of the pension benefit (Eurogroup, 2016; 2017; European Commission, 2014).

However, a public pension system is expected to experience a pattern of increasing expenditures from the early years of its existence, until it reaches a state of maturity (Plamondon et al., 2002). After a period of 65 to 70 years, under stable conditions, the expenditure of a scheme expressed as a percentage of insured earnings normally stabilizes, since the first generation of young new entrants to the scheme has passed through the various stages of participation.

Pension schemes mature very slowly, that is, over many decades (Cichon et al., 2004). Moreover, increases in pension expenditures are a perfectly normal phenomenon during the maturation phase of national pension schemes, which lasts several decades. Rising pension expenditures are not per se necessarily indicative of a financial sustainability issue.

Therefore, the design of pension financing systems should accommodate this expected growth of pension expenditure. However, pension privatization policies, implemented in a number of countries, in consequence of concern with the pattern of increasing pension expenditure (World Bank, 1994), did not deliver the expected results, as coverage[1] and pension benefits[2] did not increase, systemic

[1] Coverage (also denoted contribution density) is defined both as the proportion of workers participating in pension schemes and the proportion of the elderly receiving some kind of pension income (OECD/Inter-American Development Bank/The World Bank, 2014).

[2] These pension benefits are available to people who have reached pensionable age through: (i) earnings-related contributory pensions (guaranteeing minimum benefit levels, or replacement rates corresponding to a prescribed proportion of an individual's past earnings – in particular for those with lower earnings); and/or (ii) flat-rate pensions (mostly residency-based and financed by the general budget) and/or means-tested pensions (ILO, 2018).

risks were transferred to individuals and fiscal positions worsened (Beattie and McGillivray, 1995; ILO, 2018). Consequently, several countries are reversing privatization measures and returning to public solidarity-based systems.

In addition, recent austerity or fiscal consolidation trends affected the adequacy of pension systems and general conditions of retirement, putting at risk the fulfilment of the minimum standards in social security and, consequently, the contribution of public pension systems to the Sustainable Development Goals (SDGs) (ILO, 2017; 2018).

Few studies are available regarding the factors that influence the evolution of Portuguese pension expenditure, and whether there is a link between pension expenditure as a dependent variable and other relevant explanatory variables, including the most recent developments on relevant variables, covering the current environment and data (Garcia and Lopes, 2009; Garcia, 2017).

This paper aims to understand which variables have a relevant influence on social security pension expenditure using econometric techniques that include a vector error correction model (VECM).

In the next section we briefly describe the Portuguese public pension system. Next we review the literature covering the impact of ageing on several macroeconomic variables especially pension expenditure. In the methods section, we present our variables and methodology, as well as data analysis. In the following section, we show our estimation results. The last sections provide the discussion and the conclusion.

2 THE PORTUGUESE PENSION SYSTEM

The Portuguese pension system is an earnings-related public pension scheme with a means-tested safety net (OECD, 2015), which is financed both by contributions from employees and employers, and by transfers from the State Budget.

Throughout its existence, several measures have been enacted, allegedly to reinforce the pension system's financial sustainability, such as the creation of the public pension reserve fund in 1989, and the convergence of the civil servants' scheme with the public pension system that covers the private sector in 2005.

In 2007, a sustainability factor was introduced for the calculation of the old age pension benefit, reducing it so that it takes life expectancy into account. Indeed, Portugal's population is ageing very rapidly and shrinking, due among other things to very low fertility rates (OECD, 2019). A further change came in 2013, with a decrease in the pension benefit, although this only covered early retirement. This reform, whose effects will mainly be felt in the medium and long term, also intended to promote the financial sustainability of public finances, reducing the expected value of future pension expenditure and replacement rates. Simultaneously, as a consequence of the Portuguese bailout in 2011 (European Commission, 2011), an extraordinary solidarity contribution was also introduced which decreased all pension income.

PUBLIC SECTOR
ECONOMICS
47 (2) 177-203 (2023)

In 2013, the normal retirement age was established at 66 years in 2014, but increased to 66 years and two months in 2015, following the automatic process of adjusting the normal age of retirement by two-thirds of gains in life expectancy from age 65, measured as the average of the previous two years (Garcia, 2017).

In summary, Portugal essentially has a pay-as-you-go pension scheme (World Bank, 2006), which represents the major source of retirement income, with occupational and personal pension funds only existing to a minor extent (Blake, 2006; European Parliament, 2011; Garcia, 2017). The Portuguese system is also a defined-benefit system (European Commission, 2015), offering pensioners more measurable post-employment income benefits (Ramaswamy, 2012). Pensions are indexed to prices and gross domestic product (European Commission, 2015; OECD, 2019).

3 LITERATURE REVIEW

Demographic aging and its impact on pension expenditure brought to the debate the need to reform public pension systems (European Commission, 2012; 2015; ECB, 2015; OECD, 2015). In the case of the United States of America, Roach and Ackerman (2005) show that a wide range of existing policy options could be used to secure the finances of the U.S. social security programme over the next 75 years without major structural changes, whereby it will continue to provide beneficiaries with a stable and predictable source of retirement income. These authors believe that the system is not in crisis and that it cannot go bankrupt as long as revenues continue to be collected. Focusing on the major industrial economies, Ramaswamy (2012) stresses the ideas that lower payroll tax revenues during a period of high unemployment and rising fiscal deficits are a test of the sustainability of pay-as-you-go (PAYG) public pension schemes, while poor financial market returns and low long-term real interest rates create challenges for the defenders of defined benefit (DB) pension schemes. The author concludes that the projected increase in the old-age dependency ratio suggests that in many countries the contributions to pay-as-you-go schemes have to increase by 20% from current levels in 2020 to pay pensions. Furthermore, for occupational DB schemes that face large funding shortfalls, employer contributions will have to rise to improve the coverage ratio of these schemes. In addition, the author emphasises that as more employers progressively shift towards defined contribution (DC) schemes for providing post-employment benefits, regulatory policies might be needed to restrict the range of permissible investment options available for plan assets to avoid unintended risks being taken by the plan beneficiaries, and to set mandatory minimum contribution rates for participating in DC schemes.

Although to limit public expenses, pension benefits might be decreased, retirement income adequacy is a concern (European Parliament, 2011; Chybalski and Marcinkiewicz, 2014). In this context, Orenstein (2011) calls attention to the fact that, from 1981 to 2007, more than thirty countries worldwide fully or partially replaced their pre-existing PAYG pension systems with ones based on individual, private savings accounts in a process often labelled “pension privatisation”.

However, pension privatization did not deliver the expected results (ILO, 2018), revealing limited effects on capital markets and economic growth. In fact, coverage rates and pension benefits decreased, the risk of financial market fluctuations was shifted to individuals, and administrative costs increased. Moreover, the high costs of transition created large fiscal pressures. In addition, private pension fund administration did not improve governance as, frequently, the regulatory and supervisory functions were captured by economic groups responsible for managing the pension funds, allowing concentration in pension industry. By using an overlapping generations model with a PAYG pension system, Cipriani (2014) concludes that population ageing due to increased longevity implies a reduction in pension benefits. However, the efects of ageing on pensions may not be negative if the elderly are free to choose their retirement age, while they are always negative in the case of full retirement (Cipriani, 2016). In light of the 2008 economic crisis, Halmosi (2014) emphasises that the study of the pension systems of developed countries is a priority issue. Indeed, Grech (2015) presents evidence that the impacts of the crisis were different for Continental and Mediterranean systems, pension benefits of the latter being cut back significantly. Analysing pension reforms in Greece, Italy, Portugal, and Spain, between, 1990 and 2013, Natali and Stamanti (2014) conclude that all these countries encouraged the spread of private pensions and harmonised their fragmented public schemes. In addition, cost containment was massive, putting future adequacy at risk. In addition, Natali (2015) provides a summary of reforms in Europe since the onset of the Great Recession, providing evidence that austerity has hit both public pay-as-you-go schemes and private pre-funded schemes alike. Indeed, both have been subject to measures to contain costs (e.g., a higher pensionable age, the introduction of automatic stabilisers of future spending, reduced indexation, and higher taxes and/or contributions). In fact, Diamond (1996), much earlier, suggested the indexation of normal retirement age to life expectancy, and the investment of part of the public reserve funds in the private economy as being good measures to solve the social security pension system problem.

The implications of population ageing for economic growth are also a cause for concern. In this context, Bloom, Canning and Fink (2010) conclude that OECD countries are likely to see modest – but not catastrophic – declines in the rate of economic growth, emphasising that policy reforms (including an increase in the legal age of retirement) can mitigate the economic consequences of an ageing population. More recently, Žokalj (2016) examines the fiscal implications of the demographic shift using panel data on 25 European Union countries in the period from 1995 until 2014. The results suggest significant and positive impacts of the elderly share on expenditure for pensions and social protection.

In order to disentangle the macroeconomic impacts on the PAYG Portuguese social security system, Garcia and Lopes (2009) conclude that some cumulative measures, such as a changing of indexing rules, a better actuarial match between pensions and contributions, and measures to increase the effective age of retirement, could have a bigger impact on reducing the expected increase in pension expenditure than applying a systemic pension reform. Using a macroeconomic model of the Portuguese economy, the estimations suggest that the elimination of early retirement schemes, combined with an increase in the effective contribution rate could be a good alternative to promote the financial sustainability of the system. Economic growth strengthened by the pension reserve fund (which had an average annual nominal rate of return of 5.17% during the period 1989-2014, and relatively low administrative costs compared with funded systems), brings more advantages to the system when compared with a fully pre-funded system, which has high transition costs, with current tax payers being responsible for paying both their own and the existing pensioners' benefits (European Parliament, 2011).

This paper analyses the factors that influence the evolution of Portuguese pension expenditure, including the most recent developments on relevant variables besides the demographic aging.

4 METHODS

4.1 VARIABLES

The choice of both dependent and independent variables used in our empirical analysis follows the recent literature trend, as in European Commission (2015). In order to study the determinants of pension expenditures, we adopt the ratio between pension spending and gross domestic product at current prices as the dependent variable (pensions as percentage of gross domestic product).

The independent variables consider eight factors that could influence pension expenditure. The first group of factors follows the related literature concerning the macroeconomic and demographic characteristics:

(1) Unemployment refers to unemployed people defined as persons aged 15 to 64 without work during the reference week, available to start work within the next two weeks (or who has already found a job to start within the next three months), and has actively sought employment at some time during the last four weeks. In pay-as-you-go systems, unemployment shrinks the contribution base, negatively affecting the pension system balance.

(2) Apparent labour productivity denotes apparent productivity of labour that relates the wealth created to the labour factor. The apparent labour productivity is the real gross domestic product in terms of expenditure, at constant prices of 2011, per annual hours worked by employed people. Apparent labour productivity presents the potential to overcome the negative effects of ageing, positively affecting the pension system balance.

(3) Old age dependency ratio is the ratio between elderly people at an age when they are generally economically inactive (i.e. aged 65 and over) and the number of people of working age (i.e. 15-64 years old). This variable is expected to have a positive effect on the dependent variable.

The second group tries to disentangle the impact of the main pension system laws since 1975 (Garcia, 2017). Therefore, five dummy variables were set, each of which refers to a specific period, that is to say, the variable's value will be 1 if included in that specific period, and 0 otherwise. The events are:

(1) Revolution of April 1974 (Rev1974), which led to important social and economic changes during the second half of the 70s. This variable is expected to have a positive effect on the dependent variable.
(2) The first Social Security Act of 1984 (R1984), which established pension benefit payments in the private sector. This variable is also expected to have a positive effect on the dependent variable.
(3) The Social Security Reform of 1993 (R1993), which made changes to the existent social security system of the Public Administration (civil servants), in order to adopt the same features (eligibility and benefits) established for the private sector. This reform considers a new formula for the calculation of public employees' pensions, which is the same as that of the private sector workers' scheme. This variable is expected to have a negative effect on the dependent variable.
(4) The Third Social Security Act of 2002 (R2002), which considered parametric changes to the old age pension benefit formula, including the accrual rate and life-time earnings. This variable is expected to have a negative effect on the dependent variable.
(5) The Fourth Social Security Act of 2007 (R2002), which introduced the sustainability factor and the voluntary public regime of capitalisation. The sustainability factor is the ratio between average life expectancy at the age of 65 in 2000 and average life expectancy at the age of 65 for the year prior to the year for which the pension benefit is calculated. This Act also increases the penalty for early retirement to 6% per year. This variable is also expected to have a negative effect on the dependent variable.

4.2 METHODOLOGY

We conduct linear regression analysis using annual time series data from 1975 to 2014. This timespan takes in 40 years, starting immediately after the revolution of 1974 and ending in 2014, the year when the 3-year period of the Portugal bailout ended. Indeed, to prevent an insolvency situation in the debt crisis, Portugal applied in April 2011 for bailout programs and drew a cumulated €78 billion from the IMF (International Monetary Fund), the EFSM (European Financial Stabilisation Mechanism), and the EFSF (European Financial Stability Facility). Portugal exited the bailout in May 2014, the same year that positive economic growth reappeared following three years of recession (OECD, 2014).

The equation of the model is as follows:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \delta_0 D_{1t} + \delta_1 D_{2t} + \delta_2 D_{3t} + \delta_3 D_{4t} + \delta_4 D_{5t} + \varepsilon_t \quad (1)$$

where Y is the ratio between pension spending and gross domestic product; X_1 is the unemployment in logarithmic form; X_2 is the apparent labour productivity in logarithmic form; X_3 is the old age dependency ratio; and D_1 to D_5 represent dummy explanatory variables used to indicate the occurrence of the events described above. Similarly, the model equation is given by:

$$\begin{aligned}\frac{pensions}{gross\ domestic\ product}_t &\\ &= \beta_0 + \beta_1 log\ unemplyment_t + \beta_2 \log \text{apparent labor productivity}_t \\ &+ \beta_3 old\ age\ dependency\ ratio_t + \delta_0 Rev1974_t + \delta_1 R1984_t + \delta_2 R1993_t \\ &+ \delta_3 R2002_t + \delta_4 R2007_t + \varepsilon_t\end{aligned} \quad (1a)$$

The data source is PORDATA (Francisco Manuel dos Santos Foundation, 2010) and the descriptive statistics for the variables used in the analysis are presented in the appendix (table A1).

4.3 DATA ANALYSIS

To test for stationarity, unit root tests were undertaken (Wooldridge, 2009). Following the methodology adopted by Brooks (2014), the tests used were the augmented Dickey-Fuller test and Phillips-Perron test (table A2). The p-values analysis of both tests suggests that the null hypothesis of the presence of a unit root cannot be rejected in all variables at 10% significance level, and that stationarity is achieved with first differences through the rejection of the same null hypothesis at 5% significance level, highlighting their strong persistence (I(1) process).

Non-stationarity may render the potential econometric results statistically invalid. Typically, the linear combination of I(1) variables will be I(1), but it is desirable to obtain I(0) residuals, which are only achieved if the linear combination of I(1) variables is I(0), that is to say, if the variables are co-integrated (Brooks, 2014).

With regards to the hypothesis of the existence of more than one linearly independent co-integration relationship between more than two variables, it is appropriate to stress the issue of co-integration using the Johansen VAR test. To develop the Johansen VAR framework, the selection of the optimum number of lags is needed to avoid problems of residual autocorrelation, using the VAR Lag Order Selection Criteria procedure. The Likelihood Ratio Criteria (LR), the Final Predictor Error (FPE), and the Hannan-Quinn Information Criteria (HQ) selected two lags as an optimum limit, against the evidence of the Akaike Information Criteria (AIC) and the Schwarz Information Criteria (SC), which presented the optimum selection of three and one lag, respectively (table A3).

The Johansen co-integration test allows for the selection of the appropriate lag length and model to choose (table A4). The test result suggests that the number of appropriated lags is two (as stated above), with one co-integrating vector, and the model to adopt consists of the allowance of a quadratic deterministic trend, with intercept and trend in the co-integration equation and intercept in VAR, following Akaike Information Criteria (Brooks, 2014).

Therefore, it was decided to use an error correction model "incorporated" into a VAR framework in order to model the short and long-run relationships between variables: a Vector Error Correction Model (VECM). The VECM can be set up in the following form (Brooks, 2014):

$$\Delta\gamma_t = \Pi\gamma_{t-k} + \Gamma_1 \Delta\gamma_{t-1} + \cdots + \Gamma_{k-1} \Delta\gamma_{t-(k-1)} + u_t \tag{2}$$

where $\Pi = (\sum_{i=1}^{k} \beta_i) - I_g$ and $\Gamma_i = (\sum_{j=1}^{i} \beta_j) - I_g$.

This VECM contains *g* variables in first-differenced form on the LHS, and *k-1* lags of the dependent variables (differences) on the RHS, each with a Γ short-run coefficient matrix. Π consists of a long-run coefficient matrix, as being in equilibrium, all the $\Delta\gamma_{t-1} = 0$, and setting $E(u_t) = 0$ will leave $\Pi\gamma_{t-k} = 0$. Π illustrates the speed of adjustment back to equilibrium, that is to say, it measures the proportion of the last period´s equilibrium error that it is corrected for (Brooks, 2014).

The VECM model estimation is depicted in table 1 and encompasses the co-integration equation with dummy variables.

As all inverse roots of the characteristic polynomial are inside the unit circle, the model is stable. The residuals assumptions were tested, and it is possible to conclude that the mean of the residuals is zero (table A5). The White heteroscedasticity test p-value does not allow for the rejection of homoscedastic residuals (table A6). In addition, the covariance between residuals and explanatory variables is zero, thus satisfying the assumption of there being no relationship between them (table A7) and that the residuals are normally distributed (table A8). Finally, the null hypothesis of no residual serial correlation is not rejected at 5% significance level with the use of two lags (table A9).

As such, the estimators are efficient, and the confidence intervals and hypothesis tests using *t* and F-statistics are reliable.

Table 1
VECM estimation results

Cointegrating Eq	**CointEq1**			
Pensions to gross domestic product ratio (-1)	1.000000			
Log unemployment (-1)	-0.9342			
	(0.0849)			
	[-11.0107]			
Log apparent labuor productivity (-1)	-3.4509			
	(0.6157)			
	[-5.6050]			
Old age dependency ratio (-1)	-0.1141			
	(0.0736)			
	[-1.5509]			
@TREND(75)	0.0248			
C	18.1948			
Error Correction:	**D(Pensions to gross domestic product ratio)**	**D(Log unemployment)**	**D(Log apparent labor productivity)**	**D(Old age dependency ratio)**
CointEq1	-0.8237	0.3550	-0.0850	-0.2292
	(0.2515)	(0.2710)	(0.0245)	(0.1965)
	[-3.2759]	[1.2683]	[-3.4650]	[-1.1663]
D(Pensions to gross domestic product ratio (-1))	0.0487	-0.1340	0.0380	0.2164
	(0.2394)	(0.2665)	(0.0234)	(0.1871)
	[0.2035]	[-0.5030]	[1.6264]	[1.1568]
D(Pensions to gross domestic product ratio (-2))	-0.0235	-0.3270	0.0237	0.1410
	(0.1943)	(0.2163)	(0.0190)	(0.1518)
	[-0.1210]	[-1.5120]	[1.2523]	[0.9287]
D(Log unemployment (-1))	0.4020	0.6528	-0.0026	0.1478
	(0.2435)	(0.2711)	(0.0238)	(0.1903)
	[1.6510]	[2.4081]	[-0.1090]	[0.7768]
D(Log unemployment (-2))	-0.0061	0.1362	-0.0544	-0.1486
	(0.2421)	(0.2610)	(0.0236)	(0.1892)
	[-0.0252]	[0.5054]	[-2.3002]	[-0.7854]
D(Log apparent labour productivity (-1))	-0.2214	3.2461	-0.4638	-1.2344
	(2.3580)	(2.6252)	(0.2301)	(1.8426)
	[-0.0939]	[1.2366]	[-2.0155]	[-0.6699]
D(Log apparent labour productivity (-2))	0.5801	0.7765	-0.0560	-0.4968
	(1.6070)	(1.7890)	(0.1568)	(1.2557)
	[0.3610]	[0.4340]	[-0.3569]	[-0.3956]
D(Old age dependency ratio (-1))	0.1703	-0.3452	0.0351	0.4682
	(0.2614)	(0.2910)	(0.0255)	(0.2043)
	[0.6517]	[-1.1863]	[0.3746]	[2.2924]
D(Old age dependency ratio (-2))	0.3714	-0.0353	0.0464	0.1500
	(0.2094)	(0.2331)	(0.0204)	(0.1636)
	[1.7736]	[-0.1514]	[2.2717]	[0.9169]
C	-0.0661	-0.0101	0.0013	-0.0661
	(0.1389)	(0.1547)	(0.0136)	(0.1086)
	[-0.4756]	[-0.0655]	[0.0977]	[-0.6085]

Sample (adjusted): 1978 to 2014
Included observations: 37 after adjustments
Standard errors in ( ) & t-statistics in [ ]
Determinant residual covariance (dof adj.) 9.01E-11
Determinant residual covariance 9.35E-12
Log likelihood 259.8113
Akaike information criterion -10.36818
Schwarz criterion -7.407571

Error Correction:	**D(Pensions to gross domestic product ratio)**	**D(Log unemployment)**	**D(Log apparent labor productivity)**	**D(Old age dependency ratio)**
	0.0071	0.0127	-0.0014	0.0040
@TREND(75)	(0.0150)	(0.0167)	(0.0015)	(0.0117)
	[0.4717]	[0.7591]	[-0.9283]	[0.3408]
	-0.2682	-0.0022	0.0537	0.1629
REV1974	(0.1366)	(0.1521)	(0.0133)	(0.1067)
	[-1.9639]	[-0.0147]	[4.0319]	[1.5267]
	0.0776	-0.1999	0.0541	0.2442
R1984	(0.1194)	(0.1329)	(0.0117)	(0.0933)
	[0.6499]	[-1.5040]	[4.6463]	[2.6173]
	-0.3833	0.0323	-0.0403	-0.1179
R1993	(0.1743)	(0.1940)	(0.0170)	(0.1362)
	[-2.1987]	[0.1665]	[-2.3696]	[-0.8657]
	-0.0999	-0.0041	0.0021	-0.1449
R2002	(0.1684)	(0.1875)	(0.0164)	(0.1316)
	[-0.5931]	[-0.0221]	[0.1281]	[-1.1014]
	0.1692	-0.0406	0.0007	0.1799
R2007	(0.1028)	(0.1145)	(0.0100)	(0.0803)
	[1.6459]	[-0.3544]	[0.0705]	[2.2399]
R-squared	0.6850	0.4133	0.8028	0.8424
Adj. R-squared	0.4601	-0.0058	0.6619	0.7298
Sum sq. resids	0.3049	0.3779	0.0029	0.1862
S.E. equation	0.1205	0.1341	0.0118	0.0942
F-statistic	3.0451	0.9862	5.6986	7.4815
Log likelihood	36.2744	32.3037	122.3691	45.4012
Akaike AIC	-1.0959	-0.8813	-5.7497	-1.5892
Schwarz SC	-0.3993	-0.1847	-5.0531	-0.8926
Mean dependent	0.1243	0.0231	0.0202	0.3676
S.D. dependent	0.1640	0.1338	0.0202	0.1811

Source: Authors' computation.

5 RESULTS

The results suggest that the long-run relationship between pensions to gross domestic product ratio and old age dependency ratio is negative, whereas the long-run relationship between pensions to gross domestic product ratio and the other two variables (log unemployment and log apparent labour productivity) is positive. In fact, the normalised co-integrating model estimation (table A10 in the appendix), without dummy variables, allows the following equation to be obtained:

Pensions in percentage of gross domestic product = 1.320370 log unemployment
+ 1.818858 log apparent labour productivity – 0.221652 old age dependency ratio

The presence of a co-integrating vector illustrates an equilibrium phenomenon, as it is possible that co-integrated variables may deviate from their relationship in the short run, but that their association will return in the long run (Brooks, 2014).

The positive long-run coefficient of log unemployment suggests that unemployment has a positive impact on pension system expenditure, which is in line with the literature. High unemployment leads to negative migratory balances (mostly affecting young people), aggravating the ageing process, and consequently the declining demographics. With fewer people, investment decreases, shrinking economic growth. The causality from ageing and unemployment to productivity is confirmed by a VEC Granger Causality Test, at 5% and 10% significance level, respectively.

The positive long-run coefficient of log apparent labour productivity on pensions to gross domestic product ratio is not in line with European Commission (2015).

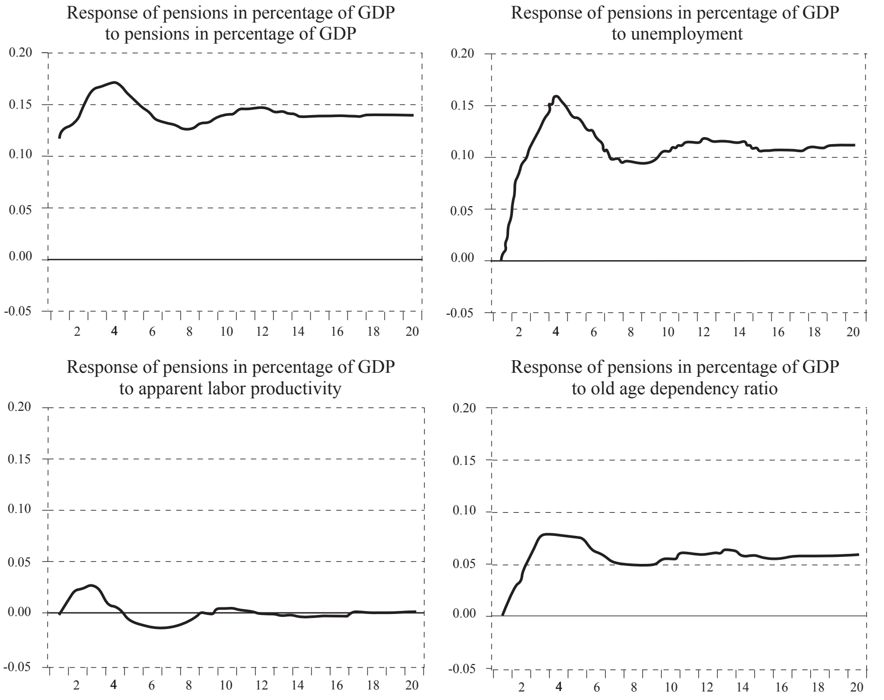
Concerning the negative coefficient of the old age dependency ratio, this might be the consequence of the parametric changes introduced to the system since 2000 (Garcia, 2017), especially the change of the normal retirement age (NRA) to 66 years, in 2013, becoming life expectancy-dependent after 2014. Therefore, an increase of the old age dependency ratio does not necessarily imply an increase of pension expenditure as a share of gross domestic product in the long-run. This measure is strongly supported by the literature as a crucial measure to guarantee the financial sustainability of pension systems, smoothing the impact of an ever-increasing number of pensioners (Diamond, 1996; Clements et al., 2015). The introduction of a sustainability factor into the benefit calculation formula, which is related to the evolution of average life expectancy (ALE), also represents a significant decrease in the pension benefit.

With regards to the short-run coefficients of the dummy variables, only the revolution of April 1974 (at 10% significance level) and the 1993 Social Security Reform (at 5%) present statistical significance, and the negative coefficients illustrate each contribution to the decrease of pension expenditure as a share of gross domestic product, where the possible causes can be the high average real gross domestic product growth rate after 1976 and until 1979 of 5.4% in the first case (PORDATA), and in the latter case, the equalisation of the official retirement ages for men and women, as well as the increase of the minimum contributory period from 10 to 15 years.

Finally, the impulse-response functions were stressed, as well as the variance decomposition for pensions to gross domestic product ratio, which is strongly dependent on the Cholesky ordering, which does not follow a specific requirement (Brooks, 2014). In order to guarantee some consistency and reasonability in the results, the order considered was from the most exogenous variable to the most endogenous one, determined by a VEC Granger Causality Test. The higher the p-value, the greater the exogeneity of the variable. The adopted order is as follows: old age dependency ratio, log unemployment, pensions to gross domestic product ratio and log apparent productivity of labour.

Figure 1

Pensions to gross domestic product ratio response to Cholesky one standard deviation impulse



Source: Authors' computation.

Following the methodology of Brooks (2014), figure 1 gives the impulse responses of the pensions to gross domestic product ratio, regarding unit shocks in the old age dependency ratio, in the log unemployment, and in log apparent productivity of labour, and their impact during 20 periods (years) ahead. Considering the signs of the responses, shocks in the old age dependency ratio have a positive impact until the 5th year, achieving its peak in the 3rd year. After this, the impact is negative, although the effect of the shock ends up dying down. A standard deviation shock in log unemployment and log apparent productivity of labour always has a positive impact on pensions to gross domestic product ratio, reaching its peak in the 4th and 3rd years, respectively, and stagnating in the long-run. Finally, the own impulse in pensions to gross domestic product ratio registers a similar impact as log unemployment, that is to say, it reaches its peak in the 4th year, and then stagnation thereafter.

When analysing this approach, the main highlight is the fact that the old age dependency ratio registers an almost irrelevant contribution for the evolution of pensions to gross domestic product ratio in the long-run, when compared with the other variables, which is surpassed by the effects of log unemployment and log apparent productivity of labour, this reinforcing the doubts about the impact of ageing on pension expenditure. It is also possible to verify the relevance of unemployment in the presence of a positive shock immediately in the first years (as stressed by European Commission, 2015), over a 20-year forecasting horizon (positive but constant impact), shrinking the contributory base and the economic growth, with a similar pattern in relation to the apparent productivity of labour, guaranteeing higher pension entitlements.

The results of the variance decomposition of the pensions to gross domestic product ratio residuals show that, for the 20-year forecasting horizon, the old age dependency ratio shocks account for only 2.86%, in the first year, and 5.35%, in the 20th year, of the variance of the pensions to gross domestic product ratio, while log unemployment contributes between 57.87% and 85.83%, reinforcing the huge importance of unemployment on pension expenditure and the reduced impact of ageing when compared with the other variables (table A11). It is also important to stress the own shocks of pensions to gross domestic product ratio, which accounts for between 39.76% and 0.93% of its movements.

The negative relationship between pensions to gross domestic product ratio and old age dependency ratio supports the hypothesis of a spurious result. Therefore, the Johansen co-integration test with dummy variables was carried out (table A12), although there is a problem in that the critical values may not be valid with exogenous series, such as dummy variables (Johansen, Mosconi and Nielsen, 2000; Giles and Godwin, 2012).

With this test, the old age dependency ratio long-run coefficient becomes positive and the sign of the other two coefficients does not change. However, it is important to take into account the econometric limitations of this change. To derive the VECM p-values, the VECM model with the coefficients as C(1) until C(16) was developed (table 2). C(1) is the coefficient of the co-integration equation (as well as the speed of adjustment back to equilibrium), C(10) is the constant, C(2) up to C(9) are the short-run coefficients of the lagged variables (until the second lag), and C(12) until C(16) are the coefficients of the dummy variables. C(11) is the trend coefficient (Brooks, 2014).

Table 2

VECM model with p-values

	Coefficient	Std. error	t-Statistic	Prob.
C(1)	-0.8237	0.2515	-3.2759	0.0036
C(2)	0.0487	0.2394	0.2035	0.8407
C(3)	-0.0235	0.1943	-0.1210	0.9049
C(4)	0.4020	0.2435	1.6510	0.1136
C(5)	-0.0061	0.2421	-0.0252	0.9801
C(6)	-0.2214	2.3580	-0.0939	0.9261
C(7)	0.5801	1.6070	0.3610	0.7217
C(8)	0.1703	0.2614	0.6517	0.5217
C(9)	0.3714	0.2094	1.7736	0.0906
C(10)	-0.0661	0.1389	-0.4756	0.6393
C(11)	0.0071	0.0150	0.4717	0.6420

	Coefficient	Std. error	t-Statistic	Prob.
C(12)	-0.2682	0.1366	-1.9639	0.0629
C(13)	0.0776	0.1194	0.6499	0.5228
C(14)	-0.3833	0.1743	-2.1987	0.0392
C(15)	-0.0999	0.1684	-0.5931	0.5595
C(16)	0.1692	0.1028	1.6459	0.1147
R-squared	0.6850	Mean dependent var	0.1243	
Adjusted R-squared	0.4601	S.D. dependent var	0.1640	
S.E. of regression	0.1205	Akaike info criterion	-1.0959	
Sum squared resid	0.3049	Schwarz criterion	-0.3993	
Log likelihood	36.2744	Hannan-Quinn criter.	-0.8503	
F-statistic	3.0451	Durbin-Watson stat	2.3277	
Prob(F-statistic)	0.0097			

Dependent Variable: D(Pensions to gross domestic product ratio)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Sample (adjusted): 1978 2014
Included observations: 37 after adjustments
D(Pensions to gross domestic product ratio) = C(1)*(Pensions to gross domestic product ratio (-1) - 0.934243024013* Log unemployment (-1) - 3.45091727663* Log apparent labor productivity (-1) -0.114073635473* Old age dependency ratio (-1) + 0.02475749296*@TREND(75) + 18.1948315066) + C(2)*D(Pensions to gross domestic product ratio (-1)) + C(3) *D(Pensions to gross domestic product ratio (-2)) + C(4)*D(Log unemployment (-1)) + C(5) *D(Log unemployment (-2)) + C(6)*D(Log apparent labour productivity (-1)) + C(7)*D(Log apparent labour productivity (-2)) + C(8) *D(Old age dependency ratio (-1)) + C(9)*D(Old age dependency ratio (-2)) + C(10) + C(11)*@TREND(75) + C(12)*REV1974 + C(13)*R1984 + C(14)*R1993 + C(15)*R2002 + C(16)*R2007

C(1), which is negative and statistically significant at 5%, confirms the long-run relationship between pensions to gross domestic product ratio, log unemployment, log apparent labour productivity, and the old age dependency ratio, as well as the existence of a correction mechanism of deviations (Wooldridge, 2009). When carrying out the Wald tests (table A13), it is not possible to reject the null hypothesis of C(4)=C(5)=0, C(6)=C(7)=0 and C(8)=C(9)=0, and the conclusion that needs to be stressed is the absence of short-run causality running from log unemployment, log apparent labour productivity, and the old age dependency ratio to pensions to gross domestic product ratio.

In addition, the results need to be analysed carefully: if the order of variables changes, then the results of impulse-response functions and variance decomposition can change drastically, mainly the variance decomposition between pensions to gross domestic product ratio and log unemployment. Nevertheless, it is noticeable that unemployment strongly influences pension expenditure behaviour.

6 CONCLUSION

The results of the estimation, after taking into consideration certain aspects such as non-stationarity, co-integration, and residuals testing, suggest that unemployment, apparent productivity of labour, and the old age dependency ratio all jointly present a long-run relationship with pension expenditure as a share of gross domestic product, but not in the short-run.

Unemployment is crucial to explain the increase of pension expenditure as a share of gross domestic product, as reinforced by the review of the literature on pensions. This interpretation is confirmed by the variance decomposition of pensions to gross domestic product ratio and also the impulse-response functions.

The apparent productivity of labour also seems to have a positive impact on pension expenditure to gross domestic product, which is not in line with European Commission (2015), supporting the assumption that gross domestic product growth is larger than pension expenditure growth in Portugal, because pensions are not fully indexed to wages after retirement.

The most intriguing result concerns the old age dependency ratio. In fact, after the development of the Johansen co-integration tests, both without dummy variables and with dummy variables, the old-age dependency ratio long-run coefficient presents different signs, giving rise to the hypothesis that ageing may not be the most relevant factor jeopardising the financial sustainability of the Portuguese public pension system. This is corroborated by the irrelevance of the influence of the old-age dependency ratio (in the long-run) on the impulse-response-functions, suggesting that the system has reached a state of maturity (Plamondon et al., 2002). Furthermore, this is in line with European Commission projections presented in the 2021 Ageing Report concerning Portugal (European Commission, 2021). Indeed, the country is expected to experience an overall decline in public pension expenditure (-3.2 percentage points) from 2019 to 2070 while the share of the age cohorts above 65 years in the total population is expected to rise from 22% to 33.1% (11.1 percentage points) in the same period.

When designing a pension system policy to reinforce the financial sustainability of the system, policy makers should take these findings into account. In other words, apparently, an increasing demographic strain seems not to impact pension expenditure as critically as unemployment. Therefore, policies to reduce unemployment should be considered as policy options to control pension expenditure, which represents a difficult way to address the financial sustainability of public pension systems. This is even more challenging in a stagflation environment, since actions intended to lower inflation may exacerbate unemployment. Future lines of research should try to do a similar analysis including more recent data and to disentangle the shape of pension expenditures over time (whether it follows an expected a logistical curve) (Cichon et al., 2004).

Disclosure statement
The authors declare that they do not have any conflict of interest.

REFERENCES

1. Beattie, R. and McGillivray, W., 1995. A risky strategy: Reflections on the World Bank Report Averting the old age crisis. *International Social Security Review*, 48(3-4), pp. 5-23. https://doi.org/10.1111/j.1468-246X.1995.tb00443.x
2. Blake, D., 2006. *Pension Economics*. Toronto: John Wiley & Sons.
3. Bloom, D. E., Canning, D. and Fink, G., 2010. Implications of population ageing for economic growth. *Oxford Review of Economic Policy*, 26(4), pp. 583-612. https://doi.org/10.1093/oxrep/grq038
4. Brooks, C., 2014. *Introductory econometrics for finance*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139540872
5. Chybalski, F. and Marcinkiewicz, E., 2014. How to measure and compare pension expenditures in cross-country analyses? Some methodological remarks. *International Journal of Business and Management,* 2(4), pp. 44-59.
6. Cichon, M. [et al.], 2004. *Financing social protection.* Geneva: International Labour Office and International Social Security Association.
7. Cipriani, G. P., 2014. Population aging and PAYG pensions in the OLG model. *Journal of Population Economics,* 27, pp. 251-256. https://doi.org/10.1007/s00148-013-0465-9
8. Cipriani, G. P., 2016. Aging, Retirement and Pay-As-You-Go Pensions. *IZA Discussion Paper,* No. 9969. https://doi.org/10.2139/ssrn.2790006
9. Clements, B. [et al.], 2015. The Fiscal Consequences of Shrinking Populations. *IMF Staff Discussion Note,* No. 21. https://doi.org/10.5089/9781513544885.006
10. Diamond, P. A., 1996. Proposals to Restructure Social Security. *Journal of Economic Perspectives,* 10(3), pp. 67-88. https://doi.org/10.1257/jep.10.3.67
11. Eurogroup, 2016. *Common principles for strengthening pension sustainability*, 16 June 2016.
12. Eurogroup, 2017. *Structural reform agenda – thematic discussions on growth and jobs: benchmarking pension sustainability,* 144/17, 20 March 2017.
13. ECB, 2015. Box 7 The 2015 Ageing Report: how costly will ageing in Europe be? *Economic Bulletin*, Issue 4/2015.
14. European Commission, 2011. The Economic Adjustment Program of Portugal. *Occasional Paper,* No. 79.
15. European Commission, 2012. The 2012 Ageing Report. *European economy,* No. 2.
16. European Commission, 2014. Identifying fiscal sustainability challenges in the areas of pension, health care and long-term care policies. *Occasional Papers,* No. 201.
17. European Commission, 2015. The 2015 Ageing Report: Economic and budgetary projections for the 28 EU Member States (2013-2060). *European economy,* No. 3.
18. European Commission, 2021. The 2021 Ageing Report: economic and budgetary projections for the EU Member States (2019-2070). *Institutional paper,* No. 148.
19. European Parliament, 2011. *Pension Systems in the EU – Contingent Liabilities and Assets in the Public and Private Sector*.

20. Francisco Manuel dos Santos Foundation, 2010. *PORDATA: the database of contemporary Portugal*.
21. Garcia, M. T. M. and Lopes, E. G., 2009. The macroeconomic impact of reforming a PAYG system: The Portuguese Case. *International Social Security Review,* 62(1), pp. 1-23.
22. Garcia, M. T. M., 2017. Overview of the Portuguese Three Pillar Pension System. *International Advances in Economic Research*, 23, pp. 175-189. https://doi.org/10.1007/s11294-017-9636-x
23. Giles, D. E. and Godwin, R. T., 2012. Testing for multivariate cointegration in the presence of structural breaks: p-values and critical values. *Applied Economics Letters*, 19(16), pp. 1561-1565. https://doi.org/10.1080/13504851.2011.639727
24. Grech, A. G., 2015. Convergence or divergence? How the financial crisis affected European pensioners. *International Social Security Review,* 68(2), pp. 43-61. https://doi.org/10.1111/issr.12065
25. Halmosi, P., 2014. Transformation of the Pension Systems in OECD Countries after the 2008 Crisis. *Public Finance Quarterly*, (4), pp. 457-469.
26. ILO, 2017. *World Social Protection Report 2017-19: Universal social protection to achieve the Sustainable Development Goals.* Geneva: International Labour Office.
27. ILO, 2018. Social protection for older persons: Policy trends and statistics 2017-19. *Social Protection Policy Papers*, No. 17.
28. Johansen, S., Mosconi, R. and Nielsen, B., 2000. Cointegration analysis in the presence of structural breaks in the deterministic trend. *The Econometrics Journal*, 3(2), pp. 216-249.
29. Natali, D. and Stamati, F., 2014. Reassessing South European pensions after the crisis: Evidence from two decades of reforms. *South European Society and Politics,* 19(3), pp. 309-330. https://doi.org/10.1080/13608746.2014.951515
30. Natali, D., 2015. Pension Reform in Europe: What has happened in the Wake of the Crisis? *CESifo DICE Report,* No. 2/2015, pp. 31-35. https://doi.org/10.1111/1368-423X.00047
31. OECD, 2014. *OECD Economic Surveys: Portugal 2014*. Paris: OECD. https://doi.org/10.1787/eco_surveys-prt-2014-en
32. OECD, 2015. *Pensions at a Glance 2015*. Paris: OECD. https://doi.org/10.1787/eco_surveys-prt-2014-en
33. OECD, 2019. *OECD Reviews of Pension Systems: Portugal, OECD Reviews of Pension Systems*. Paris: OECD. https://doi.org/10.1787/9789264313736-en
34. OECD/Inter-American Development Bank/The World Bank, 2014. Policy issues: Coverage and adequacy. In: *Pensions at a Glance: Latin America and the Caribbean*. Paris: OECD. https://doi.org/10.1787/9789264224964-5-en
35. Orenstein, M. A., 2011. Pension Privatization in Crisis: Death or Rebirth of a Global Policy Trend? *International Social Security Review*, 64(3), pp. 65-80. https://doi.org/10.1111/j.1468-246X.2011.01403.x

36. Plamondon, P. [et al.], 2002. *Actuarial practice in social security*. Geneva: International Labour Office/International Social Security Association.
37. Ramaswamy, S., 2012. The Sustainability of Pension Schemes. *BIS Working Papers*, No. 368. https://doi.org/10.2139/ssrn.1980094
38. Roach, B. and Ackerman, F., 2005. Securing Social Security: Sensitivity to Economic Assumptions and Analysis of Policy Options. *GDAE Working Paper,* No. 05-03.
39. Wooldridge, J. M., 2009. *Introductory Econometrics, A Modern Approach*, 4th ed. Boston: South-Western.
40. World Bank, 1994. *Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth.* Oxford University Press.
41. World Bank, 2006. *Pension Reform and the Development of Pension Systems: An Evaluation of World Bank Assistance*. Washington, DC: Independent Evaluation Group, World Bank.
42. Zokalj, M., 2016. The impact of population aging on public finance in the European Union. *Financial Theory and Practice*, 40(4), pp. 383-412. https:// doi.org/10.3326/fintp.40.4.2

PUBLIC SECTOR
ECONOMICS
47 (2) 177-203 (2023)

APPENDIX

Table A1
Descriptive statistics

Variables	Mean	Median	Max	Min	Std. dev.	Skewness	Kurtosis	Jarque-Bera
Dependent variable								
Pensions in % of GDP	5.05	5.15	7.70	2.20	1.28	0.16	2.82	0.28 (0.89)
Independent variables								
Log unemployment	12.75	12.72	13.66	12.09	0.38	0.65	3.14	2.82 (0.24)
Log apparent labour productivity	2.69	2.77	3.01	2.18	0.28	-0.53	2.03	3.43 (0.18)
Old age dependency ratio	22.35	22.00	30.70	16.30	4.09	0.33	1.94	2.61 (0.27)
No. of observations					40			

Note: The probability is between brackets.

Source: Authors' computation.

Table A2
Unit root augmented Dickey-Fuller and Phillips-Perron tests

	Dickey-Fuller test			Phillips-Perron test		
	Deterministic component	**p-Value**	**t-Stat**	**Deterministic component**	**p-Value**	**Adj. t-Stat**
Dependent variable						
Pensions in % of GDP	constant and trend	0.44	-2.27	constant and trend	0.29	-2.59
First difference	constant	0	-6.23	constant	0	-6.24
Independent variables						
Old age dependency ratio	constant and trend	0.98	-0.45	constant and trend	0.99	0.41
First difference	constant and trend	0.005	-4.53	constant and trend	0.004	-4.58
Log unemployment	constant and trend	0.39	-2.36	constant and trend	0.69	-1.78
First difference	none	0.000	-3.97	none	0.00	-3.97
Log apparent labour productivity	constant and trend	0.86	-1.33	constant and trend	0.83	-1.46
First difference	constant and trend	0.017	-4.01	constant and trend	0.017	-4.01

Table A3
Var lag order selection criteria procedure

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-86.6845	NA	0.0044	5.9269	6.8066	6.2340
1	212.3587	448.5649	6.82e-10	-9.7977	-8.2142*	-9.2450
2	240.9137	36.4869*	3.73e-10*	-10.4952	-8.2079	-9.6969*
3	257.5250	17.5341	4.35e-10	-10.5292	-7.5381	-9.4852
4	278.0591	17.1118	4.79e-10	-10.7811*	-7.0862	-9.4914

** Indicates lag order selected by the criterion.*

Endogenous variables: Pensions in percentage of gross domestic product Log unemployment Log apparent labour productivity Old age dependency ratio

Exogenous variables: REV1974 R1984 R1993 R2002 R2007

Sample: 1975 2014; Included observations: 36

AIC: Akaike information criterion; SC: Schwarz information criterion; LR: sequential modified; LR test statistic (each test at 5% level); FPE: Final prediction error; HQ: Hannan-Quinn information criterion.

Table A4
Johansen co-integration test summary

Data Trend:	**None**	**None**	**Linear**	**Linear**	**Quadratic**
Test Type	No Intercept	Intercept	Intercept	Intercept	Intercept
	No Trend	No Trend	No Trend	Trend	Trend
Trace	0	1	1	1	1
Max-Eig	0	1	1	1	1
*Critical values based on MacKinnon-Haug-Michelis (1999)					
Information Criteria by Rank and Model					
Data Trend:	**None**	**None**	**Linear**	**Linear**	**Quadratic**
Rank or	No Intercept	Intercept	Intercept	Intercept	Intercept
No. of CEs	No Trend	No Trend	No Trend	Trend	Trend
Log Likelihood by Rank (rows) and Model (columns)					
0	192.5650	192.5650	195.6342	195.6342	199.7912
1	199.4257	208.0199	210.6990	212.6689	216.0536
2	204.8373	214.6224	217.2509	223.7600	226.8486
3	207.8471	218.6760	220.1204	228.7651	229.9657
4	210.2947	221.3637	221.3637	231.0177	231.0177
Akaike Information Criteria by Rank (rows) and Model (columns)					
0	-8.6792	-8.6792	-8.6289	-8.6289	-8.6374
1	-8.6176	-9.0281	-9.0108	-9.0632	-9.0840
2	-8.4777	-8.8985	-8.9325	-9.1762	-9.2351*
3	-8.2079	-8.6311	-8.6552	-8.9603	-8.9711
4	-7.9078	-8.2899	-8.2899	-8.5956	-8.5956
Schwarz Criteria by Rank (rows) and Model (columns)					
0	-7.2860*	-7.2860*	-7.0615	-7.0615	-6.8958
1	-6.8761	-7.2430	-7.0951	-7.1040	-6.9941
2	-6.3879	-6.7216	-6.6685	-6.8251	-6.7969
3	-5.7698	-6.0624	-6.0429	-6.2174	-6.1847
4	-5.1214	-5.3293	-5.3293	-5.4608	-5.4608

Sample: 1975 2014

Included observations: 37

Series: Pensions in percentage of gross domestic product Log unemployment Log apparent labour productivity Old age dependency ratio

Lags interval: 1 to 2

Selected (0.05 level) Number of Co-integrating Relations by Model*

PUBLIC SECTOR
ECONOMICS
47 (2) 177-203 (2023)

Table A5
Descriptive statistics – Residuals

Variables	Mean	Median	Max	Min	Std. dev.	Skewness	Kurtosis	Jarque-Bera
Resid01	2.25E-17	-0.007	0.236	-0.191	0.092	0.399	3.113	1.006 (0.605)
Resid02	-3.00E-18	0.015	0.245	-0.220	0.102	0.198	2.897	0.259 (0.879)
Resid03	5.06E-18	-0.001	0.018	-0.021	0.010	-0.065	3.14	0.042 (0.979)
Resid04	2.69	2.77	3.01	2.18	0.28	2.90	2.32	0.775 (0.679)
No. of observations				37				

Note: In parenthesis the probability.

Table A6
White heteroscedasticity test (no cross terms)

Joint test:		
Chi-sq	**df**	**Prob.**
257.1420	250	0.3646

Sample: 1975 2014

Included observations: 37

Table A7
Covariance matrix between variables and residuals

Variables	Resid01	Resid02	Resid03	Resid04
Pensions in % of GDP	0.0022	-0.0009	-6.561464E-05	0.0004
Old age dependency ratio	-0.0026	-0.0062	8.657489E-05	0.0079
Log unemployment	-0.0007	-0.0002	-6.941109E-05	0.0010
Log apparent labor productivity	0.0002	0.0004	6.809052E-05	-0.0003

PUBLIC SECTOR
ECONOMICS
47 (2) 177-203 (2023)

Table A8

Residual normality test

Component	Skewness	Chi-sq	df	Prob.
1	0.3998	0.9858	1	0.3208
2	-0.0396	0.0097	1	0.9217
3	0.2675	0.4412	1	0.5066
4	-0.4303	1.1417	1	0.2853
Joint		2.5784	4	0.6307
Component	**Kurtosis**	**Chi-sq**	**df**	**Prob.**
1	3.1133	0.0198	1	0.8881
2	2.6726	0.1653	1	0.6844
3	2.4715	0.4305	1	0.5117
4	2.7765	0.0770	1	0.7814
Joint		0.6926	4	0.9522
Component	**Jarque-Bera**	**df**	**Prob.**	
1	1.0056	2	0.6048	
2	0.1749	2	0.9163	
3	0.8717	2	0.6467	
4	1.2187	2	0.5437	
Joint	3.2710	8	0.9162	

Orthogonalisation: Cholesky (Lutkepohl)

Null Hypothesis: residuals are multivariate normal

Sample: 1975 to 2014

Included observations: 37

Table A9

Residual serial correlation LM test

Lags	LM-Stat	Prob
1	26.5385	0.0469
2	23.6704	0.0970
3	14.3367	0.5737
4	12.1379	0.7344
5	17.8499	0.3328
6	15.0697	0.5195
Probs from chi-square with 16 df.		

Null Hypothesis: no serial correlation at lag order h

Sample: 1975 2014

Included observations: 37

Table A10

Johansen Co-integration Test without dummy variables

Unrestricted Co-integration Rank Test (Trace)

Hypothesized no. of CE(s)	**Eigenvalue**	**Trace Statistic**	**0.05 Critical Value**	**Prob.****
None *	0.5848	62.4530	55.2458	0.0102
At most 1	0.4421	29.9281	35.0109	0.1580
At most 2	0.1551	8.3383	18.3977	0.6481
At most 3	0.0553	2.1040	3.8415	0.1469

Trace test indicates 1 co-integrating eqn(s) at the 0.05 level
* Denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Co-integration Rank Test (Maximum Eigenvalue)

Hypothesized no. of CE(s)	**Eigenvalue**	**Max-Eigen**	**0.05**	
		Statistic	**Critical Value**	**Prob.****
None *	0.5848	32.5249	30.8151	0.0306
At most 1	0.4421	21.5898	24.2520	0.1082
At most 2	0.1551	6.2343	17.1477	0.7936
At most 3	0.0553	2.1040	3.8415	0.1469

Max-eigenvalue test indicates 1 co-integrating eqn(s) at the 0.05 level
* Denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Co-integrating Coefficients (normalised by b'*S11*b=I):

Pensions in percentage of GDP	**Log unemployment**	**Log apparent labor productivity**	**Old age dependency ratio**	
6.4595	-8.5289	-11.7489	1.4318	
1.6370	-6.7668	-37.7933	-1.0975	
6.4758	-3.6887	-25.9968	-0.8533	
-1.8548	3.5122	20.2981	-2.5845	

Unrestricted Adjustment Coefficients (alpha):

D(Pensions in percentage of GDP)	-0.0493	0.0088	-0.0167	-0.0274
D(Log unemployment)	0.0496	0.0308	-0.0121	-0.0156
D(Log apparent labor productivity)	-0.0074	0.0067	0.0009	0.0012
D(Old age dependency ratio)	-0.0292	0.0144	0.0329	-0.0019

1 Co-integrating Equation(s): Log likelihood 216.0536

Normalized co-integrating coefficients (standard error in brackets)

Pensions in percentage of GDP	**Log unemployment**	**Log apparent labor productivity**	**Old age dependency ratio**	
1.000000	-1.3204 (0.163)	-1.8189 (0.936)	0.2217 (0.082)	

Adjustment coefficients (standard error in brackets)

D(Pensions in percentage of GDP)	-0.3182 (0.1666)			
D(Log unemployment)	0.3206 (0.1218)			

PUBLIC SECTOR ECONOMICS
47 (2) 177-203 (2023)

Adjustment coefficients (standard error in brackets)					
D(Log apparent labor productivity)	-0.0477 (0.0165)				
D(Old age dependency ratio)	-0.1883 (0.1141)				

Source: Authors' computation.

Sample (adjusted): 1978 2014

Included observations: 37 after adjustments

Trend assumption: Quadratic deterministic trend

Series: Pensions in percentage of gross domestic product Log unemployment Log apparent labour productivity Old age dependency ratio

Lags interval (in first differences): 1 to 2

Table A11

Variance for the Pensions in percentage of gross domestic product residuals

Years ahead	Pensions in percentage of GDP	Log unemployment	Log apparent labour productivity	Old age dependency ratio	St. errors
1	39.76	57.39	0.00	2.86	0.12
2	14.09	81.98	1.80	2.13	0.21
3	6.51	86.96	4.92	1.61	0.30
4	4.12	87.13	5.94	2.81	0.39
5	3.15	86.15	6.97	3.73	0.45
6	2.68	85.39	7.33	4.59	0.49
7	2.41	85.04	7.55	5.00	0.52
8	2.19	85.06	7.56	5.18	0.54
9	2.02	85.24	7.58	5.15	0.57
10	1.84	85.49	7.57	5.10	0.60
11	1.67	85.65	7.61	5.07	0.63
12	1.52	85.73	7.66	5.09	0.66
13	1.40	85.72	7.72	5.15	0.69
14	1.31	85.71	7.77	5.22	0.71
15	1.23	85.70	7.80	5.27	0.73
16	1.16	85.72	7.82	5.30	0.76
17	1.09	85.75	7.83	5.32	0.78
18	1.04	85.79	7.85	5.33	0.80
19	0.98	85.81	7.87	5.34	0.82
20	0.93	85.83	7.88	5.35	0.84

Table A12

Johansen Co-integration Test with Dummy Variables

Hypothesized no. of CE(s)	**Eigenvalue**	**Trace Statistic**	**0.05 Critical value**	**Prob.****
None *	0.7395	84.7403	55.2458	0.0000
At most 1	0.4658	34.9636	35.0109	0.0506
At most 2	0.2490	11.7627	18.3977	0.3270
At most 3	0.0311	1.1681	3.8415	0.2798
Trace test indicates 1 co-integrating eqn(s) at the 0.05 level				
*Denotes rejection of the hypothesis at the 0.05 level				
**MacKinnon-Haug-Michelis (1999) p-values				
Unrestricted Co-integration Rank Test (Maximum Eigenvalue)				
Hypothesised no. of CE(s)	**Eigenvalue**	**Max-Eigen Statistic**	**0.05 Critical value**	**Prob.****
None *	0.7395	49.7768	30.8151	0.0001
At most 1	0.4658	23.2009	24.2520	0.0684
At most 2	0.2490	10.5945	17.1477	0.3447
At most 3	0.0311	1.1681	3.8415	0.2798
Max-eigenvalue test indicates 1 co-integrating eqn(s) at the 0.05 level				
*Denotes rejection of the hypothesis at the 0.05 level				
**MacKinnon-Haug-Michelis (1999) p-values				
Unrestricted Co-integrating Coefficients (normalised by b'*S11*b=I):				
Pensions in percentage of GDP	**Log unemployment**	**Log apparent labour productivity**	**Old age dependency ratio**	
12.693	-11.8587	-43.8040	-1.4480	
-1.6522	-4.7928	1.2328	-0.8655	
6.8935	-12.1842	-77.0428	3.6249	
-2.6401	3.0243	51.6066	6.0229	
Unrestricted Adjustment Co-efficients (alpha):				
D(Pensions in percentage of GDP)	-0.0649	0.0490	0.0171	0.0036
D(Log unemployment)	0.0280	0.0668	0.0094	0.0001
D(Log apparent labour productivity L)	-0.0067	-0.0010	-0.0013	-0.0013
D(OAD)	-0.0181	-0.0224	0.0240	-0.0069
1 Co-integrating Equation(s):		Log likelihood	259.8113	
Normalised co-integrating coefficients (standard error in brackets)				
Pensions in percentage of GDP	**Log unemployment**	**Log apparent labour productivity**	**Old age dependency ratio**	
1.000000	-0.9342	-3.4509	-0.1141	
	(0.0849)	(0.6157)	(0.0736)	
Adjustment coefficients (standard error in brackets)				
D(Pensions in percentage of GDP)	-0.8237			
	(0.2515)			
D(Log unemployment)	0.3550			
	(0.2799)			

Adjustment coefficients (standard error in brackets)				
D(Log apparent labour productivity)	-0.0850			
	(0.0245)			
D(Old age dependency ratio)	-0.2292			
	(0.1965)			

Sample (adjusted): 1978 2014

Included observations: 37 after adjustments

Trend assumption: Quadratic deterministic trend

Series: Pensions in percentage of gross domestic product Log unemployment Log apparent labour productivity Old age dependency ratio

Exogenous series: REV1974 R1984 R1993 R2002 R2007

Warning: Critical values assume no exogenous series

Lags interval (in first differences): 1 to 2

Unrestricted Co-integration Rank Test (Trace)

Table A13

Wald test for the VECM short-run coefficients

Test Statistic	Value	df	Probability
F-statistic	1.3647	(2, 21)	0.2772
Chi-square	2.7294	2	0.2555
Null hypothesis: C(4)=C(5)=0			
Null hypothesis summary:			
Normalised restriction (= 0)		**Value**	**Std. err.**
C(4)		0.4020	0.2435
C(5)		-0.0061	0.2421
All restrictions are linear in coefficients.			
Test Statistic	**Value**	**df**	**Probability**
F-statistic	0.0660	(2, 21)	0.9363
Chi-square	0.1320	2	0.9361
Null Hypothesis: C(6)=C(7)=0			
Null Hypothesis Summary:			
Normalised Restriction (= 0)		**Value**	**Std. err.**
C(6)		-0.2214	2.3580
C(7)		0.5801	1.6070
Test Statistic	**Value**	**df**	**Probability**
F-statistic	1.8520	(2, 21)	0.1817
Chi-square	3.7040	2	0.1569
Null Hypothesis: C(8)=C(9)=0			
Null Hypothesis Summary:			
Normalised Restriction (= 0)		**Value**	**Std. err.**
C(8)		0.1703	0.2614
C(9)		0.3714	0.2094

Leading indicators of financial stress in Croatia: a regime switching approach

TIHANA ŠKRINJARIĆ, Ph.D.[*]

Article[**]
JEL: C54, G01, G15
https://doi.org/10.3326/pse.47.2.3

[*] The views expressed are those of the author and do not necessarily reflect the official views of the institution the author works at. The author is grateful to two anonymous referees who have contributed to the quality of the final version of the paper. The paper was written when the author was at Croatian National Bank, Department for Financial Stability.

[**] Received: June 1, 2022
Accepted: November 21, 2022

Tihana ŠKRINJARIĆ
Bank of England, Stress Test Strategy Division, 20 Moorgate, London, EC2R 6DA, United Kingdom
e-mail: tihana.skrinjaric@bankofengland.co.uk
ORCiD: 0000-0002-9310-6853

PUBLIC SECTOR
ECONOMICS
47 (2) 205-232 (2023)

Abstract

This research focuses on the prediction of the probability of (re)entering high financial stress (via a large set of cyclical risk accumulation indicators). The focus is placed on a specific single-country analysis to obtain answers to questions about which indicators are best in explaining the future probability of (re)entering a high-stress regime. This allows the policymaker to get a better focus on the best-performing variables. It is challenging to monitor a whole set of indicators of cyclical risk build-up; the results could bring into focus a smaller group of the essential variables. The contribution of this paper is in finding a set of indicators that help in forecasting financial stress, in terms of switching from one regime to another. The regime-switching models' results indicate that some credit specifications, house price dynamics, and debt burden could be best monitored for the case of Croatian data.

Keywords: financial stress, macro-prudential policy, Regime-switching models, Croatia

1 INTRODUCTION

Macroprudential policy has the difficult task of identifying sources of financial system instability alongside detecting appropriate indicators to measure the accumulation of systemic risks. Timely and high-quality measures and instruments can then be devised and put into action so that the potential risks are mitigated, and the adverse effects that risk materialization causes to the rest of the real economy are reduced. That is why much research today focuses on quantifying the build-up of vulnerabilities in the financial system, as well as on quantifying risk materialization in the forms of financial stress that occurs in the system itself. One goal is to predict the turning points of a financial cycle or financial crisis and another is to measure the state of the financial system's instability. The amount of literature that focuses on one or the other important aspects has grown in the last decade as macroprudential policy has developed. However, there is still a lack of research linking the build-up of vulnerabilities to the risk materialization approach. This paper tries to fill that gap. The research has endeavoured to combine the two approaches in order to try to evaluate the potential predictability of risk materialization measured via the financial stress indicator, based on the indicators that are preferred in EWS (early warning system) models. The first paper to attempt this was that of Duprey and Klaus (2017), in which a potential is seen in combining the best of both approaches. Thus, this research applies the regime-switching methodology of modelling the financial stress indicator for Croatia, to evaluate predictive possibilities of a number of indicators usually used in the EWS approach to the build-up of financial system vulnerabilities.

The main contribution of this study is in its identification of finding a set of indicators that help forecast financial stress, switching from one regime to another, while utilizing techniques to reduce their number. This study observes many different variables and their transformations in order to find the best ones. This allows the policymaker to focus on the variables that were found to be best in the modelling process. It is challenging to monitor a whole set of indicators of cyclical risk build-up but the results of this research could bring into clear focus a smaller group of essential variables. Furthermore, most existing research focuses on predicting the value of future stress and concludes that it is challenging to predict stress levels. However, as Christensen and Li (2014) point out, the decision-making process should rely on point forecasts and insights into the likelihood of the occurrence of the stress event. That is why this study focuses on estimating the effects of the dynamics of indicators on the future probability of entering a stress event, which is rarely found in the literature; to the knowledge of the author, the only other existing study framed in such terms is Duprey and Klaus (2017; 2022). Furthermore, an extensive analysis is made of over several hundred variants of indicators of cyclical risk accumulation. In that way, the quality of regime-switching models is examined in a fashion that has never been accomplished previously (to the author's knowledge). Namely, the usual criteria used in the model comparison are contrasted over six categories of measures, for four different lags of the indicators, with a particular focus on the empirical distributions of the results. This enables the policymaker to determine the characteristics of each category for every lag. Moreover, those individual variables that were found to be the best in adding relevant information to the model are used in a simple simulation, involving the construction of a composite indicator of risk accumulation that can be used to monitor financial cycle dynamics.

There are a couple of reasons why we focus on the regime-switching approach: Croatian data, and the combination of both approaches. First, regime-switching models have been extensively used and found to be useful in modelling business cycles (see, e.g. Doz, Ferrara and Pionnier, 2020) and financial markets (Ang and Timmermann, 2012)[1]. Furthermore, research has found that the interaction between financial stress and the real economy is not linear. Quite the opposite, the relationship is asymmetric and exhibits nonlinear behaviour, as seen in Giglio, Bryan and Pruitt (2016), Cardarelli, Elekdag and Lall (2011), or O'Brien and Wosser (2021). Moreover, Vermeulen et al. (2015) show that spikes in financial stress occur abruptly. Finally, Misina and Tkacz (2009) think that nonlinear models are more suitable for capturing the asymmetry in the behaviour of financial market participants. Thus, regime-switching models capture the nonlinearities and asymmetric behaviour that cannot be adequately modelled via linear models. Another important issue is that forecasting future crises does not rely solely on the value of a variable (i.e. the value of the stress indicator itself). Still, the probability of crisis occurrence is also important (Gneitinga and Ranjanb, 2011) and it is more important when making policy decisions. Next, Croatian data are observed as this area of macroprudential policy practice is not sufficiently explored for this

[1] This is due to the intuitive interpretation of dynamics and characteristics of economic and financial variables via expansions and contractions in real economy or bull and bear markets in financial markets (see Ang and Timmermann, 2012). Structural and abrupt changes in the economy and/or financial markets due to political issues, legislative changes, new methodological approaches, etc. can be captured via the regime-switching approach (Baele, 2003 is an example of the economic and monetary integration of Western Europe affecting the European financial markets).

PUBLIC SECTOR
ECONOMICS
47 (2) 205-232 (2023)

country. On the other hand, the country has known an extensive macroprudential approach over the last two decades[2]. There exist just a few related papers regarding the Croatian financial system: Dumičić (2015a), where the financial stress indicator is constructed for the first time for the Croatian case; Duraković (2021) upgraded the stress indicator in the approach of Holló, Kremer and Lo Duca (2012) and its interaction with the real economy was tested and found to be significant and asymmetric; and Dumičić (2015b), where the first attempt at constructing a composite indicator of cyclical systemic risk accumulation was made. Countries with an economy similar to the Croatian could thus acquire some insight into potential outcomes regarding their financial systems.

Although researchers are often prejudiced against single country analysis, there is actually a potential in focusing on one system, as compared to panel data analysis. For example, Klomp (2010) examined 110 different countries and found heterogeneity in the causes of banking crises, alongside an overview of previous related research that found different variables to be significant in predicting future crises, alongside different estimated signs. Although panel analyses are very important due to their advantages over single-country analysis, countries cannot copy measures and instruments one from another directly due to country specificities. In fact, previous research that focused on early warning systems of crises has often incorporated country-specific analysis: Kaminsky and Reinhart (1999) estimate country-specific thresholds in minimizing the noise-to-signal ratio in the modelling process. There are significant differences between the country-specific and global thresholds in the EWS results found in Davis and Karim (2008), who, focusing on banking crises conclude that generalized global models cannot replace country-specific macroprudential surveillance.

2 LITERATURE REVIEW

The first group of research that focuses on the build-up of vulnerabilities in the financial system is focused on exploring the possibilities of predicting turning points of the financial cycle. The rise of interest in forecasting financial crises has risen especially after the GFC (Global Financial Crisis), due to the severe consequences it entailed for the rest of the real economy. Early-warning systems became very popular, as their purpose is to give an early enough warning of a turning point in the financial cycle (Reinhart and Rogoff, 2009). Here, the idea is to evaluate

[2] Croatia has a specific, i.e., unique experience regarding macroprudential policymaking and monitoring cyclical risk accumulation. It stands out due to it belonging to a group of countries that had the most intensive use of instruments before the global financial crisis (Vujčić and Dumičić, 2016). This means that macroprudential policy was active during the boom and bust phases of the financial cycle. Analysis regarding Croatian data could provide insights into the effects of macroprudential policy during all phases of the cycle, the effects on the financial stress, and other analyses of interactions of this policy with the rest of the economy. As Vujčić and Dumičić (2016) state, Croatian experience shows that policymakers shouldn't focus only on textbook approaches to macroprudential policy conduct but, to keep their minds open, also on analysis such as the one in this study. However, a not-typical one could be one of the starting points. Moreover, Bambulović and Valdec (2020) state that Croatia is an interesting country for a study of the effects of macroprudential policies on credit growth, as the Croatian National Bank employed many measures in the pre-GFC period. More specifically, greater macroprudential activity started in 2003, as it was seen that monetary policy would not be efficient alone (see Kraft and Galac, 2011).

individual or group possibilities in signalling a future crisis of selected variables that should capture the changes in the accumulation of cyclical systemic risk. Empirical work has focused on the six categories of measures of the build-up of system-wide risk as recommended in ESRB (2014). The best predictors of previous crises are closely observed in the practices of central banks to make decisions about macroprudential instruments to mitigate systemic risks, with a special focus on cyclical systemic risk. For example, some central banks utilize from 6 to 35 individual indicators (Arbatli-Saxegaard and Muneer, 2020), with many of them monitoring over 20 indicators when making decisions about the countercyclical capital buffer. Thus, one strain of literature has focused on defining a binary variable depending on the definition of financial, banking, currency, or related crises dates (based on the criteria in ESRB, 2018; ECB, 2017; and Dimova, Kongsamut and Vandenbussche, 2016). The EWS approach is, thus, based on predetermined dates of crisis or non-crisis state in the financial system, and the idea is to evaluate the signalling properties of potential indicators regarding the binary variable defining a vulnerable period before a crisis happens.

The second group of works in the literature concerned with monitoring systemic stress is focused on the composite indicator of financial stress. An influential paper by Holló, Kremer and Lo Duca (2012) has opened a path to measuring the realization of risks in financial markets. Here, the composite indicator indicates the current state of instability of the financial system, together with the frictions and strains affecting it. Holló, Kremer and Lo Duca (2012) state that such a measure is helpful in determining episodes of financial crises. This means that the realizations of the stress captured in the financial stress indicator can be utilized to estimate the crisis and non-crisis states of the system, as well as corresponding dates. This is useful due to previous research having found evidence of the negative effects of financial stress on the economic activity (Borio and Drehmann, 2009; Bloom, 2009; Hubrich and Tetlow, 2015; Chavleishvili and Manganelli, 2019; Dumičić, 2015a; and an overview in Škrinjarić, 2022), which is especially true for the downside risks of economic growth (e.g. Adrian, Boyarchenko and Giannone, 2019; and a general overview in Plagborg-Moller et al., 2020)[3].

Misina and Tkacz (2009) utilize a linear and threshold regression approach in forecasting the FSI (financial stress index) based on values in k (= 1 to 12) periods ahead of selected indicators. The empirical part of their research focused on Canadian data on credit developments, asset prices, GDP, and crude oil prices (various definitions of variables, with quarterly and yearly growth rates). Forecast quality was contrasted among all different variants of considered models, with horizons of $k = 4$ and 8 being the best for credit and asset price variables. Another single-country analysis is found in Hanschel and Monnin (2005), where the authors model stress in the banking sector

[3] That is why forecasting future financial stress would be helpful in order to adjust the instruments and measures so that the financial system is more robust and stable over time, alongside reducing the costs of financial crises. Moreover, financial stress indicators are used in the quick release of the countercyclical capital buffers (CCyB), as at the beginning of the COVID-19 crisis in some countries (see Arbatli-Saxegaard and Muneer, 2020).

of Switzerland. 13 OECD countries are examined in Slingenberg and de Haan (2011), in which the authors in a fashion similar to the models in Misina and Tkacz (2009) observe the predictive power of similar variables. The main findings include that credit growth has some predictive power for most of countries, but other variables are significant for some countries and not for others. This also goes in favour of single-country analysis. Lo Duca and Peltonen (2011) define a stress event when the stress variable exceeds its 90-th percentile value in the analysis, and then construct a binary variable depending on this value. The EWS approach was done here as well, in which in total 28 countries were observed. Asset prices and credit developments were found to be the best predictors of future crises. A signal extraction approach was made in Christensen and Li (2014), where 13 OECD countries were observed in understanding the behaviour of selected economic indicators preceding financial stress. However, a financial stress event is defined via a subjective selection of a threshold value that is between some values used in related approaches[4]. Vašíček et al. (2017) observed 25 OECD countries in the BMA (Bayesian model averaging) approach in determining the leading indicators of financial stress. Here, the authors obtain the following results: the panel approach does not yield good results in terms of successful prediction of future FSI movements; country-level analysis produces better results; and in general, FSI movements are hard to predict, as out of sample predictions are worse than in-sample analysis. Pietrzak (2021) has compared several approaches to predicting financial distress for 43 countries, in the period 1990 to 2016. Thus, this is not an analysis methodologically comparable to this one, but the results are interesting, as the results show that specific variables such as capital adequacy, leverage, and return on assets predict future financial distress[5]. Duprey and Klaus (2017; 2022) is the closest study to this one, the authors evaluating 27 indicators for 15 EU countries and in the unbalanced sample from Q1 1970 to 4Q 2018 analysing the potential of the binary-logit model and the regime-switching approach. The main results indicate that the debt service ratio, credit dynamics, and property market variables are useful in predicting higher financial stress periods. The analysis that included post-GFC data was better than the pre-GFC one. Based on these results, the authors concluded that before the GFC, it was more difficult to predict the future movement of the financial stress probabilities.

Dumičić (2015a) has constructed an earlier version of a financial stress indicator (FSI) for Croatia. In the second phase of the research, FSI indicators were used in a regime-switching model to see their performance in boom and bust phases, alongside the sources of risk increases during regimes of more stress. Based on an extensive analysis of higher and lower stress sub-periods, the author concluded that the FSIs defined in this paper are a good starting point for use in practice.

[4] A stress event is defined if the stress indicator exceeds its mean value enlarged by its standard deviation multiplied by a constant *k*. This could be arbitrary, as opposed to the regime-switching that is based on the optimization of the likelihood function in which optimal switching behaviour is governed by the data. This is what the authors concluded at the end of the study: future research should focus on switching models.

[5] A drawback of the study is found in determining the stress periods when the financial condition indicator exceeds the 90th percentile of the country's indicator distribution. The regime-switching approach does not ask for such interventions from the researcher's side.

PUBLIC SECTOR
ECONOMICS
47 (2) 205-232 (2023)

Duraković (2021) is a newer variant of a composite indicator of systemic stress. The author followed the Holló, Kremer and Lo Duca (2012) approach in constructing the indicator with some specifics for Croatian data. In this work, the author observed the threshold VAR model in estimating the effects of higher or lower financial stress on real economy growth. Asymmetric findings confirm the aforementioned nonlinearities in the introduction. Compared to the previous two papers, this paper looks beyond current realizations of financial stress in the financial markets in Croatia. Here, we observe the information obtained from MRS models regarding the effects of cyclical risk indicators that could have affected switching from one regime to another. This gives the policymaker some insights into the possibility of affecting those indicators via certain measures or instruments, which could, in turn, reduce future financial stress.

Generally speaking, after reviewing the related literature, it seems reasonable to analyse the effects of early warning indicators on the probability of future higher or lower stress, as the value of the stress variable is hard to predict. However, the selected indicators were found useful in the early warning literature, regarding the signalling approach to future crisis modelling. Variables utilized in this study are based on the literature within this group of papers.

3 DATA AND METHODOLOGY DESCRIPTION

3.1 DATA DESCRIPTION

Quarterly data on the following six (according to the ESRB (2014) Recommendation) categories of measures of cyclical risk build-up have been collected from CNB (2022): credit dynamics, potential overvaluation of property prices, external imbalances, the strength of bank balance sheets, private sector debt burden and potential mispricing of risks. As the availability of data varies depending on the variable category (or even within the category itself), to produce comparability of the estimation results, all of the variables after the initial transformations were brought to the same initial starting point of 4Q 2005. The last available data are for 3Q 2021. Although this results in a short timespan dataset, future work should focus on this issue. Besides the indicators, values on HIFS[6] (Croatian indicator of financial stress) have been collected from internal CNB calculations, and quarterly averages were calculated for the HIFS levels. This indicator is in line with the methodology of Holló, Kremer and Lo Duca (2012) and is described in detail in Duraković (2021). Table 1 gives an overview and description of the variables used in the study alongside the data transformations. Every category consists of transformed variables in terms of one-year growth rates or differences, annualized growth rates or changes, and the HP (Hodrick-Prescott, 1997) gap.

The smoothing parameter values in the filtering were based on previous studies of the length of the financial cycle. For example, Galán (2019) finds that the value of

[6] Calculation of HIFS is based on volatilities and similar measures and correlations derived from daily data on bond yields, stock market returns, money market interest rates, exchange rates, and interbank market rates and fragility. It does not include any of the individual indicators (or their transformations) used in MRS modelling.

25.600 for the smoothing of the credit-to-GDP gap parameter is better in the signalling the future crisis, whereas Wezel (2019) found that CEE (Central and Eastern Europe) countries have shorter credit cycles, which is also in favour of a smaller smoothing parameter. Furthermore, other studies utilize the values of 85,000, 125,000, and 400,000 for the smoothing parameter in the HP filtering, as the length of the financial cycle is assumed to be 2 to 4 times longer than the business cycle: Drehmann et al. (2010) focused on OECD countries and higher values of the parameter, whereas Galati et al. (2016) for 5 eurozone countries conclude that their financial cycle lasts between 8 and 25 years. Moreover, Schüler (2018), in his empirical analysis, concludes that the value of 400.000 for the smoothing parameter in the ESRB (2014) Recommendation could result in omitting relevant fluctuations in relevant financial series. Thus, based on these mixed results, this study allows for different smoothing parameters in the HP filtering process.

Furthermore, it should be noted that the decision-making process of the policymaker is in real-time, meaning that all of the filters are calculated recursively, i.e. as one-sided gaps. Smaller values of the smoothing parameters are related to the assumption of a shorter length of the financial cycle or the fact that some variables could be linked more to the business cycle (parameter 1,600). In total, 241 different variables are examined in the study. All variables are interpreted in such a way that the greater the value, the greater the risk accumulation. This means that some variables were multiplied with value -1 to provide such interpretation (e.g. interest rate margins, net exports, etc.).

Table 1

Brief description of variables used in the study

Category	Variables	Transformations		
Credit dynamics	Narrow and broad credit Narrow and broad credit-to-GDP ratio Household (HH) credit Non-financial corporations (NFC) credit HH credit-to-GDP ratio NFC credit-to-GDP ratio (nominal and real variants of credit)	1YG 1YC	A2YG A2YC	HP gap 25K 85K 125K 400K
Potential overvaluation of property prices	House price index (nominal and real) Price to rent ratio (nominal and real) Price to income ratio (nominal and real) Price to cost ratio (nominal and real) Construction work index ratio Indicator of overvaluation of property prices	1YG 1YC	A2YG A2YC	HP gap 1600 25K 85K 125K 400K
External imbalances	Gross external debt (GED) Net external debt (NED) GED-to-GDP ratio NED-to-GDP ratio Terms of trade Current account to GDP ratio Net export to GDP ratio	1YG 1YC	A2YG A2YC	HP gap 1600 25K 85K 125K 400K

Category	Variables	Transformations		
Strength of bank balance sheets	Deposit to credit ratio Capital to assets ratio Assets to GDP ratio	1YG 1YC	A2YG A2YC	HP gap 1600 25K 85K 125K 400K
Private sector debt burden	Debt service ratio (household and nonfinancial corporations separately) Total debt to income ratio Total debt to gross operating surplus ratio	1YG 1YC	A2YG A2YC	HP gap 25K 85K 125K 400K
Mispricing of risks	Stock market index CROBEX HH credit interest rate margin[7] NFC credit interest rate margin Bank prudence indicator[8]	1YG 1YC	A2YG A2YC	HP gap 1600 25K 85K

Note: HP denotes the Hodrick-Prescott filter, 25K, 85K, 125K, and 400K denote that chosen smoothing parameter was equal to 25.600, 85.000, 125.000, and 400.000 in the HP filter, 1600 denotes the value of the smoothing parameter in HP filter. Narrow credit includes bank credits to households and the private sector, broad credit includes all credit institution claims on the private sector and the gross external debt of the private sector. 1YG, 1YC, A2YG, and A2YC denote one-year growth rate, one-year change, annualized 2-year growth rate, and annualized 2-year change respectively.

Source: CNB (2022).

3.2 METHODOLOGY OF REGIME SWITCHING

The basic approach of the present research is to observe how different measures of risk accumulation affect the future probability of transitioning from one regime to another regarding the financial stress variable. Thus, the Markov regime-switching models (MRS) are the most appropriate, and we follow the usual literature in the brief description of MRS (Gray, 1996; Kim, 1994; Franses and van Dijk, 2000). It is assumed that the financial stress variable (HIFS) is defined via the following data generating process (DGP):

$$HIFS_t = a_i + \varphi HIFS_{t-1} + \varepsilon_{i,t}, i \in \{1,2\}, \varepsilon_{i,t} \sim \mathrm{N}(0,\sigma_i^2) \tag{1}$$

where the average value and the variance are regime-dependent. Other specifications in (1) can be assumed, depending on the best DGP of the empirical data. For example, in the empirical part of the research, before the estimation is made, the Box-Jenkins approach will be employed to determine the best ARMA(p,q) model. The probability transition matrix between the two assumed regimes (low and high stress) is defined as follows:

$$P(S_t \mid S_{t-1}, X_{t-4}) = \begin{bmatrix} p_{11} & 1-p_{11} \\ 1-p_{22} & p_{22} \end{bmatrix} \tag{2}$$

[7] Margins were calculated as the difference between the interest rates on credits to HH or NFC and the Euribor interest rate, as the national referent interest rate ceased to exist in 2020. Comparable studies (e.g. Kupkovič and Šuster, 2020) also use Euribor as the referent interest rate.
[8] Defined in Pfeifer and Hodula (2018).

PUBLIC SECTOR
ECONOMICS
47 (2) 205-232 (2023)

where

$$p_{ij}(t) = \frac{\exp(X'_{t-4}\beta_{ij})}{\sum_{s=1}^{2}\exp(X'_{t-4}\beta_{is})} \tag{3}$$

is the one-step-ahead probability of entering regime *j* in period *t* from regime *i* in period *t*-1, X_{t-4} is a vector of variables that determine the switching behaviour between the two regimes, as in Duprey and Klaus (2017; 2022). In practice, *X* will be one variable as an indicator in testing its predictive probability of future shifts between the low and high regime of the financial stress variable. The selection of lag 4 quarters in practice for the variable *X* is explained in Duprey and Klaus (2017; 2022) as being consistent with data that the financial stress, conditional on the shift in regimes, does not depend on X_{t-4} beyond the information contained in the shift in the regime. Moreover, it can be added that the variables used as *X* are usually published with lags. This means that using *t*-1 or *t*-2 periods ahead does not make sense, as the stress realization in quarter *t* would be documented before the data on X_{t-1} or X_{t-2} were available. On the other hand, the idea of using variables in *X* is founded on the early warning model (EWM) literature, which finds that such variables should signal the future turning point of the financial crises or cycle from 12 to 5 quarters ahead (ECB, 2017). Longer horizons such as 12 quarters ahead are proven to have low forecasting power (see Misina and Tkascz, 2009). Finally, short time series such as the Croatian data cannot afford to observe long horizons, due to data loss at the beginning of the sample. Vašíček et al. (2017) add that there needs to be a balance between the informative criterion (information provided in a variable declines with a longer forecast horizon) and the criterion of allowing for timely policy action. Finally, several other lags will be considered in the empirical analysis, *t*-5, *t*-6 and *t*-7.

When estimating the switching model, we will observe the estimated signs and significance of betas in equation (3). If we denote regime 1 as being the more stressful (a greater value of the variance as well as a greater average value of the stress series), positive significant betas corresponding to the indicator variables that are tested in the analysis indicate that their greater values increase the probability of transitioning from the less risky to riskier regime; and increasing the probability of staying in the same riskier regime as well. The model is estimated via the usual approach of the likelihood maximization of the function where the *HIFS* value is estimated with respect to given information *I* in the previous period, alongside the information about being in regime *i*, to estimate unknown parameters in θ $f(HIFS_t \mid I_{t-1}, \Delta_t = i; \theta) = P(HIFS_t \mid I_{t-1}, \Delta_t = i; \theta)$, where $\theta = (\beta(\Delta), p_{ij}, \sigma_i^2)$. The log likelihood function is defined as the sum of log values of probabilities $P(HIFS_t \mid I_{t-1}, \Delta_t = i; \theta)$ for every regime. As it is unknown what the true regime is in period *t*-1, true probabilities to be estimated are unknown. Thus, the inferential and smoothed probabilities are estimated recursively as in the definition of Gray (1996). It should be noted that the decision-making process is in real time, which means that smoothed probabilities are not available in such a scenario. Thus, the macroprudential policymaker is observing the one-step-ahead probability and eventually filtered.

4 EMPIRICAL RESULTS

4.1 INITIAL RESULTS

Firstly, the Box-Jenkins approach was made in determining the adequate ARMA(p,q) model for the HIFS variable behaviour. Table A1 in the appendix depicts the Schwartz information criterion table, where the AR(2) model is chosen as the best one. Thus, the rest of the analysis will be made according to this model specification. Figure 1 depicts the dynamics of the stress variable HIFS in the observed period (solid black curve), with the one-step-ahead and filtered probabilities of being in the regime of higher stress (grey curves). These probabilities are estimated according to a two-regime model, in which it is assumed that the average value of the HIFS and the variance of the error term are regime dependent. The AR terms are assumed to be non-regime dependent. The corresponding estimated model is shown in table 2 and it will be the basis for comparisons with its extensions with the risk build-up variables. Other model specifications were estimated and contrasted against the one given in table 2, in terms of the information criterion, log-likelihood, and Wald test for parameter equality over the regimes. It was found that this is the best specification of the DGP of HIFS if it is assumed that it follows a regime-switching process. Ang and Bekaert (2002a; 2002b) defined RCM (regime classification measure) to quantify how well the regime-switching model classifies the regimes. It is calculated as follows:

$$RCM = 100K^2 \frac{1}{T} \sum_{t=1}^{T} \prod_{i=1}^{K} p_{it} \tag{4}$$

where K is the number of regimes in the model, T is the number of observations, p_{it} is the smoothed probability of being in regime i at time t. The value is between 0 and 100, and the closer to zero it is, the better the classification of the regimes. The value is given in table 2, alongside other results.

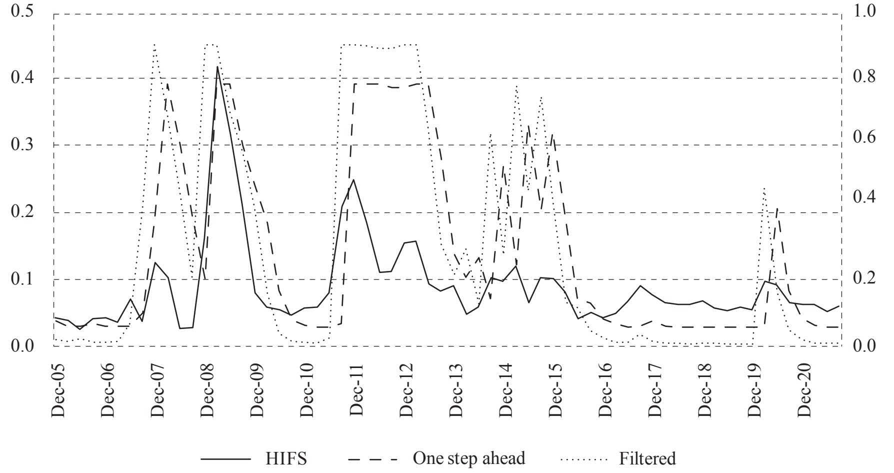
Firstly, figure 1 indicates that the regime-switching approach successfully captures the dynamics in the stress variable. The reactions of financial markets when the GFC, the EU sovereign debt crisis hit, and other impactful events, such as the COVID-19 crisis are pronounced in the dynamics. Moreover, the probability of being in the regime of higher stress is greater for the two aforementioned crises, whereas turbulences in 2014-2015 and the COVID-19 crisis were characterized by lower values of the high-stress probability. This indicates that the regime-switching approach is appropriate for such analysis[9]. The results in table 2 also provide insights into the baseline model: the RCM measure is very low, the values of the average HIFS value and its volatility are fairly different and significant in both regimes.

[9] The results in figure 1 are comparable to those in Dumičić (2015a) in sub-periods of greater and lower stress. Furthermore, based on the discussion in section 4.3., in a comparison of the results to the discussion in the mentioned paper, findings on indicators in table 3 of this research confirm the importance of these indicators for macroprudential policy-making, alongside the measures that the CNB conducted in the observed sample. For more details, please see section 4.3.

PUBLIC SECTOR
ECONOMICS
47 (2) 205-232 (2023)

Figure 1

HIFS values (left hand side) and the one-step-ahead and filtered probability of being in the regime of higher stress (right hand side)



Source: CNB (2022) and author's calculation (for probabilities).

Table 2

Regime switching estimation results for a 2-regime AR(2) model for HIFS

Parameter/values	Regime 1	Regime 2
Constant	0.1718 (0.014)***	0.0657 (0.006)***
Variance	0.0081 (0.005)*	0.0004 (0.000)***
β_{ij}	0.7912 (0.149)***	0.0385 (0.027)
AR(1)	0.7351 (0.110)***	
AR(2)	-0.2186 (0.102)**	
AIC	-247.566	
Log L	131.873	
RCM	14.776	

Note: β_{ij} *is the estimated parameter for constant governing the regime switching in eq. (3).* AIC *and* Log L *denote the Akaike information criterion and* log likelihood *value respectively. Values in parentheses denote standard error of the estimator.*

, ** and * denote statistical significance at 10%, 5% and 1% respectively.*

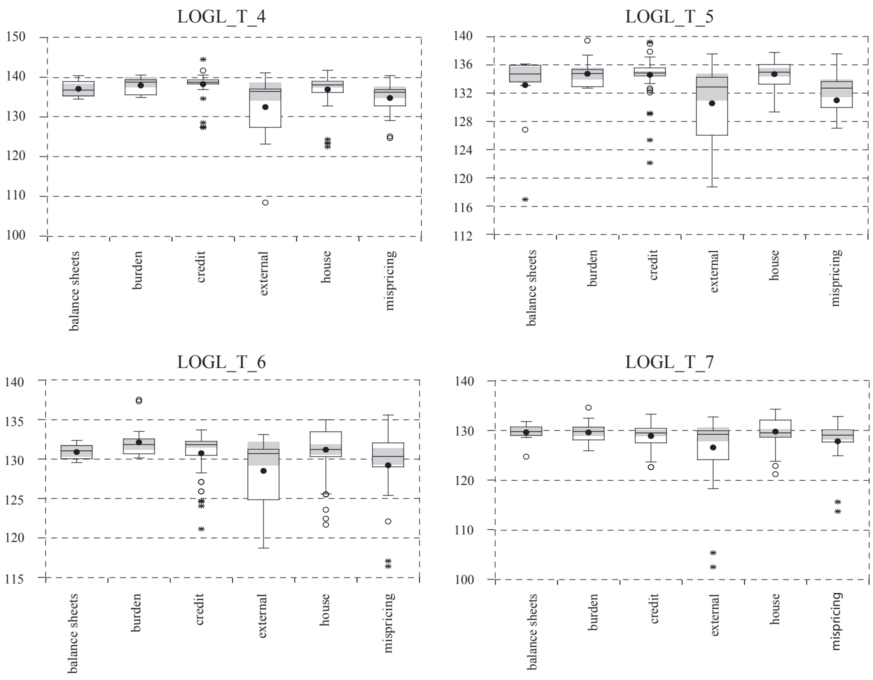
Source: Author's calculation.

4.2 INCLUSION OF INDICATOR VARIABLES RESULTS

Next, the baseline model is challenged by including the variables of cyclical risk accumulation in equation (3) of the model to govern the one-step-ahead probability of entering the risker regime (or staying in it). Thus, for every variable in the previous section, the 2-regime AR(2) model is estimated, including the single variable (in some of its transformations from table 1) with a lag of 4, 5, 6, and 7 quarters. That is, 964 models were estimated in total. As not all variables are comparable in terms of their unit of measurement, we first compare the log likelihoods and AIC information criterion across all variables, lags, and their category according to table 1. Figure 2 depicts the box-plots for the mentioned variations for the log likelihood values, and figure 3 does the same, but for the AIC values[10]. It can be seen that the value of Log L increases as the lag of the indicator variable shortens across all of the categories of measures. This is in line with the EWS research (ESRB, 2018; Lang et al., 2019; Alessi and Detken, 2019; Candelon, Dumitrescu and Hurlin, 2012), where the indicators should show a consistent increase in their value before the peak of the financial cycle occurs and the risk materializes. A more detailed depiction of the values from figure 2 is given in figure A1 in the appendix. A detailed picture of the Log L values across all individual indicators in the regime-switching specification. The best specification is for the *t*-4 lag case, as the values are greater than the baseline model in most of the cases, and the variability is the smallest for most of the measure categories.

Figure 2

Box plot for log likelihood values across all 964 models



Note: Shaded areas denote 95% confidence interval for the median value. "balance sheets", "burden", "credit", "external", "house" and "mispricing" denote the six categories of measures from table 1: strength of balance sheets, private sector debt burden, credit dynamics, external imbalances, potential overvaluation of house prices and mispricing of risks respectively. t_4, t_5, t_6 and t_7 denote models that include indicator variables lagged for 4, 5, 6 and 7 quarters respectively.

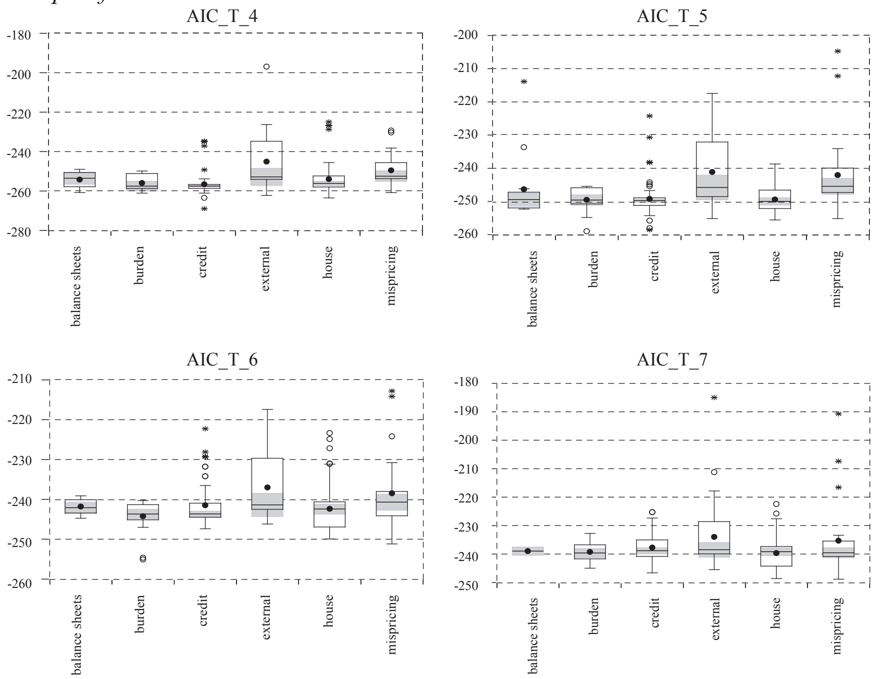
Source: Author's calculation.

[10] As the baseline model has LogL value of 131.873 (see table 2), inclusion of indicator variables in the model is considered successful if the new model has a higher value than the baseline model to ensure comparability across all specifications, all models are contrasted on the same length of *T*.

Figure 2 also shows that the best-performing models are those regarding the credit, house price, and debt burden variable categories. In contrast, external imbalances are those that have the greatest variation of results. This means that in modelling the HIFS regime-switching behaviour, it would be useful to include credit and debt burden variables with a 4-quarter lag. However, the rest of the categories with the exception of the aforementioned external imbalances, are also found to have good Log L values. AIC values in figure 3 tell a similar story (when the same analysis is done on the interquartile range and the average and median values compared to the original model's value of -247.6). Here, we can also see that for *t*-7 and *t*-6, most of the variables have greater AIC values than the baseline model, contributing to the validity of *t*-4 and *t*-5 variants.

FIGURE 3

Box plot for AIC values across all 964 models



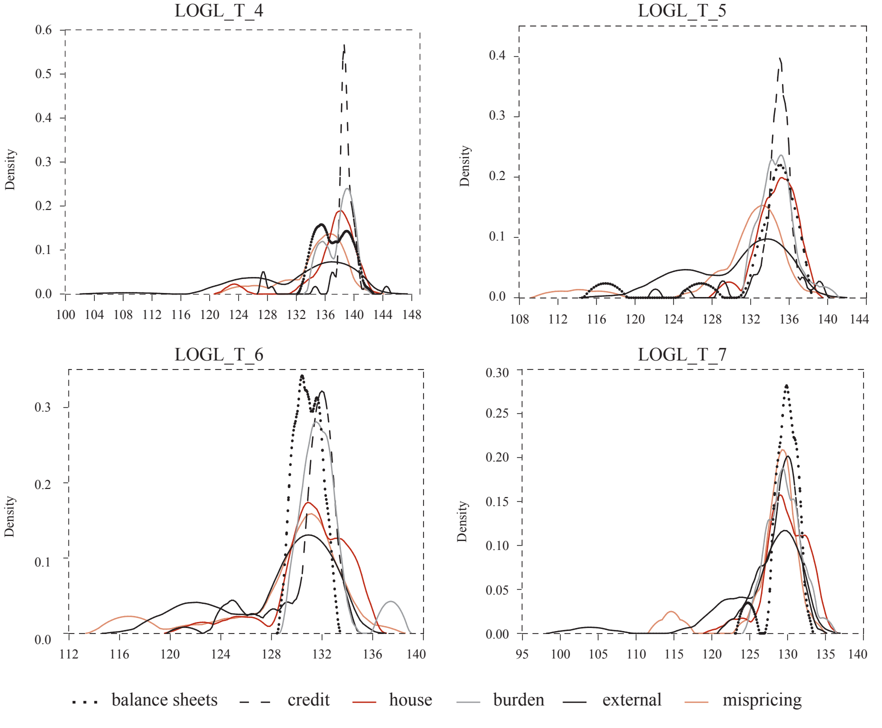
Note: Shaded areas denote 95% confidence interval for the median value. "balance sheets", "burden", "credit", "external", "house" and "mispricing" denote the six categories of measures from table 1: strength of balance sheets, private sector debt burden, credit dynamics, external imbalances, potential overvaluation of house prices and mispricing of risks respectively. t_4, t_5, t_6 and t_7 denote models that include indicator variables lagged for 4, 5, 6 and 7 quarters respectively.

Source: Author's calculation.

Not all variables are best performing for the *t*-4 case, which is seen in figure 4. It observes the empirical density functions of each risk category for all four lags observed in the study. This helps determine which lag could be best for a variable category, or individual variables when deciding which variant it is best to use in practice. This figure tells the following for the credit series as we move from *t*-7 to *t*-4, the density function becomes tighter with higher peaks. However, the best performance is found for *t*-5 for the balance sheets category, as here, most of the observations are above the value of 132, with a high peak, compared to higher peaks of lower values for *t*-6 and *t*-7 cases. Other measures have the best performance for *t*-4 and *t*-5 for the house prices category; the external and mispricing categories have the worst performances overall, due to high dispersion, especially regarding the left tails.

Figure 4

Empirical density functions of log likelihood values across all 964 models



Note: "balance sheets", "burden", "credit", "external", "house" and "mispricing" denote the six categories of measures from table 1: strength of balance sheets, private sector debt burden, credit dynamics, external imbalances, potential overvaluation of house prices and mispricing of risks respectively. t_4, t_5, t_6 and t_7 denote models that include indicator variables lagged for 4, 5, 6 and 7 quarters respectively.

Source: Author's calculation.

In general, the promising results of models with the inclusion of indicator variables are in line with other studies that analyse some form of non-linear models, such as Davis and Karim (2008) and Vašíček et al. (2017). To summarize the results, it is evident that the performance of the models gets better when the lag of the indicator variable gets smaller. There is a trade-off, however, as the information about some variables has a certain lag. The results found here are in line with previous literature. Credit dynamics is found to be the best predictor of financial crises in previous literature (Borio and Lowe, 2002; Borio and Drehmann, 2009;

Aldasoro, Borio and Drehmann, 2018), with newer studies including Schularick and Taylor (2012), alongside Drehmann and Juselius (2014), where the debt service to income ratio is one of the best early warning indicators, as found in this study. The private sector debt burden category is relevant, as found by Detken et al. (2014), or Giese et al. (2014). Moreover, the mispricing and external imbalances categories were found to have poor performance, as found in Slingenberg and de Haan (2011). The reasoning could be found in the poor performance of the stock market index, which has had little dynamics in the last 10-12 years. The lack of dynamics in a variable could result in no variation being caught in the modelling process. The results here are also in line with the research of Miszina and Tkacz (2009), who find that indicators such as asset prices are better predictors of future financial stress when nonlinearities are included in the analysis.

However, previous analysis observes the entire distribution performance of each category. In the next step, we observe those individual indicators that are useful in transition probability forecasting (equation (3)). We extract those indicators that have significant parameters in equation (3) and that are positive, indicating that greater values of those indicators affect the future probability of staying within the stressful regime or entering it. Put differently, when estimating the regime-switching model, the matrix in (2) includes the estimation of the parameter that is related to entering the regime of higher stress if the DGP was in the lower stress regime in the previous period. Furthermore, the matrix includes the parameter related to staying in the regime of higher stress if the DGP was previously in the same regime in the previous period. These parameters are depicted in table 3. The left panel in table 3 shows the value of the estimated parameters from equation (3) for the case of entering the higher risk from the lower risk regime for all four types of lags: t-4 to t-7. The right panel in table 3 contains values of the parameters regarding the case of staying in the higher risk regime. As the indicator variables are in different units of measurement, they are not directly comparable. The values are shown so that the dynamics of the parameter value can be observed. The dynamics are stable across all individual variables in the left panel in table 3. In contrast, an increasing trend is found in the right panel of table 3 for some variables (e.g., credit and house price series). Stable dynamics indicate a stable effect of a variable in the model, which is needed for its credibility. Increasing dynamics of the parameter in the right panel in table 3 means that the effects of individual indicators on the probability of staying in higher stress regime increases as the lag gets smaller. Such a result is in line with previous early warning system research[11] that concludes that the best-performing signalling indicators increase their values the most several quarters before the crisis hits. In the context of regime-switching, this could be interpreted as the persistence in those indicators accumulating the total effects for the future transition probability.

[11] See literature review section, and introduction.

These results are in line with Duprey and Klaus (2017; 2022), who find that the credit and housing dynamics are the best predictors for entering a more stressful period; with Pietrzak (2021) who discovers that variables related to real estate markets (alongside earnings and profitability) are helpful in financial stress forecasting; specifically real estate prices as in Vašiček et al. (2017), as well as with some earlier studies such as Adalid and Detken (2007); and more recent, such as Christensen and Li (2014), and Slingenberg and Haan (2011), who also find that house price returns and credit growth are best predictors of financial stress. The house price-to-income ratio significance in this study is in line with Anundsen et al. (2016), who also found that the higher the value of this ratio, the more the probability of financial crisis increases. Although the results of this study are aligned with this mentioned research, it should be noted that the comparisons are made on a broader scale, i.e., this research looks at the effects of individual indicators on risk probabilities. In contrast, most of the work mentioned looks at the values of financial stress. As it is difficult to predict future values of financial stress, as the mentioned literature agrees, this research has the advantage of predicting just higher or lower stress regime probability, not actual values of financial stress.

In economic language, the results in table 3 are useful to determine which variables should be closely monitored over the financial cycle, as their dynamics affects the probability of entering or staying in the higher stress regime. The policymaker can observe these variables and their variations more closely, compared to the rest of those available, as monitoring many data at once is resource and time-consuming. Moreover, the results in table 3 show that the policymakers face difficulties when actively tracking the dynamics of most of the risk indicators. This is due to different units of measurement for each of the series, as this indicates non-synchrony of their characteristics regarding the assumption of the length of the financial cycle. This has proven in the literature to be a very difficult task, in Croatia as well (CNB, 2022b). Moreover, the policymaker needs to choose other relevant criteria to find the best indicator from figures 5 and 6. This will depend on the decision maker's preferences and tracking of other indicator dynamics. One possible solution could be observing interval estimates of such indicators where different smoothing parameters are used. An example of such an application is found in CNB (2022b), in which flexibility is introduced by observing such intervals. When more data become available[12], the policymaker could re-estimate all of the models to see which indicators are best for specific purposes, including the one in this study.

[12] And more data on the characteristics of individual indicators during different phases of the financial cycle, so that the duration of the cycle could be estimated better.

Table 3

Significant coefficients, probability of transitioning from lower to higher risk regime (upper panel) and significant coefficients, probability of staying in the higher risk regime (lower panel)

Indicator	*t*-7	*t*-6	*t*-5	*t*-4
HPI 2y growth rate	0.187	0.164	0.152	0.163
HPI real, gap 25K	0.176	0.130	0.131	0.161
Capital/Assets, gap 1600	0.147	0.100	0.130	0.119
CNFP real 1y change	0.072	0.091	0.171	0.131
Basel gap, 25K	0.277	0.238	0.381	0.456
Broad credit 1y change	0.069	0.075	0.066	0.071
Broad credit real 1y change	0.046	0.050	0.044	0.047
Narrow credit gap, 125K	0.654	0.748	0.740	0.689
Narrow credit gap, 25K	0.740	0.833	0.836	0.777
Narrow credit gap, 400K	0.645	0.737	0.727	0.675
Narrow credit gap, 85K	0.663	0.757	0.754	0.705
CNFC 2y growth rate	0.128	0.139	0.156	0.147
H D-to-I 2y change	0.019	0.018	0.019	0.016
CNFC, 25K	0.185	0.181	0.187	0.163
P-to-Income, 400K	0.110	0.055	0.143	0.165
Deposits/Credit, 1600	0.019	0.018	0.020	0.023
HPI real 1y growth rate	0.149	0.150	0.177	0.187
P-to-I 2y growth rate	0.190	0.181	0.198	0.231
P-to-I real 2y growth rate	0.164	0.171	0.199	0.238
Narrow credit 1y growth rate	0.149	0.175	0.230	0.268
Narrow credit 2y growth rate	0.149	0.175	0.230	0.268
NX, cumsum	0.407	0.359	0.321	0.311
Net ext debt, 125K	0.332	0.310	0.404	0.332
Net ext debt, 25K	0.310	0.404	0.332	0.332
Net ext debt, 400K	0.332	0.310	0.404	0.332
Net ext debt, 85K	0.332	0.310	0.404	0.332
NX	0.327	0.291	0.301	0.392
HH 1y change	0.509	0.294	0.300	0.556

Note: 1600, 25K, 85K, 125K and 400K denote the value of the smoothing parameter in HP gap, in values of 1,600, 25,000, 85,000, 125,000 and 400,000 respectively. 1y and 2y are one and two years, CNFP is credit to nonfinancial corporations, HPI is house price index, Capital/Assets is the capital-to-assets ratio. t-4, t-5, t-6 and t-7 denote models that include indicator variables lagged for 4, 5, 6 and 7 quarters respectively. HH denotes credit to households, Net ext debt denotes net external debt, NX is net exports share in GDP, whereas NX cumsum is sum of net exports over 4 quarters share in sum of GDP over 4 quarters, P-to-I is the price to income ratio, HPI is house price index, Deposits/Credit denotes the deposit-to-credit ratio, P-to-Income is house price-to-income ratio, CNFC is credit to nonfinancial corporations, H D-to-I is household debt to income ratio.

Source: Author's calculation.

4.3 DISCUSSION

Anundsen et al. (2016) state that specific information is present in each single country's history of financial crises. Thus, it should be said that in the period before GFC, Croatia experienced an increase in internal and external vulnerabilities due to excessive credit activity and foreign borrowing. This is captured in the results, in the statistical significance of this risk category. Moreover, FDI (foreign direct investment) inflows at the beginning of the 2000s due to the privatization of telecoms and oil companies, and government foreign borrowing increased rapidly. Then, the GFC hit, and re-financing problems spilled over from mortgage markets to interbank money markets worldwide, including Croatia. This was when the HIFS indicator spiked in the analysed period (figure 1). Some of the macroprudential measures taken in the pre-GFC period had the goals of reducing credit growth and capital inflows, supporting the banking system's liquidity and increasing its capitalization. By looking at all of the measures being put into place before the GFC (see Kraft and Galac, 2011 for details), it is seen that a lot of fine-tuning was done in the years that preceded the crisis. Kraft and Galac (2011) and Galac (2010) agree that some measures did manage to slow down the credit growth in that period, alongside the marginal reserve requirement. All this information indicates that the CNB utilized effective measures. However, many problems of that era, many of which were out of Croatia's scope, contributed to the consequent crisis. If the knowledge from the estimated results here had been known back then, maybe the macroprudential policy would have been even tighter. This conclusion stems from the facts of what the policymakers were doing back then, just by observing the happenings in the financial system. As they were very active and prudent, having the output from models such as the one in this study would perhaps lead to more significant tightening. However, the measures taken back then for surely have contributed to lowering the probability of entering the more stressful regime.

The period from 2010 to 2015 was characterized by several stressful events for the Croatian markets: the government debt crisis in the Eurozone, the rise of CDS premiums on parent bank bonds of the largest banks in Croatia, the fall of the credit rating of Croatia in 2013 and increased costs of foreign borrowing. In this period, the macroprudential policy was loosening for the most part, as CNB was lowering the minimum required amount of foreign currency claims and was lowering reserve requirement rates so that the released funds would be utilized for economic recovery. Bambulović and Valdec (2020) found that domestic and foreign banks increased lending activity after the loosening measures. However, as Dumičić (2015a) pointed out, the recovery was prolonged due to a lack of structural reforms and a deterioration of fiscal indicators. This period indicates that although CNB had good timing of measures, the good conduct of macroprudential policy is not enough to bring the economy on the right path if other conditions are not fulfilled. The last sub-period, from 2016 until the end of the sample, was the most tranquil period until the COVID-19 shock hit. Here, the economic activity was in recovery, and lending to the private sector started to increase. In addition, CNB introduced tightening measures due to the new Basel accords and legislation. Finally, the COVID-19 shock was exogenous and could not have been forecasted with the cyclical risk indicators. This shows that there is a need for the policymaker to track structural and cyclical risks and mitigate them in a timely fashion so that the shocks in the financial stress series come from mostly exogenous shocks.

Consequences for policymakers are as follows. The results show that some indicators of cyclical risk accumulation provide information about future financial stress dynamics. The policymaker could use this information to narrow the most important indicators that need to be tracked over time and, consequently, tailor policies that would mitigate those risks. Although CNB had timely measures that focused on the majority of the risk indicators over the entire observed period, better coordination among macroprudential, monetary and fiscal policies is needed to achieve healthier economic growth. As some asymmetry in results is found, this indicates that the tightening and loosening policy should not be considered similarly.

The results indicate that policymakers will have to be flexible in policy tailoring, due to the difficulties of predicting financial stress. However, the results here do tell a similar story, in line with previous empirical research: credit and housing variables are extremely important in future financial risk materialization prediction and thus, useful for countercyclical capital buffer (CCyB) calibration (Bonfim and Monteiro, 2013)[13].The best indicators in this research are in line with the credit dynamics, such as the credit-to-GDP gap and growth rates of credit (Borio, 2012; Borio and Drehmann, 2009; Babecký et al., 2014); and property price dynamics, such as the growth rates and gaps regarding house prices, rents, construction work index (Borio, 2012; Jordá, Schularick and Taylor, 2015; Behn et al., 2013). Furthermore, external debt imbalances have been examined in the literature for a long time due to the analysis of currency crises. Variables regarding this group of potential indicators include different transformations of a country's gross and net external debt, terms of trade, current account, and net export (Giese et al., 2014; Laeven and Valencia, 2008; Tölö, Laakkonen and Kalatie, 2018). In addition, the balance sheets of credit institutions were often analysed as well, as they give insights into potential weaknesses of the structure of the balance sheets. Here, the leverage ratio and other relevant ratios regarding deposits, assets, and credits of banks were analyzed (see Laina, Nyholm and Sarlin, 2015; Alessandri et al., 2015; Drehmann and Juselius, 2014; Rychtarik, 2014). Furthermore, the private sector debt burden has been recognized as another important group of measures, which includes debt service ratios (see Lombardi, Mohatny and Shim, 2017; Detken et al., 2014; Drehmann and Juselius, 2012; 2014). Finally, another group of variables that is important in evaluating their potential in signalling future crises is the mispricing of risks. It includes measures that relate to the credit institutions' skewed views (interest rate margins), but private investors as well (in terms of stock market index-derived measures). Here, research finds that

[13] That is why additional analysis was done by constructing a composite indicator of cyclical risks based on the results from this study and it has a potential for being used in practice. Details are available upon results.

optimism and pessimism changes of economic agents over the financial cycle also affect the build-up of cyclical imbalances (see Pfeifer and Hodula, 2018; Drehmann et al., 2010).

5 CONCLUSION

Forecasting future financial stress is extremely relevant and topical. However, macroprudential policy is still relatively new compared to other economic policies, with respect to data availability and to the possibility of investigating transmission mechanisms of instruments and general interactions with different policies. Thus, any step forward in these areas contributes to a better understanding of the matter, so policymakers can tailor better instruments to achieve important objectives. This paper contributes to the literature by finding indicators that help forecast financial stress, in terms of switching from one regime to another, alongside utilizing techniques to reduce their number.

Instead of defining important dates of crises or specific events happening in the financial system, we allow the model to optimize the selection of which data should belong to each of the regimes. This lets the policymaker allow the data-generating process "to speak" freely without imposing assumptions or restrictions on the modelling process. Next, we focus on the properties of estimated regime-switching models across chosen lag lengths of indicator variables, and across risk group categories. This gives concise information about the quality of selected variables overall in terms of their forecasting capabilities. Nevertheless, the financial stress variable is very hard to predict, as it can react to different political, economic, and other shocks in the economy. That is why the approach made here was not to forecast the value of the stress variable itself but to try to incorporate additional information into the probability of (re)entering the higher stress regime. The results are in line with related literature, as the credit and housing price variables stood out the most here, as they did in previous research on related topics. But here, this confirmation is obtained from another approach to the problem. Although more than 900 models have been estimated and compared, with more than 240 variables considered in the analysis, one could say that the results are not so promising due to not more than 20 variables having been found to have meaningful results. This does not have to reflect poor data availability or other similar reasoning. On the contrary, it could reiterate the problems of macroprudential policymaking and research as a still insufficiently analysed area of research and application.

However, we are aware of the shortfalls of such approaches due to lack of data, shorter periods, bias of results, etc. The Croatian time series is short compared to other countries which only enabled in-sample analysis. Single-country analysis suffers from having only a few crisis periods in the sample, which is not something new (see Claessens, 2009). However, a starting point is needed, and macroprudential instruments need to be calibrated on the specific characteristics of a country. Moreover, a single-country analysis can be enhanced by combining both approaches (EWS indicators and the regime-switching of stress) has the potential to obtain overall good forecasting results. As Kauko (2014) observes, most of research on EWS utilizes binary variables as the dependent ones in the analysis. As crises rarely occur, meaning that not many observations are included in the regime of crisis occurrence, the combination of financial stress data as the dependent variable can be helpful. Vermeulen et al. (2015) agree on this topic. Finally, the EWS approach utilized often in research depends on the policymaker's preferences regarding false alarms and missing signals. On the other hand, regime-switching does not include such subjectivity. The heterogeneity in different countries may not capture specific variables or their transformations as relevant in the modelling process.

Macroprudential policymakers could use results from this study to track specific indicators in greater detail and try to estimate the stance concerning the instruments used over the observed period and the goals set during the boom-and-bust phases of the financial cycle. The Croatian case showed that the policymakers had a good focus on specific problems that were occurring during the entire observed sample. Most tightening and loosening measures were tailored according to some of the indicators found to be the best-performing ones in this study.

Future work should seek how to do a similar analysis such as this one on a larger scale, considering that all different variable categories, variable transformations, and lag selection make this a big data problem. This could be relevant for international institutions or those who need to monitor more countries simultaneously. Other institutions focusing on one country could apply this approach regularly, since the problem is focused on the country the researcher or policymaker is familiar with the most.

Disclosure statement

The author declares that there is no conflict of interest.

REFERENCES

1. Adalid, R. and Detken, C., 2007. Liquidity shocks and asset price boom/bust cycles. *ECB Working Paper*, No. 732. https://doi.org/10.2139/ssrn.963147
2. Adrian, T., Boyarchenko, N. and Giannone, D., 2019. Vulnerable growth. *American Economic Review*, 109(4), pp. 1263-1289. https://doi.org/10.1257/aer.20161923
3. Aldasoro, I., Borio, C. and Drehmann, M., 2018. Early warning indicators of banking crises: expanding the family. *BIS Quarterly Review*, March 2018.
4. Alessandri, P. [et al.], 2015. A note on the implementation of a Countercyclical Capital Buffer in Italy. *Bank of Italy Occasional Paper,* No. 278. https://doi.org/10.2139/ssrn.2649949
5. Alessi, L. and Detken, C., 2019. Identifying excessive credit growth and leverage. *Journal of Financial Stability*, 35. pp. 215-225. https://doi.org/10.1016/j.jfs.2017.06.005
6. Ang, A. and Bekaert, G., 2002a. Short Rate Nonlinearities and Regime Switches. *Journal of Economic Dynamics and Control*, 26, pp. 1243-1274. https://doi.org/10.1016/S0165-1889(01)00042-2
7. Ang, A. and Bekaert, G., 2002b. International Asset allocation With Regime Shifts. *Review of Financial Studies*, 15, pp. 1137-1187. https://doi.org/10.1093/rfs/15.4.1137
8. Ang, A. and Timmermann, A., 2012. Regime changes and financial markets. *Annual Review of Financial Economics,* 4, pp. 313-337. https://doi.org/10.1146/annurev-financial-110311-101808
9. Anundsen, A. K. [et al.], 2016. Bubbles and crises: The role of house prices and credit. *Journal of Applied Econometrics,* 31, pp. 1291-1311. https://doi.org/10.1002/jae.2503
10. Arbatli-Saxegaard, E. C. and Muneer, M. A., 2020. The countercyclical capital buffer: A cross-country overview of policy frameworks. *Staff memo*, No. 6.
11. Babecký, J. [et al.], 2014. Banking, Debt, and Currency Crises in Developed Countries: Stylized Facts and Early Warning Indicators. *Journal of Financial Stability*, 15, pp. 1-17. https://doi.org/10.1016/j.jfs.2014.07.001
12. Baele, L., 2003. Volatility Spillover Effects in European Equity Markets: Evidence from a Regime-Switching Model. *Ghent University, Working paper*, No. 189.
13. Bambulović, M. and Valdec, M., 2020. Testing the characteristics of macroprudential policies' differential impact on foreign and domestic banks' lending in Croatia. *Public Sector Economics*, 44(2), pp. 221-249. https://doi.org/10.3326/pse.44.2.4
14. Behn, M. [et al.], 2013. Setting Countercyclical Capital Buffers Based on Early Warning Models: Would It Work? *ECB Working Paper*, No. 1604.
15. Bloom, N., 2009. The Impact of Uncertainty Shocks. *Econometrica,* 77(3), pp. 623-685. https://doi.org/10.3982/ECTA6248
16. Bonfim, D. and Monteiro, N., 2013. The Implementation of the Countercyclical Capital Buffer: Rules versus Discretion. *Financial Stability Report (November)*, pp. 87-110.

17. Borio, C. and Drehmann, M., 2009. Assessing the Risk of Banking Crises – Revisited. *BIS Quarterly Review*, (March), pp. 29-46.
18. Borio, C. and Lowe, P., 2002. Asset prices, financial and monetary stability: exploring the nexus. *BIS Working Papers*, No. 114. https://doi.org/10.2139/ssrn.846305
19. Borio, C., 2012. The Financial Cycle and the Marco-economy: What have we learnt. *BIS Working Paper Series*, No. 395.
20. Candelon, B., Dumitrescu, E-I. and Hurlin, C., 2012. How to Evaluate an Early-Warning System: Toward a Unified Statistical Framework for Assessing Financial Crises Forecasting Methods. *IMF Economic Review*, 60, pp. 75-113. https://doi.org/10.1057/imfer.2012.4
21. Cardarelli, R., Elekdag, S. and Lall, S., 2011. Financial stress and economic contractions. *Journal of Financial Stability*, 7(2), pp. 78-97. https://doi.org/10.1016/j.jfs.2010.01.005
22. Chavleishvili, S. and Manganelli, S., 2019. Forecasting and stress testing with quantile vector autoregression. *ECB Working Paper*, No. 2330. https://doi.org/10.2139/ssrn.3489065
23. Christensen, I. and Li, F., 2014. Predicting financial stress events: a signal extraction approach. *Journal of Financial Stability*, 14, pp. 54-65. https://doi.org/10.1016/j.jfs.2014.08.005
24. Claessens, S., 2009. Discussion of Credit, asset prices, and financial stress. *International Journal of Central Banking*, pp. 123-129.
25. CNB, 2022a. *Statistical data*. Zagreb: Croatian National Bank.
26. CNB, 2022b. *Macroprudential diagnostics*, No. 16. Zagreb: Croatian National Bank.
27. Davis, E. P. and Karim, D., 2008. Comparing early warning systems for banking crises. *Journal of Financial Stability*, 4(2), pp. 89-120. https://doi.org/10.1016/j.jfs.2007.12.004
28. Detken, C. [et al.], 2014. Operationalising the countercyclical capital buffer: indicator selection, threshold identification and calibration options. *ESRB Occasional Paper,* No. 2014/5. https://doi.org/10.2139/ssrn.3723336
29. Dimova, D., Kongsamut, P. and Vandenbussche, J., 2016. Macroprudential Policies in Southeastern Europe. *Working Paper,* No. 16/29. https://doi.org/10.5089/9781498342872.001
30. Doz, C., Ferrara, L. and Pionnier, P., 2020. Business cycle dynamics after the Great Recession: An extended Markov-Switching Dynamic Factor Model. *OECD Statistics Working Papers*, No. 2020/01. https://doi.org/10.1787/9626dda3-en
31. Drehmann, M. [et al.], 2010. Countercyclical capital buffers: exploring options. *BIS Working Papers,* No. 317. https://doi.org/10.2139/ssrn.1648946
32. Drehmann, M. and Juselius, M., 2012. Do debt service costs affect macroeconomic and financial stability? *BIS Quarterly Review,* September 2012.
33. Drehmann, M. and Juselius, M., 2014. Evaluating early warning indicators of banking crises: Satisfying policy requirements. *International Journal of Forecasting,* 30(3), pp. 759-780. https://doi.org/10.1016/j.ijforecast.2013.10.002

34. Dumičić, M., 2015a. Financial stress indicators for small, open, highly euroized countries: the case of Croatia. *Financial Theory and Practice*, 39(2), pp. 171-203. https://hrcak.srce.hr/139668
35. Dumičić, M., 2015b. Financial Stability Indicators – The Case of Croatia. *Journal of Central Banking Theory and Practice*, (1), pp. 113-140. https://doi.org/10.1515/jcbtp-2016-0006
36. Duprey, T. and Klaus, B., 2017. How to predict financial stress? An assessment of Markov switching models. *ECB Working Paper,* No. 2057. https://doi.org/10.2139/ssrn.2968981
37. Duprey, T. and Klaus, B., 2022. Early warning or too late? A (pseudo-) real-time identification of leading indicators of financial stress. *Journal of Banking and Finance*, 138, pp. 106-196. https://doi.org/10.1016/j.jbankfin.2021.106196
38. Duraković, E., 2021. Croatian Systemic Stress Index. *CNB Working Paper,* No. I-61.
39. ECB, 2017. A new database for financial crises in European countries. ECB/ESRB EU crises database. *ECB Occasional Paper Series*, No. 194.
40. ESRB 2018. *The ESRB handbook on operationalising macroprudential policy in the banking sector*. European Systemic Risk Board.
41. ESRB, 2014. Recommendation of the European Systemic Risk Board of 18 June 2014 on guidance for setting countercyclical buffer rats (ESRB/2014/1). *Official Journal of the European Union* C 293/1.
42. Franses, P. H. and van Dijk, D., 2000. *Nonlinear time series models in empirical finance*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511754067
43. Galac, T., 2010. The Central Bank as Crisis Manager in Croatia – A Counterfactual Analysis. *Croatian National Bank Working Paper,* W-27.
44. Galán, J. E., 2019. Measuring Credit-to-GDP Gaps. The Hodrick-Prescott Filter Revisited. *Documentos de Trabajo,* No. 1906. https://doi.org/10.2139/ssrn.3384613
45. Galati, G. [et al.], 2016. Measuring financial cycles in a model-based analysis: Empirical evidence for the United States and the Euro Area. *Economics Letters,* 145, pp. 83-87. https://doi.org/10.1016/j.econlet.2016.05.034
46. Giese, J. [et al.], 2014. The credit-to GDP gap and complementary indicator for macroprudential policy: evidence from the UK. *International journal of finance and economics,* 19(1), pp. 25-47. https://doi.org/10.1002/ijfe.1489
47. Giglio, S., Bryan, K. and Pruitt, S., 2016. Systemic Risk and the Macroeconomy: An Empirical Evaluation. *Journal of Financial Economics*, 119(3), pp. 457-471. https://doi.org/10.1016/j.jfineco.2016.01.010
48. Gneitinga, T. and Ranjanb, R., 2011. Comparing Density Forecasts Using Threshold- and Quantile-Weighted Scoring Rules. *Journal of Business and Economic Statistics*, 29(3), pp. 411-422. https://doi.org/10.1198/jbes.2010.08110
49. Gray, S. F., 1996. Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process. *Journal of Financial Economics*, 42, pp. 27-62. https://doi.org/10.1016/0304-405X(96)00875-6

50. Hanschel, E. and Monnin, P., 2005. Measuring and Forecasting Stress in the Banking Sector: Evidence from Switzerland. *BIS Papers*, No. 22, pp. 431-449.
51. Hodrick, R. and Prescott, E., 1997. Postwar U.S. Business Cycles. An Empirical Investigation. *Journal of Money, Credit, and Bankin*g, 29, pp. 1-16. https://doi.org/10.2307/2953682
52. Holló, D., Kremer, M. and Lo Duca, M., 2012. CISS – A composite indicator of systemic stress in the financial system. *ECB Working Paper Series*, No. 1426.
53. Hubrich, K. and Tetlow, R., 2015. Financial stress and economic dynamics: The transmission of crises. *Journal of Monetary Economics*, 70(C), pp. 100-115. https://doi.org/10.1016/j.jmoneco.2014.09.005
54. Jordá, O., Schularick, O. and Taylor, A. M., 2015. Leveraged Bubbles. *Journal of Monetary Economics*, 76(Supplement), pp. S1-S20. https://doi.org/10.1016/j.jmoneco.2015.08.005
55. Kaminsky, G. L. and Reinhart, C. M., 1999. The Twin Crises: The Causes of Banking and Balance-of-Payments Problems. *American Economic Review*, 89(3), pp. 473-500. https://doi.org/10.1257/aer.89.3.473
56. Kauko, K., 2014. How to foresee banking crises? A survey of the empirical literature. *Economic Systems*, 38(3), pp. 289-308. https://doi.org/10.1016/j.ecosys.2014.01.001
57. Kim, C-J., 1994. Dynamic Linear Models with Markow-Switching. *Journal of Econometrics*, 60, pp. 1-22. https://doi.org/10.1016/0304-4076(94)90036-1
58. Klomp, J., 2010. Causes of banking crises revisited. *The North American Journal of Economics and Finance*, 21(12), pp. 72-87. https://doi.org/10.1016/j.najef.2009.11.005
59. Kraft, E. and Galac, T., 2011. Macroprudential Regulation of Credit Booms and Busts, The Case of Croatia. *Policy Research Working Paper,* No. 5772. https://doi.org/10.1596/1813-9450-5772
60. Kupkovič, P. and Šuster, M., 2020. Identifying the Financial Cycle in Slovakia. *NBS working papers*, No 2/2020.
61. Laeven, L. and Valencia, F., 2008. Systemic Banking Crises: A New Database. *IMF Working Paper,* No. WP/08/224. https://doi.org/10.5089/9781451870824.001
62. Laina, P., Nyholm, J. and Sarlin, P., 2015. Leading Indicators of Systemic Banking Crises: Finland in a Panel of EU Countries. *Review of Financial Economics*, 24, pp. 18-35. https://doi.org/10.1016/j.rfe.2014.12.002
63. Lang, J-H. [et al.], 2019. Anticipating the bust: a new cyclical systemic risk indicator to assess the likelihood and severity of financial crises. *Occasional Paper Series,* No. 219. https://doi.org/10.2139/ssrn.3334835
64. Lo Duca, M. [et al.], 2017. A new database for financial crises in European countries. *Occasional paper series,* No. 194. https://doi.org/10.2139/ssrn.1773446
65. Lo Duca, M. and Peltonen, T., 2011. Macrofinancial vulnerabilities and future financial stress: assessing systemic risks and predicting systemic events. *Macroprudential regulation and policy*, 60, pp. 82-88. https://doi.org/10.2139/ssrn.1803075

66. Lombardi, M. J., Mohatny, M. and Shim, I., 2017. The real effects of household debt in the short and long run. *BIS working paper,* No 607.
67. Misina, M. and Tkacz, G., 2009. Credit, asset prices, and financial stress. *International Journal of Central Banking*, 5, pp. 95-122.
68. O'Brien, M. and Wosser, M., 2021. Growth at Risk & Financial Stability. *Financial stability notes*, 2021(2).
69. Pfeifer, L. and Hodula, M., 2018. A profit-to-provisioning approach to setting the countercyclical capital buffer: the Czech example. *CNB working paper,* No. 5/2018. https://doi.org/10.2139/ssrn.3723430
70. Pietrzak, M., 2021. Can Financial Soundness Indicators Help Predict Financial Sector Distress? *IMF Working Paper,* No. WP/21/197. https://doi.org/10.5089/9781513593005.001
71. Plagborg-Moller, M. [et al.], 2020. When is growth at risk? Brookings Papers on Economic Activity, (Spring), pp. 167-229. https://doi.org/10.1353/eca.2020.0002
72. Reinhart, C. M. and Rogoff, K. S., 2009. *This Time Is Different: Eight Centuries of Financial Folly.* Princeton University Press. https://doi.org/10.1515/9781400831722
73. Rychtarik, Š., 2014. Analytical background for the counter-cyclical capital buffer decisions in Slovakia. *Ročník,* 22(4), pp. 10-15.
74. Schularick, M. and Taylor, A. M., 2012. Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870-2008. *The American Economic Review*, 1022, pp. 1029-1061. https://doi.org/10.1257/aer.102.2.1029
75. Schüler, Y., 2018. Detrending and financial cycle facts across G7 countries: mind a spurious medium term! *ECB Working Paper Series*, No. 2138. https://doi.org/10.2139/ssrn.3152820
76. Slingenberg, J. W. and de Haan, J., 2011. Forecasting Financial Stress. *De Nederlandsche Bank Working Paper,* No. 292. https://doi.org/10.2139/ssrn.1951800
77. Škrinjarić, T., 2022. Macroeconomic effects of systemic stress: a rolling spillover index approach. *Public Sector Finance*, 46(1), pp. 109-140. https://doi.org/10.3326/pse.46.1.4
78. Tölö, E., Laakkonen, H. and Kalatie, S., 2018. Evaluating Indicators for Use in Setting the Countercyclical Capital Buffer. *International Journal of Central Banking,*142, pp. 51-111.
79. Vašiček, B. [et al.], 2017. Leading indicators of financial stress: new evidence. *Journal of Financial Stability*, 28, pp. 240-257. https://doi.org/10.1016/j.jfs.2016.05.005
80. Vermeulen, R. [et al.], 2015. Financial Stress Indices and Financial Crises. *Open Economies Review*, 26(3), pp. 383-406.
81. Vujčić, B. and Dumičić, M., 2016. Managing Systemic Risks in the Croatian Economy. *BIS Paper,* No. 861.
82. Wezel, T., 2019. Conceptual Issues in Calibrating the Basel III Countercyclical Capital Buffer. *IMF Working paper,* No. 2019/086.

Table A1

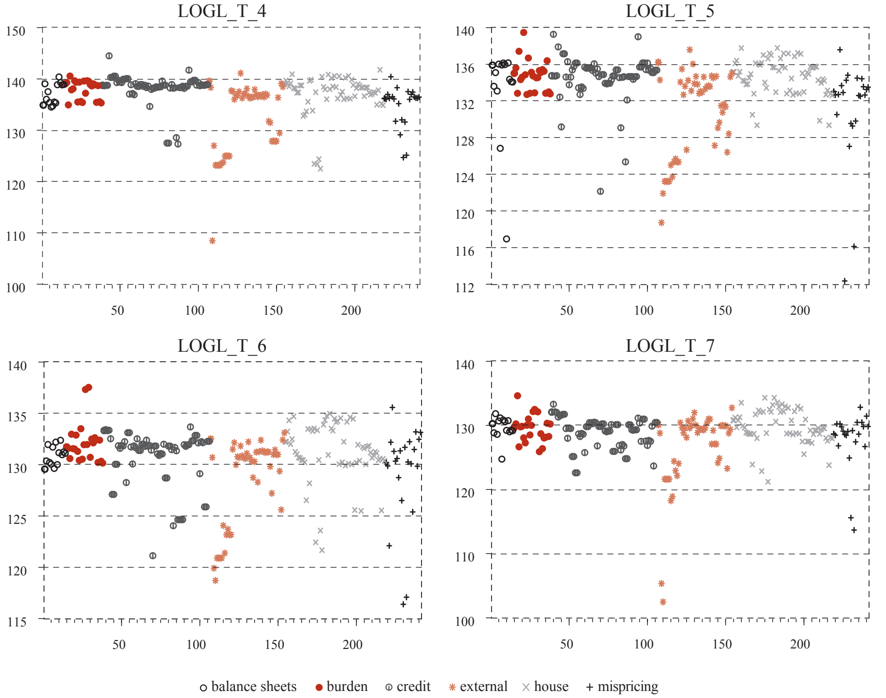
SIC information criterion on different specifications of ARMA(p,q) model for variable HIFS

AR/MA	0	1	2	3
0	-1.9645	-2.5218	-2.6794	-2.6643
1	-2.5912	-2.6729	-2.6539	-2.5977
2	-2.6839	-2.6389	-2.6038	-2.5532

Source: Author's calculation.

Figure A1

Log likelihood statistics for regime switching models with variable indicators included in the equation (3)



Note: "balance sheets", "burden", "credit", "external", "house" and "mispricing" denote the six categories of measures from table 1: strength of balance sheets, private sector debt burden, credit dynamics, external imbalances, potential overvaluation of house prices and mispricing of risks respectively. t_4, t_5, t_6 and t_7 denote models that include indicator variables lagged for 4, 5, 6 and 7 quarters respectively.

Source: Author's calculation.

An analysis of COFOG expenditures in former Yugoslavian countries

MARKO CRNOGORAC, Ph.D.*
SANTIAGO LAGO-PEÑAS, Ph.D.*

Article**
JEL: H50
https://doi.org/10.3326/pse.47.2.4

* This research has been supported financially by the Xunta de Galicia, Grant number ED431C 2021/50. The authors thank the D.1 unit-candidate and pre-candidate countries of DG ECFIN of the European Commission for its help in obtaining some of the data used in this research as well as the contributing institutions of the sample countries who participated in our data-collection process. The usual disclaimer applies. The authors are grateful to two anonymous referees who have contributed to the quality of the final version of the paper.

** Received: June 1, 2022
Accepted: January 16, 2023

Marko CRNOGORAC
Governance and Economics Research Network (GEN), Universidade de Vigo, Campus Universitario As Lagoas s/n, 32004 Ourense, Spain
e-mail: mcrnogorac@uvigo.es
ORCiD: 0000-0002-5302-2081

Santiago LAGO-PEÑAS
Governance and Economics Research Network (GEN), Universidade de Vigo, Campus Universitario As Lagoas s/n, 32004 Ourense, Spain
e-mail: slagop@uvigo.es
ORCiD: 0000-0003-4601-8655

PUBLIC SECTOR ECONOMICS
47 (2) 233-254 (2023)

Abstract

In this paper, we carry out an in-depth analysis of public expenditure in all former Yugoslavian countries. Our purpose is threefold: first, to verify the existence of common patterns of spending; second, to investigate the cyclicality hypothesis of fiscal policy in non-OECD countries; and third, to analyse both political and economic determinants of expenditure composition. Our results show a weak convergence in structures, the countercyclical behaviour of public expenditures, and the influence of electoral cycles, business cycles, and the degree of nationalization of party systems on the composition of public expenditure.

Keywords: COFOG expenditures, former Yugoslavia

1 INTRODUCTION

Former Yugoslavian countries (Bosnia and Herzegovina, Croatia, Montenegro, North Macedonia, Serbia, and Slovenia) are all at different stages of the process of integration into the European Union (EU). While fiscal variables are a key issue for that process, related literature, both academic and policy-oriented, is rather limited. We consider the possibility of comparing evolutions from a common starting point, the dissolution of former Yugoslavia, as a type of natural experiment. The idea is to use the sequential acts of secession as a clearly defined cause that might change the outcomes of managing public expenditures from country to country. Our objective would be to assess whether the policies would converge or diverge post secession, as well as to provide conclusions explaining the fiscal behaviour.

In the past few years, we have contributed to a partial filling of the gap. Crnogorac and Lago-Peñas (2019a) analysed fiscal policy from the perspectives of the evolution of the main fiscal aggregates, composition of tax revenues, and budget elasticity in relation to the economic cycle. The focus was on the main aggregates of public deficit, revenues, and expenditures. Although data constraints allowed public revenues to be explored in more detail than public expenditures, we concluded that fiscal policies are driven by the spending component. Furthermore, in Crnogorac and Lago-Peñas (2019b) we focused on tax evasion and in Crnogorac and Lago-Peñas (2020) we analysed the determinants of tax morale as factors explaining tax evasion in the former Yugoslavian countries.

The focus of this paper is on public expenditure. In particular, we deal with the lack of publicly available data on COFOG[1] categories in Western Balkan countries to build a database facilitating new empirical research on public expenditure convergence, cyclicality, and the determinants of public spending choices. This paper takes a cross-country perspective dealing with the main fiscal aggregate, expenditures, instead of focusing on single country issues, like most recent literature. To the best

[1] The Classification of the Functions of Government (COFOG) by the United Nations (2000) is a current functional disaggregation of total general government expenditures into 10 categories.

of our knowledge, this is the first comprehensive and quantitative analysis of public expenditures in former Yugoslavian countries covering those dimensions. In more detail, our first target is to explore whether there are common patterns in spending among the sample countries, using both convergence and cluster analysis. Second, we analyse the impact of the economic cycle on overall public expenditures and each of their COFOG categories, as we did in work (Crnogorac and Lago-Peñas, 2019a) for revenue categories. Third, we shed light on the economic and political determinants of the composition of public expenditures.

The rest of the paper is organized as follows: section 2 deals with the convergence of public expenditures and their COFOG categories. Section 3 analyses whether the economic cycle is relevant in explaining the dynamics of the functions of public expenditures in the context of recent literature on developed and developing countries. Section 4 provides an estimate of the determinants of three aggregates of COFOG categories based on multiple specifications from recent literature. Section 5 concludes, providing the most relevant findings.

2 THE EVOLUTION OF PUBLIC EXPENDITURE IN FORMER YUGOSLAVIAN COUNTRIES

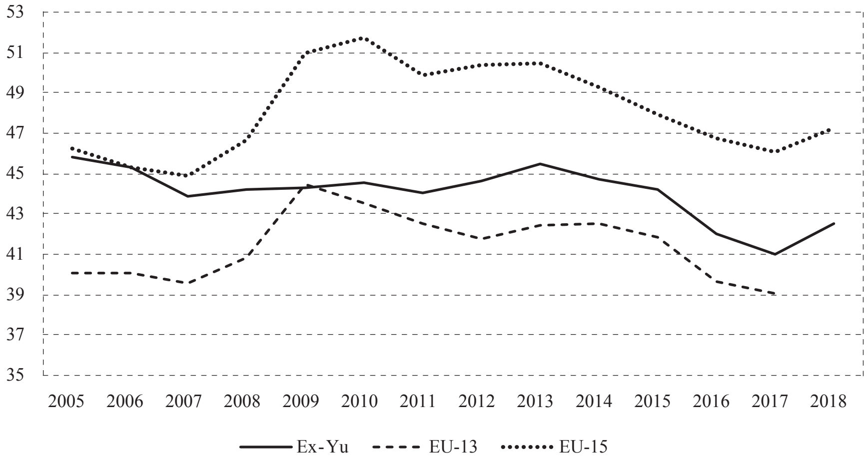
The current disaggregation of total general government expenditure by function was established in 1999. While only developed countries collected and presented data this way from the start, the COFOG methodology has spread progressively. All former Yugoslavian countries joined this trend in recent years. Reforms in many areas on their paths to join the EU have involved relevant advances in statistical data collection, including information on public expenditure. The only previous study to include at least partial data for the functional structure of the six former Yugoslavian countries is Afonso and Jalles (2014). Its sample of 155 countries included the six countries where the analysis was focused on only three specific COFOG categories: education, health, and social security. Moreover, because of their memberships of international organisations, Slovenia and Croatia have been included in other cross-sectional studies concerning public expenditure (Hessami, 2014; del Granado, Martinez-Vazquez and McNab, 2018), but a specific focus on them is lacking.

In order to analyse the evolution of public expenditures over time, we conduct several complementary analyses using a data sample[2] of six former Yugoslavian countries between 2011 and 2019, which we complement with other countries, depending on the analysis. The sample size is limited by the lack of data on all the countries in previous years, as well as the intention of the authors to exclude data post 2020 due to extensive shocks on the expenditure side. First, we turn to expenditure compositions in order to observe if common patterns exist among the six former Yugoslavian countries and the average of core EU countries, replacing

[2] The authors thank the D.1 unit-candidate and pre-candidate countries of DG ECFIN of the European Commission for its help in obtaining some of the data used in this research as well as the contributing institutions of the sample countries who participated in our data-collection process.

GDP with total expenditures in the denominator in order to control for differences in the level of expenditure ratios and focus on the structure of expenditures. The purpose of this analysis is to verify whether our sample countries have similarities in terms of expenditures with EU members, starting with the best-performing. Second, after obtaining conclusions on divergence in the descriptive analysis, we pool the former Yugoslavian countries with new EU member states for a cluster analysis to confirm whether there are identifiable common patterns among lower-performing EU member states. On average, these two groups of countries have similar expenditure shares over GPD and trajectories over time (figure 1), which makes them more suitable for pooling in a cluster analysis. Third, in the last stage of the convergence analysis, we use an analysis analogue to a β-convergence, where we look solely at the former Yugoslavian countries.

Figure 1
Public expenditure as a share of GDP (%)



2.1 COMPOSITION OF PUBLIC EXPENDITURES

Firstly, we compare the figures for each former Yugoslavian country with the average of the EU-15[3] countries in 2019. To focus on the composition of public expenditure controlling for differences in the size of total public expenditure, figures over total expenditures are used now. Social Protection expenditures in former Yugoslavian countries are below the average of EU-15 countries. Concerning Health expenditures, most of them follow the EU average, except for Montenegro. The remaining two-digit expenditure items are General Public Affairs and Education. Bosnia and Herzegovina, Montenegro, and Serbia seem to deviate from the core EU average by having higher expenditures on the former and lower on the latter. In terms of Economic Affairs, all former Yugoslavian countries except for Bosnia and Herzegovina exceed the EU-15 average. This category is the heart of

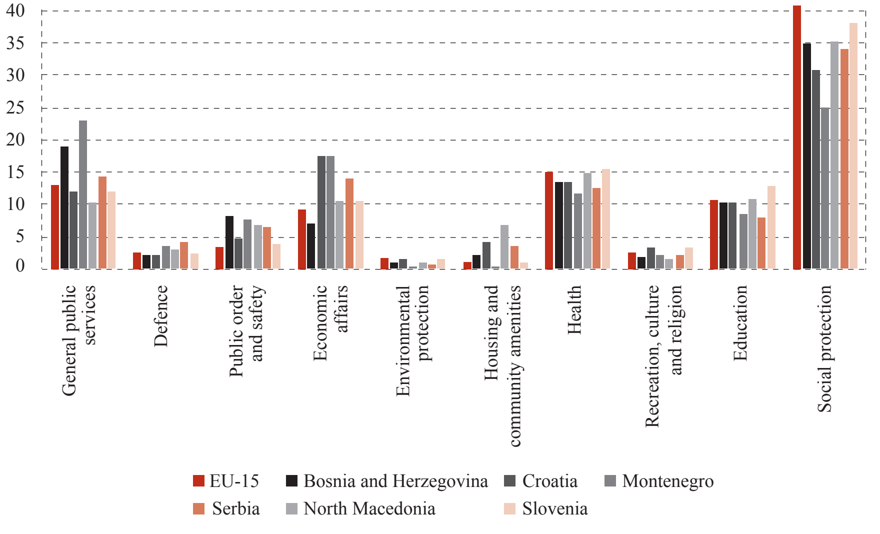
[3] The EU-15 countries, also known as core EU countries, are: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and United Kingdom.

PUBLIC SECTOR
ECONOMICS
47 (2) 233-254 (2023)

productive spending,[4] which is inherent to developing countries. Another indicator of overspending in the public sector is the Public Order and Safety category in all countries except Slovenia. Similarly, Housing expenditures are well above the EU-15 average, except in Slovenia and Montenegro. All former Yugoslavian countries spend below the core EU average on Environmental Protection, indicating that environmental awareness should be increased. Defence expenditures and Recreation, Culture, and Religion expenditures in former Yugoslavian countries seem to have the smallest deviations from the EU-15 averages.

Figure 2

Expenditure categories as a share of total expenditures in 2019, former Yugoslavian countries vs core EU countries (EU-15), %



2.2 CLUSTER ANALYSIS

We merge the countries of interest with a set of EU-13 countries[5] to perform a cluster analysis and test whether common patterns in financial efforts made in the different areas of expenditure can be detected. Hence, the final sample includes 17 countries, of which 13 are new EU member states. The remaining four are former Yugoslavian countries, while Slovenia and Croatia are listed among the new EU member states since they joined in 2004 and 2013, respectively. We focus on both total expenditure and its disaggregation in 10 functions.[6] In all cases, figures are expressed as a share of GDP. If common patterns among former Yugoslavian countries are strong enough, the existence of a cluster inside the pool should be reflected, despite

[4] The category mainly includes labour, agriculture, forestry, fishing, fuel, energy, mining, manufacturing, construction, transport, communication, other industries, and R&D.

[5] EU-13 is a group of 13 countries that have joined the EU since 2004: Bulgaria, Croatia, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia, and Slovenia.

[6] The 10 expenditure categories as per COFOG are: General Public Services; Defence; Public Order and Safety; Economic Affairs; Environmental Protection; Housing and Community Amenities; Health; Recreation, Culture and Religion; Education; and Social Protection.

the difference in degrees of EU integration among former Yugoslavian countries. The computations were made using the Stata 14 statistical package.

We perform two cluster analyses for the first (2011) and final (2019) years of available data. Computations use the average linkage clustering method and Euclidean distance as the similarity or dissimilarity measure. The corresponding dendrograms are presented in figures 3 and 4. Results for 2011 show that former Yugoslavian countries are not grouped in one specific cluster. The same picture emerges in 2019. In short, there are significant differences in the composition of public expenditure across countries and they do not clearly tend to decrease (or increase) over time.

Figure 3

Dendrogram of cluster analysis for 2011

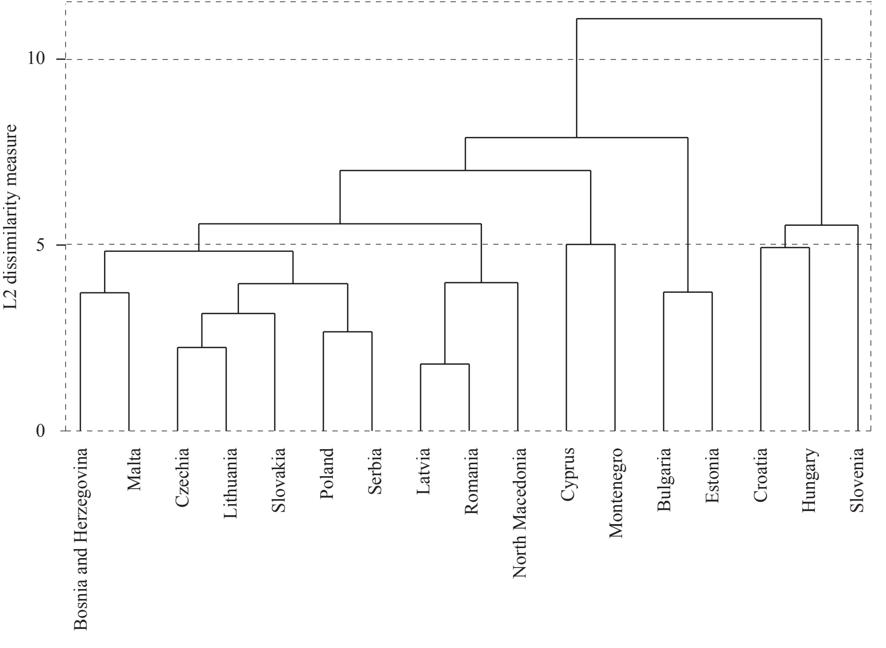
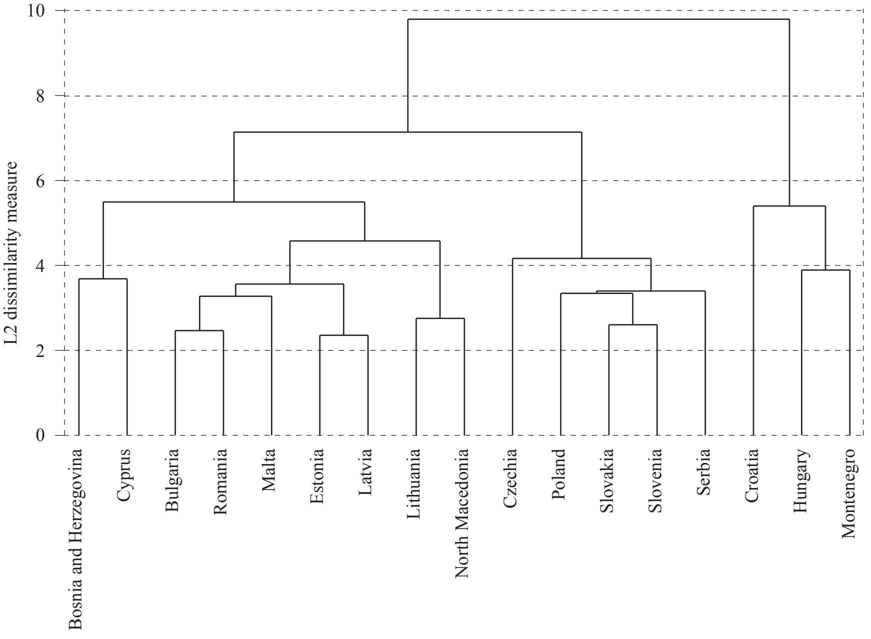


FIGURE 4
Dendrogram of cluster analysis for 2019



2.3 CONVERGENCE ANALYSIS

In order specifically to test convergence in public expenditures among the six former Yugoslavian countries, we rely upon the concept of β-convergence (Sala-i-Martin, 1996). However, the methodology of Sala-i-Martin (1996) is not followed *strictu sensu* since we do not employ log values, and rather use ratios of expenditure over GDP. Namely, we analyse convergence by relating variations in the corresponding ratio on the starting point. We focus again on both total expenditure and its disaggregation into 10 functions in terms of GDP. The panel dataset is balanced, with the initial year of the time series being 2011 and the final one 2019. The benchmark specification is the following one, where E – public expenditures over GDP; t – current year; α – constant term:

$$E_t - E_{t\text{-}1} = \alpha + \beta \times E_{t\text{-}1} + \varepsilon_t \quad (1)$$

Convergence is confirmed when the coefficient on the lagged expenditure is negative and significant, meaning lower staring values tend to involve higher increases in the ratio, and vice versa. The main results are reported in table 1. There is statistically significant convergence in seven categories and non-significant convergence in the remaining three. Overall government expenditures are convergent at a marginal significance, with a p-value of 0.10. In sum, there is empirical evidence of limited convergence of expenditures among former Yugoslavian countries.

Table 1

β-convergence in public expenditure

	Total expenditures	General public services	Defence	Public order and safety	Economic affairs	Environmental protection
Intercept	4.35 (1.55)	0.77 (1.8)*	0.22 (2.47)**	0.13 (1.18)	2.96 (3.93)***	0.05 (1.55)
Et-1	-0.11 (1.67)	-0.11 (1.8)*	-0.21 (2.92)***	-0.06 (1.6)	-0.54 (4.25)***	-0.16 (2.21)**

	Housing and community amenities	Health	Recreation, culture and religion	Education	Social protection
Intercept	0.04 (0.61)	0.95 (2.28)**	0.08 (2.11)**	0.17 (0.92)	0.84 (1.47)
Et-1	-0.04 (0.75)	-0.17 (2.35)**	-0.1 (2.98)***	-0.06 (1.36)	-0.07 (1.74)*

Obs.: 53; Method: OLS.

*Note: ***, **, * indicate statistical significance at 1%, 5% and 10%, respectively. T-statistics computed using OLS residuals is reported in parenthesis.*

3 SENSITIVITY OF PUBLIC EXPENDITURE TO THE ECONOMIC CYCLE

The literature on the cyclicality of fiscal policy makes a clear distinction between developing and developed countries. In a seminal paper, Gavin and Perotti (1997) argued that fiscal policies tend to be countercyclical in industrial (developed) countries and procyclical in developing countries. The main conclusion to be drawn from recent literature is that developing countries always tend to have a more procyclical fiscal policy than developed countries. Although developed countries are usually characterized by countercyclical fiscal policy, they may have a procyclical fiscal policy as well. However, the procyclical effect of the business cycle on fiscal policy is always stronger in developing countries.

Viren (2014) analysed EU-15 countries over the period 1970-2011 to find that expenditure elasticities with respect to output growth appear at -0.58 for a one-year horizon. Furthermore, in the analysis of Mourre, Astarita and Princen (2014), expenditure elasticity for European countries is also negative at -0.50. Both results are negative, suggesting countercyclicality. Égert (2014) found that discretionary fiscal policy in OECD countries is neutral or mildly countercyclical. Balassone and Kumar (2007) also used a mixed sample of industrial and developing countries between 1975-1997 to find clear evidence of how policy tends to be a cyclical or mildly countercyclical in bad times but procyclical in good times. Nevertheless, there is evidence that in many emerging market countries, fiscal policy is procyclical in bad times as well. They argue that this is due to inaccurate assessments of the economic cycle as well as a lack of access to external financial markets in crisis times, when funding is needed to exercise a countercyclical policy. The latter point was made first by Gavin and Perotti (1997), who claimed it was easier for developed countries to adopt expansionary fiscal policies during recessions (countercyclical) due to better access to financial markets. Inchauste et al. (2004), using data for 51 developing countries around the world between 1970 and 2002, confirmed that there is little scope for countries with high financial risks to undertake countercyclical fiscal policies.

When it comes to developing countries, results usually imply procyclicality. However, there are a limited number of studies solely exploring these countries. An exception is the work of Alesina, Campante and Tabellini (2008), who analyse 87 countries, where in many developing countries fiscal policy is procyclical. Furthermore, the procyclicality is driven by expenditures rather than revenues. Usually, the cyclicality of fiscal policy in developing countries is analysed within larger samples. The procyclical effect in large samples was first decomposed by Kaminsky, Reinhart, and Végh (2004), who found that fiscal policy was procyclical among low- and middle-income countries. Jaimovich and Panizza (2007) confirmed their findings. Although they found fiscal policy to be procyclical for all 118 developing and industrial countries in their sample over the 1970-2003 period, they did make one very clear differentiation. The strongest procyclical behaviour was found in low-income countries, with a statistically significant coefficient of 1.49, followed by middle-low income countries, and the lowest coefficient of 0.31 for procyclical fiscal policy was found in middle–high income countries. Ilzetzki and Vegh (2008) confirm the previous findings by using quarterly data from Q1 1960 to Q4 2006 for 27 developing and 22 high-income countries. Procyclicality in public spending is present in developing countries and, moreover, there is evidence of reverse causality. The procyclical spending in developing countries implies that fiscal policy influences the business cycle negatively. Dabla-Norriset et al. (2010) also confirm expenditure procyclicality in countries with low-quality budget institutions, while they find the same procyclicality to be lower in countries with high-quality budget institutions.

Concerning the mixed sample of former Yugoslavian countries, the expenditures have been shown to be countercyclical. The output gap elasticity values with respect to expenditures are pronounced, being at -1.27 or -1.20, depending on the estimate (Crnogorac and Lago-Peñas, 2019a). Similar findings were discovered by Arčabić and Banić (2021), who confirmed that Croatian policymakers try to keep fiscal policy both sustainable and countercyclical, but only in expansionary times.

We apply the same methodology when analysing the sensitivity of public finance, which is a simple and frequently used one for computing elasticities. The output gap, computed using the Hodrick and Prescott (1997) filter with the parameter k set at 4 according to the Ravn and Uhlig (2002) frequency rule, is chosen as a measure of the economic cycle. The variable was expressed through a mathematical statement as: GDP series over the filtered GDP series, minus 1, in brackets, and multiplied by 100. We used the GDP deflator in order to obtain figures in constant prices and avoid distortion by inflation. The panel dataset is unbalanced, starting in 2005 and ending in 2019. Hence, the econometric specification to be estimated is the following:

PUBLIC SECTOR
ECONOMICS
47 (2) 233-254 (2023)

$$E_t = \alpha_i + \beta \times output\ gap_{it} + \rho \times E_{t\text{-}1} + \varepsilon_{it} \quad (2)$$

where E is the category of expenditure analysed, α is for country fixed effects, β is the coefficient of the independent variable, ρ_i is the coefficient of the lagged dependent variable, and ε_{it} is the random error. Both public expenditure and the output gap are expressed as ratios over GDP.

The analysis was performed using the ordinary least square (OLS) estimators. However, this method may lack in a relevant shortcoming regarding endogeneity. Including a lagged dependent variable alongside fixed effects may lead to the so-called Nickell bias (Nickell, 1981), which is of the order 1/T. Since T is not very small in the case of our sample, we turn to a panel GMM estimator.[7] Several specification tests were performed to check the robustness of the results to potential endogeneity of the output gap in OLS and GMM estimators. Regarding the redundancy of both individual and period fixed effects, the corresponding tests confirmed the relevance of individual fixed effects, while period fixed effects could be omitted. Cross-section slope homogeneity was tested using a Wald test, which assumes common slopes in the null hypothesis. There is no cross-country homogeneity, confirming the previous conclusion from the convergence and cluster analyses that convergence in expenditures does not hold over time. The null hypothesis of the Pesaran-CD test that the error component may be cross-sectionally correlated is rejected (p value = 0.14). The Breusch-Godfrey test of AR(1) autocorrelation shows that it is not a serious concern. Taking into account the small N dimension of the panel dataset, we employed both the Hansen and the Arellano-Bond AR(2) autocorrelation tests in the GMM estimate. The test outcomes validate the results, meaning instruments are uncorrelated with the error term and there is no autocorrelation. The main results, obtained through equation (2), are reported in table 2.

Firstly, it is shown that the ratio of expenditures to GDP is negatively influenced by the economic cycle. The negative value of the coefficient implies a countercyclical fiscal policy in the considered period. GMM estimates support the OLS results. Secondly, after replacing the total expenditures in equation (1) with their COFOG categories, we find the output gap is significant in only three categories: Health, Education, and Social Protection. Together, these amount to a 57% share of total expenditures over the period (13, 11, and 33% respectively). The coefficient on Social Protection expenditures is significant and negative, confirming the countercyclicality. We could argue that Social Protection expenditures decrease when the output gap is positive, because employment is higher in good economic times. Coefficients on Health and Education expenditures are lower in intensity

[7] In preliminary estimates, we also tried to use the maximum likelihood with structural equation modelling (ML SEM). This estimator is employed only with balanced panels where T is relatively small (e.g. less than 10) and there are no missing data (Allison, Williams and Moral-Benito, 2017; Moral-Benito, Allison and Williams, 2019). It could be used to replace GMM with datasets where N is less than 100. However, the shortness of the sample caused lack of convergence in estimates.

and marginally significant. This can be justified by the nature of these expenditures, since the healthcare and education systems imply the existence of long-term strategies that do not sustain frequent changes. The results concerning these categories go in hand with the socialist history of the sample countries, Yugoslavia being a strong welfare state. Thirdly, three categories show ambivalent results regarding coefficient significance across the OLS and GMM estimates: General Public Services; Public Order and Safety; and Recreation, Culture, and Religion.

Table 2
Estimate of equation 1

	Total expenditures		General public services		Defence		Public order and safety	
Method	**OLS**	**GMM**	**OLS**	**GMM**	**OLS**	**GMM**	**OLS**	**GMM**
Output gap	-0.55 (-3.01)***	-0.67 (-4.11)***	-0.12 (-1.56)	-0.14 (-2.28)**	0.02 (1.4)	0.03 (1.49)	-0.02 (-1.24)	-0.03 (-2.41)**
Lagged endogenous	0.37 (3.11)***	0.33 (7.33)***	0.5 (4.73)***	0.66 (3.71)***	0.77 (10.21)***	0.74 (7.84)***	0.63 (6.55)***	0.61 (3.14)***
R2	0.85		0.87		0.82		0.93	
Individual fixed effects	Yes (0.0009)		Yes (0.008)		Yes (0.0717)		Yes (0.0297)	
Period fixed effects	No		No		Yes (0.0772)		No	
Wald	0.0118							
Pesaran CD	0.1350							
B-G test	0.8532							
Hansen test		0.3052		0.6803		0.6356		0.7494
A-B AR(2)		0.3101		0.6226		0.1620		0.1576
Observations	64	58	64	58	64	58	64	58
	Economic affairs		**Environmental protection**		**Housing and community amenities**		**Health**	
Method	**OLS**	**GMM**	**OLS**	**GMM**	**OLS**	**GMM**	**OLS**	**GMM**
Output gap	-0.07 (-0.48)	0.06 (0.51)	0.002 (0.22)	0.001 (0.15)	-0.01 (-0.45)	0.01 (0.27)	-0.07 (-2.57)**	-0.1 (-5.61)***
Lagged endogenous	0.2 (1.59)	0.1 (0.68)	0.24 (2.16)**	0.48 (3.09)***	0.35 (3.13)***	0.3 (1.79)*	0.42 (3.56)***	0.12 (0.82)
R2	0.45		0.88		0.94		0.79	

	Economic affairs		Environmental protection		Housing and community amenities		Health	
Method	**OLS**	**GMM**	**OLS**	**GMM**	**OLS**	**GMM**	**OLS**	**GMM**
Individual fixed effects	Yes (0.0128)		Yes (0.0000)		Yes (0.0000)		Yes (0.0029)	
Period fixed effects	No		No		No		No	
Hansen test		0.8943		0.0810		0.1551		0.2843
A-B AR(2)		0.2414		0.2150		0.8262		0.1292
Observations	64	58	64	46	64	58	64	58

	Recreation, culture and religion		Education		Social protection	
Method	**OLS**	**GMM**	**OLS**	**GMM**	**OLS**	**GMM**
Output gap	-0.004 (-0.33)	-0.01 (-2.7)**	-0.03 (-1.68)*	-0.03 (-2.23)**	-0.21 (-3.65)***	-0.27 (-1.96)*
Lagged endogenous	0.59 (5.08)***	0.34 (4.04)***	0.72 (9.04)***	0.74 (14.52)***	0.5 (4.45)***	0.23 (0.86)
R^2	0.92		0.94		0.93	
Individual fixed effects	Yes (0.0513)		Yes (0.0484)		Yes (0.0051)	
Period fixed effects	No		No		No	
Hansen test		0.0921		0.0662		0.4226
A-B AR(2)		0.3618		0.5178		0.1556
Observations	64	46	64	58	64	46

*Note: Redundancy F-test p-values for individual fixed effects and period fixed effects is reported in respective parenthesis. ***, **, * indicate statistical significance at 1%, 5% and 10%, respectively. T-statistics computed using OLS residuals is reported in parenthesis. Instruments in GMM are constructed for the lagged dependent variable from the second and third lagged values in the form of differences in all estimates, except for Health, Environmental protection and Recreation, culture and religion where we use the third and fourth lagged values. Additionally, third and fourth lagged values in the form of levels are applied as instruments for Recreation, culture and religion, Environmental protection and Social Protection. Lastly, period fixed effects are applied in the case of Housing and community amenities, Environmental protection and Social Protection.*

4 ON THE DETERMINANTS OF EXPENDITURE COMPOSITION

The aim of our last analysis is to shed new light on the determinants of the composition of public expenditures in former Yugoslavian countries. As a first step, we revise the existing literature in order to discover which would be the best determinants of expenditure functions. In table 3 below, we have synthetized some of the most relevant recent articles that analysed the determinants of COFOG expenditures. Some of the papers define the endogenous variables as a share of total expenditures (del Granado, Martinez-Vazquez and McNab, 2018; Cordis, 2014; Hessami, 2014; Enkelmann and Leibrecht, 2013; Lago-Peñas and Lago-Peñas, 2009), while others choose to define them as a share of GDP (Mauro, 1998; Ferreiro, Garcia-Del-Valle and Gomez, 2009). Most of the papers focus on one or several variables of interest, such as corruption, decentralization, GDP, or electoral and political variables. Concerning control variables, the most usual are ageing, population size, unemployment rate, and total expenditures.

Table 3
Specifications of models estimating determinants of expenditure functions

Dependent variable	Independent variable of interest	Controls	Source
education (ratio of education expenditures to total public expenditures) health (same) and composition (ratio of education and health expenditures to total public expenditures)	decentralization (share of local to total exp)	population population density age structure gross domestic product (GDP) per capita openness to international trade OECD membership dummy	del Granado, Martinez-Vazquez and McNab (2018)
expenditure share in total expenditures	corruption convictions	log of state population the log of real state gross domestic product (GDP) per capita the percentage of state population aged 25+ with at least a high school diploma the percentage of state population between ages 0-17	Cordis (2014)
expenditures share of total expenditures	corruption perception index	interest rate on government bonds the population density the age-dependency ratio the log of real GDP per capita and the unemployment rate regional dummies for the South, the Midwest, and the West	Hessami (2014)

PUBLIC SECTOR ECONOMICS
47 (2) 233-254 (2023)

Dependent variable	Independent variable of interest	Controls	Source
total expenditure / COFOG category	real GDP in national currency election years	lagged dependent population openness the age structure (share of young and elderly in total population) unemployment rate growth rate of total expenditures	Enkelmann and Leibrecht (2013)
expenditure (share of GDP)	ln(population) ln(GDP per capita) openness OECD membership index of ethnic fractionalization fraction of population over 65		Ferreiro, Garcia-Del-Valle and Gomez (2009)
government transfers to households	party linkage	lagged dependent age unemployment total expenditures federalism dummy district magnitude	Lago-Peñas and Lago-Peñas (2009)
education expenditures	corruption index		Mauro (1998)

Source: Authors' literature review.

As the dependent variable, we rely upon category over GDP. When attempting to estimate the 10 COFOG categories, we opted to aggregate them in two main aggregates. Housing and community amenities; Health; Recreation, culture, and religion; and Social protection are aggregated as Social expenditure (SE). Second, Economic affairs and education are aggregated into Productive expenditure (PE)[8]. As for independent variables, they were grouped as main and control variables. Decomposition of equation [3] and a summary of all variables in table 4 are shown below.

$$E_{it} = \alpha_i + \beta \times DETERMINANTS_{it} + \gamma \times CONTROLS_{it} + \varepsilon_{it} \qquad (3)$$

E – expenditure aggregation of interest (SE or PE)
DETERMINANTS – group of main variables (ELECTIONS, PARTY_NAT, CORRUPT, GDPEUR)
CONTROLS – group of control variables (OPEN, UNEM, TOTEXP)

[8] In the preliminary testing phase, we conducted regressions using a variable for public goods and general administration (PGGA), that included the remaining COFOG categories of General public services, Defence, Public order and safety, and Environmental protection. Nevertheless, the determinants and controls showed not to be significant for this spending category that has a steady level over time in all countries. Therefore, we opted not to include it in the final results table.

α – constant coefficient
β, γ – coefficients of the independent variables
ε_{it} – error term.

Table 4
Variable description and data sources

Variable	Definition	Description	Source
PGGA	Public goods and general administration	Aggregation of General public services, Defence, Public order and safety, and Environmental protection, measured in GDP %	National Financial Institutions, Eurostat
SE	Social expenditure	Aggregation of Housing and community amenities, Health, Recreation, culture and religion and Social protection, measured in GDP %	National Financial Institutions, Eurostat
PE	Productive expenditure	Aggregation of Economic affairs and Education, measured in GDP %	National Financial Institutions, Eurostat
ELECTIONS	Parliamentary elections	Dummy variable taking value of 1 for parliamentary election year, and 0 for other	National Election Commissions
PARTY_NAT	Party nationalization index	The party nationalization index calculated as per Bochsler (2010), where politically decentralized countries are described with a lower index value	National Election Commissions
CORRUPT	Corruption perception index	The corruption perception index, where less corrupt countries are described by a higher index value	Transparency International
GDPEUR	Gross domestic product	Gross domestic product in millions of euros, at current market prices	Eurostat
OPEN	Trade openness	Trade openness index, where higher values are described for more open economies	Eurostat
UNEM	Unemployment rate	Percentage of population aged over 15 years that is unemployed	International Labour Organization
TOTEXP	Total expenditures	Total expenditures, measured in GDP %	National Financial Institutions, Eurostat

There are four main variables in our sample. The elections variable is a dummy to account for the effects of potential additional spending in election years. Second, the party nationalization index is included in the specification to measure the effect of territorial diversity of the party system across territorial units. It is an indicator that shows on party level whether the support for a political party in an election is equally distributed across the territory of a country. We use the standardized party nationalization score by Bochsler (2010), which is calculated with the use of a template provided by that author. By applying data from each parliamentary election at municipal level[9], we obtain an index value between 0 and 1 for all our countries. A higher value of the obtained party nationalization index would imply a homogeneous party system with similar results at country and local level. Third, we include the corruption perception index published annually by Transparency International to account for the effects of corruption. The variable takes values between 0 and 10 for each country, where a higher corruption index value implies less corruption in a country. The last determinant variable is GDP, which is obtained from Eurostat and is used to account for the influence of economic success on public expenditures.

As for control variables, there are three commonly used ones. Trade openness is a control variable obtained as the sum of exports and imports divided by the GDP of a country. The amount of trade should make it possible to see how an open economy influences the public spending structure. The unemployment rate aims to capture the effects of the business cycle. Lastly, total expenditures as a percentage of GDP are used as a control variable to check the influence of the level of public spending in each country on its composition. Table 5 provides the summary statistics of all variables.

Table 5

Summary statistics

Variable	Mean	Median	Std. dev.	Maximum	Minimum	Observations
PGGA	10.71	9.94	2.97	19.35	6.8	70
PE	10.2	10.06	2.67	21.7	5.3	70
SE	22.94	23.38	3.14	29.1	17.8	70
ELECTIONS	0.27	0	0.45	1	0	70
PARTY_NAT	0.8	0.8	0.11	0.92	0.54	70
CORRUPT	4.55	4.2	0.94	6.7	3.2	70
GDPEUR	26,949.2	34,376.4	16,582.47	54,237.9	3,125.1	70
OPEN	1.07	1.03	0.23	1.61	0.66	70
UNEM	16.59	16.16	7.8	32.18	4.37	70
TOTEXP	43.85	44.91	5.28	60.3	33.1	70

To test for potential multicollinearity, variance inflation factors (VIF) were computed. Only variables with VIF values below 5 were considered acceptable for the final specification. A Breusch-Godfrey test on residuals discarded autocorrelation problems. In order to deal with potential cross-section heteroscedasticity and contemporaneous correlation, both OLS residuals and panel-corrected standard errors

[9] We use data on parliamentary elections in former Yugoslavian republics from 2005 to 2018. The estimates are made using territory data from lower level units, such as municipalities instead of districts and/or cities. For example, the number of local units in the latest elections up to 2018 for each country was 144 for Bosnia and Herzegovina (2018), 560 for Croatia (2016), 88 for North Macedonia (2016), 24 for Montenegro (2016), 180 for Serbia (2016), and 88 for Slovenia (2018).

were computed and are shown in table 6. Redundancy of both individual and period fixed effects was rejected in all three cases, justifying their presence.[10]

Table 6

Determinants of COFOG expenditures

	SE	PE
C	12.8484 (2.1)** [1.67]	-32.065 (-3.2)*** [-2.62]**
ELECTIONS	0.498 (2.25)** [2.09]**	-0.3541 (-0.98) [-0.83]
PARTY_NAT	16.8527 (3.49)*** [3.21]***	-4.8893 (-0.62) [-0.56]
CORRUPT	-0.3769 (-1.2) [-1.13]	0.2448 -0.48 [0.46]
GDPEUR	-0.0001 (-1.66) [-1.88]*	0.0001 -0.92 [1.28]
OPEN	-0.8379 (-0.39) [-0.38]	0.1487 -0.04 [0.05]
UNEM	-0.1432 (-3.05)*** [-3.17]***	0.2187 (2.84)*** [2.46]**
TOTEXP	0.0992 (2.09)** [1.94]*	0.8844 (11.39)*** [10.77]***
Observations	70	70
Individual fixed effects	>0.0001	>0.0001
Period fixed effects	>0.0001	0.003
B-G test	0.66	0.16
R-squared	0.97	0.91

*Note: Robust t-statistics computed using panel corrected standard errors (PCSE) are reported in brackets, whereas t-statistics computed using OLS residuals are reported in parenthesis. ***, **, * indicate the statistical significance at 1%, 5% and 10%, respectively. Redundancy F-tests for individual fixed effects and period fixed effects are reported in respective rows. B–G test denotes the Breusch-Godfrey test of AR(1) autocorrelation. Estimates are performed using EViews 9.5.*

The main results are shown in table 6. The election dummy is statistically significant, indicating that it explains the social expenditures. Furthermore, the sign is positive, meaning that social expenditures are increased during electoral cycles. This makes sense since the consequences of such spending are more visible to the voters than the other two aggregates; the governing party can potentially make a good impression

[10] In order to ensure robustness of our results, SUR (Seemingly Unrelated Equations) estimates were also made. The results were similar.

PUBLIC SECTOR
ECONOMICS
47 (2) 233-254 (2023)

before new elections by increasing the social benefits. Social expenditures are affected by political party nationalization, while productive expenditure is not. Results regarding the corruption index are not statistically significant. A negative relationship between GDP and social expenditures is confirmed at 10% statistical significance. This is expected for two reasons. First, the expenditures are expressed as a percentage of GDP, where an increase of the GDP in levels would mean a decrease in the share of social expenditures. Second, if a country is experiencing economic growth, its expenditures on social causes are expected to decrease. The coefficient of trade openness has the expected positive sign in relation to productive expenditures. Exports and imports are important components of demand which would require more productive spending from a country budget. However, this expected outcome is not confirmed by statistical significance. The unemployment rate is relevant in explaining the movements of both social and productive expenditures. A positive sign in the latter case implies that a higher unemployment rate would require more productive spending. This is in line with our findings in the previous section, where we found that the expenditure policy in all former Yugoslavian countries is countercyclical. Therefore, an expected reaction in a recession when the unemployment is high would be to increase productive spending.

Lastly, it is shown that the social and productive expenditures are positively influenced by the growth of the share of total expenditure in GDP. This finding is interesting insofar as it reveals which aggregates of the COFOG expenditure categories are the ones driving the total expenditures. Furthermore, the findings from the previous section also confirm that most of the COFOG categories belonging to the social and productive expenditure aggregates are influenced by the economic cycle. The changes in GDP are relevant for explaining the changes in total expenditures and their driving components.

5 CONCLUSIONS

The extensive research provided in this paper has covered three aspects of public expenditures in former Yugoslavian countries. First, the convergence analysis confirmed our hypothesis that the public expenditures in these countries between 2011 and 2019 diverge in terms of dynamic, static, and structural aspects. Second, analysis of the sensitivity of public expenditures and the COFOG categories to the economic cycle confirmed that public spending was countercyclical in all former Yugoslavian countries between 2005 and 2019, with Health, Education, and Social Protection expenditures driving the results. Third, the determinants of two main COFOG expenditure aggregates (social expenditure and productive expenditure) were explored. Social expenditures are increased in election years with the expectation of attracting voters to the policymakers awarding them. Furthermore, more social expenditures are expected to be made in less diverse party systems. Nevertheless, social spending is expected to decrease its share in expenditures in a booming economy, which is characterised by a growing GDP and shrinking unemployment. Lastly, productive expenditures are increased when unemployment is high, in order to combat recessions, which confirms our countercyclicality findings.

PUBLIC SECTOR
ECONOMICS
47 (2) 233-254 (2023)

Finally, this extensive research has opened the door to multiple interesting conclusions for policymakers and further research. Making the COFOG data for the Western Balkan subgroup of former Yugoslavian countries available in common databases would be the main step in that direction. The intriguing results on countercyclicality in the period where the neighbouring EU countries were mainly following a pro-cyclical fiscal policy would be interesting to check. Whereas this may not be a typical finding for developing countries, the social, health and education expenditures explain how acting as automatic stabilisers they drive the expenditures against the economic cycle. Furthermore, in recession times it is a much more difficult (and unpopular) task to implement austerity measures for policymakers (especially in electoral cycles), complementing the said conclusion. Social expenditures seem to reflect changes in various aspects of a country's economy, having a major impact on overall expenditures. Insofar as the determinants are concerned, the political party data on elections is currently the best proxy for government decentralization in some countries. Could it be that the legacies of the former Yugoslav centralized and socialist model are still present to this day? Are the interactions they are having with the European integration process the ones that determine the economic development of these countries? There are several conclusions pointing in this direction: the non-clustering with new EU member states; weak convergence of public expenditures among the ex-Yu, predominant influence of social expenditures; strong political aspect in explaining the increase in expenditures; and less diverse parliamentary systems (more authoritarian?) being the ones that can cope with increased spending. Now that certain expenditure categories have been identified as the main tools in the hands of policymakers, a more detailed analysis will be necessary in the future. Putting more emphasis on the effects of public spending would help achieve the required discipline and discover how resource allocation can be optimized.

Disclosure statement
The authors declare that there is no conflict of interest.

REFERENCES

1. Afonso, A. and Jalles, J. T., 2014. Fiscal composition and long-term growth. *Applied Economics*, 46(3), pp. 349-358. https://doi.org/10.1080/00036846.2013.848030
2. Alesina, A., Campante, F. R. and Tabellini, G., 2008. Why is fiscal policy often procyclical? *Journal of the European Economic Association,* 6(5), pp. 1006-1036. https://doi.org/10.1162/JEEA.2008.6.5.1006
3. Allison, P. D., Williams, R. and Moral-Benito, E., 2017. Maximum likelihood for cross-lagged panel models with fixed effects. *Socius,* 3. https://doi.org/10.1177/2378023117710578
4. Arčabić, V. and Banić, F., 2021. Characteristics of fiscal policy in Croatia: does it depend on the phase of the business cycle? *Public Sector Economics*, 45(4), pp. 433-457. https://doi.org/10.3326/pse.45.4.3
5. Balassone, F. and Kumar, M., 2007. *Promoting fiscal discipline.* Washington: International Monetary Fund.
6. Bochsler, D., 2010. Measuring party nationalisation: A new Gini-based indicator that corrects for the number of units. *Electoral studies,* 29(1), pp. 155-168. https://doi.org/10.1016/j.electstud.2009.06.003
7. Cordis, A. S., 2014. Corruption and the composition of public spending in the United States. *Public Finance Review*, 42(6), pp. 745-773. https://doi.org/10.1177/1091142114531320
8. Crnogorac, M. and Lago-Peñas, S., 2019a. Fiscal policy in former Yugoslavian countries (2001-2014): stylised facts and budget elasticities. *Economic Research,* 32(1), pp. 947-967. https://doi.org/10.1080/1331677X.2019.1583584
9. Crnogorac, M. and Lago-Peñas, S., 2019b. Tax evasion in the countries of Former Yugoslavia. *Zbornik radova Ekonomskog fakulteta u Rijeci: časopis za ekonomsku teoriju i praksu,* 37(2), pp. 823-851. https://doi.org/10.18045/zbefri.2019.2.823
10. Crnogorac, M. and Lago-Peñas, S., 2020. Determinants of tax morale in former Yugoslavian countries. *Eastern European Economics*, 58(2), pp. 174-196. https://doi.org/10.1080/00128775.2019.1671868
11. Dabla-Norris, E. [et al.], 2010. Budget institutions and fiscal performance in low-income countries. *IMF working papers*, No. 240. https://doi.org/10.5089/9781455209408.001
12. del Granado, F. J. A., Martinez-Vazquez, J. and McNab, R. M., 2018. Decentralized governance, expenditure composition, and preferences for public goods. *Public Finance Review*, 46(3), pp. 359-388. https://doi.org/10.1177/1091142116639127
13. Égert, B., 2014. Fiscal policy reaction to the cycle in the OECD: pro-or counter-cyclical? *Mondes en developpement*, (3), pp. 35-52. https://doi.org/10.3917/med.167.0035
14. Enkelmann, S. and Leibrecht, M., 2013. Political expenditure cycles and election outcomes: Evidence from disaggregation of public expenditures by economic functions. *Economics Letters*, 121(1), pp. 128-132. https://doi.org/10.1016/j.econlet.2013.07.015
15. Ferreiro, J., Garcia-Del-Valle, M. T. and Gomez, C., 2009. Is the composition of public expenditures converging in EMU countries? *Journal of Post Keynesian Economics*, 31(3), pp. 459-484. https://doi.org/10.2753/PKE0160-3477310305

16. Gavin, M. and Perotti, R., 1997. Fiscal policy in latin america. *NBER macroeconomics annual,* 12, pp. 11-61. https://doi.org/10.1086/654320
17. Hessami, Z., 2014. Political corruption, public procurement, and budget composition: Theory and evidence from OECD countries. *European journal of political economy*, 34, pp. 372-389. https://doi.org/10.1016/j.ejpoleco.2014.02.005
18. Hodrick, R. J. and Prescott, E. C., 1997. Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit, and Banking*, 29(1), pp. 1-16. https://doi.org/10.2307/2953682
19. Ilzetzki, E. and Végh, C. A., 2008. Procyclical fiscal policy in developing countries: Truth or fiction? *NBER Working paper*, No. 14191. https://doi.org/10.3386/w14191
20. Inchauste, M. G. [et al.], 2004. The cyclical and long-term behavior of government expenditures in developing countries. *IMF Working paper*, No. 202. https://doi.org/10.5089/9781451874433.001
21. Jaimovich, D. and Panizza, U., 2007. Procyclicality or Reverse Causality? *IDB Working Paper*, No. 501. https://doi.org/10.2139/ssrn.1820869
22. Kaminsky, G. L., Reinhart, C. M. and Végh, C. A., 2004. When it rains, it pours: procyclical capital flows and macroeconomic policies. *NBER macroeconomics annual*, 19, pp. 11-53. https://doi.org/10.1086/ma.19.3585327
23. Lago-Peñas, I. and Lago-Peñas, S., 2009. Does the nationalization of party systems affect the composition of public spending? *Economics of Governance*, 10(1), pp. 85-98. https://doi.org/10.1007/s10101-008-0051-x
24. Mauro, P., 1998. Corruption and the composition of government expenditure. *Journal of Public Economics*, 69(2), pp. 263-279. https://doi.org/10.1016/S0047-2727(98)00025-5
25. Moral-Benito, E., Allison, P. and Williams, R., 2019. Dynamic panel data modelling using maximum likelihood: an alternative to Arellano-Bond. *Applied Economics*, 51(20), pp. 2221-2232. https://doi.org/10.1080/00036846.2018.1540854
26. Mourre, G., Astarita, C. and Princen, S., 2014. Adjusting the budget balance for the business cycle: the EU methodology. *Economic Papers,* No. 536.
27. Nickell, S. J., 1981. Biases in dynamic models with fixed effects. *Econometrica*, 49(6), pp. 1417-1426. https://doi.org/10.2307/1911408
28. Ravn, M. O. and Uhlig, H., 2002. On adjusting the hodrick-prescott filter for the frequency of a observations. *The Review of Economics and Statistics*, 84(2), pp. 371-375. https://doi.org/10.2139/ssrn.273355
29. Sala-i-Martin, X., 1996. The classical approach to convergence analysis. *The economic journal*, 106(437), pp. 1019-1036. https://doi.org/10.2307/2235375
30. United Nations, 2000. Classifications of expenditure according to purpose: classification of the functions of government (COFOG), classification of individual consumption according to purpose (COICOP), classification of the purposes of non-profit institutions serving households (COPNI), classification of the outlays of producers according to purpose (COPP). *Statistical Papers*, No. 84.
31. Viren, M., 2014. How different are the fiscal policy effects? *Revue de l'OFCE*, (1), pp. 135-157. https://doi.org/10.3917/reof.132.0135

Fiscal decentralization and economic growth: evidence from Brazilian states

PEDRO JORGE HOLANDA FIGUEIREDO ALVES, Ph.D.*
JEVUKS MATHEUS ARAUJO, Ph.D., Professor*
ANA KAROLINA ACRIS MELO, Ph.D. student*
EDUARDA MASHOSKI, Ph.D.*

Article**
JEL: H7, H77
https://doi.org/10.3326/pse.47.2.5

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

** Received: June 30, 2022
Accepted: February 6, 2023

Pedro Jorge HOLANDA FIGUEIREDO ALVES
Catholic University of Brasilia, QS 07, Lote 01, Taguatinga Sul – Taguatinga, Brasília – DF, 71966-700, Brazil
e-mail: Pedrojorge_holanda@hotmail.com
ORCiD: 0000-0001-9340-030X

Jevuks Matheus ARAUJO
Federal University of Paraiba, Campus I Lot, Cidade Universitaria, PB, 58051-900, Brazil
e-mail: Jevuks@gmail.com
ORCiD: 0000-0002-5618-4502

Ana Karolina ACRIS MELO
Federal University of Paraiba, Campus I Lot, Cidade Universitaria, PB, 58051-900, Brazil
e-mail: ana.acris@live.com
ORCiD: 0000-0002-9688-1249

Eduarda MASHOSKI
Federal University of Paraiba, Campus I Lot, Cidade Universitaria, PB, 58051-900, Brazil
e-mail: Eduarda_machoski@hotmail.com
ORCiD: 0000-0001-7061-4727

Abstract

This paper investigates the relationship between fiscal decentralization and economic growth in Brazilian states from 1996 to 2015. Using five decentralization measures and the GMM-System model to address the endogeneity problem, we have identified a positive relationship between the indicators of fiscal decentralization and economic growth and observed that the industry and service sectors are the most affected by this decentralization. Our results suggest that local governments with more autonomy make states more efficient, thus increasing economic growth.

Keywords: decentralization, economic growth, Brazilian states, GMM-System

1 INTRODUCTION

Over the past few decades, developing countries have changed their institutional settings as one way of allocating more political power and fiscal autonomy to sub-national governments. Their policies follow the normative theory, which suggests that it is possible to improve the efficiency of the public sector and promote long-term economic development by decentralizing fiscal power. The familiarity of sub-national governments with local conditions and preferences is one of the key factors supporting this theory (Ma and Mao, 2018; Shah, 2006; Gadenne and Singhal, 2014; Jametti and Joanis, 2016).

The Brazilian Federal Constitution from 1988 provided a favorable movement towards decentralization, by delegating to the federated entities the responsibility to implement public policies, focusing on the particularities of local demands. Thus, the Constitution has separated the administrative functions into three levels of government, allowing states and municipalities to distribute taxes to promote local development. However, given the heterogeneity of the country and its great inequalities, decentralization may not have been as effective as expected.

From a theoretical perspective, there are mixed results regarding how governments could achieve the best outcome. Qiao, Ding and Liu (2019) argue that fiscal decentralization negatively affects the size of the government itself but that higher levels of democracy will mitigate these effects of fiscal decentralization. According to Christl, Köppl-Turyna and Kucsera (2020), it promotes public efficiency. Here, tax rules combined with decentralization impair efficiency, a phenomenon known as the ratchet effect. Finally, Colombo and Martinez-Vazquez (2020) relate higher levels of decentralization of expenditures and revenues to lower public levels of spending and R&D. There is also the work by Thanh and Canh (2020), showing that fiscal decentralization has had a positive effect on economic growth in regions where public governance is of high quality.

This article contributes to the literature on fiscal decentralization by addressing the specific issue of a developing country with characteristics conducive to fiscal decentralization. Aligned with the scenario of approaching a large developing country, we want to understand how it affected the growth rate of GDP in

Brazilian states between 1996 and 2015. Unlike most applied research, five decentralization measures have been used and while the endogeneity problem has been handled with the GMM-System model. In addition, we have advanced the debate on understanding the heterogeneous effects on economic sectors.

Our results show that decentralization has promoted the economic growth of Brazilian states, which is a result that corresponds to theoretical studies. The results suggest the main mechanism in this is the strengthening of fiscal autonomy. It should be emphasized that indicator PI is the most positively signed, implying a strong effect from the expenditure side on economic growth. In addition, the results show a greater effect on the service sector.

We must conduct our estimations carefully since economic growth and fiscal decentralization usually present endogeneity problems. Ligthart and van Oudheusden (2017) and Thanh and Canh (2020) reinforced this, with no evidence to reject the hypothesis that economic growth and fiscal decentralization are not endogenous[1]. As well as providing evidence about the relationship between economic growth and fiscal decentralization in the world context, our work is a pioneer on the subject in a national context.

This paper has seven parts, including this introduction. The following section summarizes the process of fiscal decentralization in Brazil and also presents recent data related to it. The third section presents an overview of empirical studies that address fiscal decentralization. In section four, we present the methodological aspects used in the article and describe the variables. In the fifth part, we discuss the main results and those by sector. The last section contains final considerations.

2 BACKGROUND

During the 1960s and 1970s, the Brazilian political system had fiscal centralization as its administrative model. Since the 1980s, however, after a significant expansion of public functions, the federal government has shared its financial resources and administrative responsibilities with the states and municipalities. During the process of democratization, the movement of decentralization increased, and, with the Federal Constitution from 1988, the Brazilian federation changed. However, as Araujo and Siqueira (2016) stated, this process was uncoordinated since the central government did not manage it properly. Therefore, sub-national governments controlled it and benefited from the new rules.

Decentralization in Brazil is a complex matter due to the serious socioeconomic and geographic inequalities that characterize the country. Regions such as the North and Northeast have, historically, greater inequality and lower economic growth. The simple process of decentralization does not necessarily allow local governments to be self-sufficient. Therefore, reconciling decentralization and reducing social

[1] The authors cite papers with comparable results.

PUBLIC SECTOR
ECONOMICS
47 (2) 255-280 (2023)

inequality are the main challenges. In Brazil, the latter was conducted mainly by the increase in national transfers and not from an increase in its tax collection capacity. In table 1, the percentage of municipal and state revenue is shown, according to the major regions of Brazil, in the period from 1985 to 2015.

Table 1

Distribution of municipal and state average revenues according to the regions of Brazil, 1985-2015, by percentage

Municipalities	**85-94**	**95-99**	**00-04**	**05-09**	**10-15**	**85-94**	**95-99**	**00-04**	**05-09**	**10-15**
Brazil	11.37	22.66	18.08	18.07	19.48	60.26	60.92	65.67	66.86	64.13
Midwest	9.53	24.84	12.11	12.98	15.71	66.83	75.12	74.28	72.31	71.90
Northeast	6.90	12.12	9.54	9.68	18.05	60.56	79.25	79.27	80.74	84.70
North	5.80	13.27	9.35	10.33	31.69	56.73	76.73	77.45	79.05	86.47
Southeast	22.04	26.43	23.81	23.84	19.75	50.59	52.64	58.44	59.42	49.62
South	12.57	19.86	15.19	15.41	17.40	66.57	67.31	65.41	65.88	71.21
States	**85-94**	**95-99**	**00-04**	**05-09**	**10-15**	**85-94**	**95-99**	**00-04**	**05-09**	**10-15**
Brazil	79.14	65.59	63.35	62.34	61.91	18.66	24.22	22.31	24.53	22.43
Midwest	59.79	50.11	58.70	62.58	58.44	31.08	40.97	27.74	22.75	20.98
Northeast	56.38	50.25	47.28	45.79	48.49	34.64	43.69	39.48	43.19	40.26
North	45.18	42.68	41.85	41.15	41.61	62.08	49.87	48.53	48.87	44.35
Southeast	82.70	77.23	72.10	71.02	70.09	9.07	13.13	11.95	13.99	11.94
South	84.69	62.59	67.33	67.21	68.29	10.50	18.34	19.61	22.35	19.90

Note: Data from the National Treasury Secretariat.

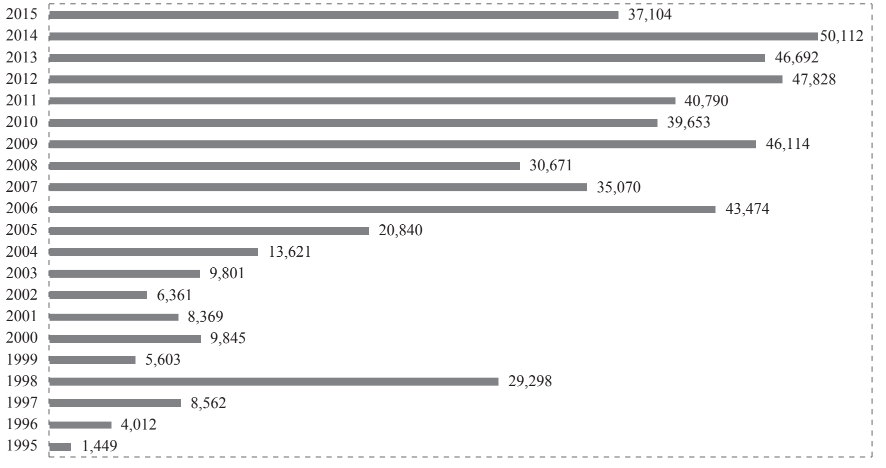
The tax revenues shown in table 1 reveal that the Brazilian municipalities and states had a meaningful change in their collected values. Compared to those before the Federal Constitution of 1988, the municipal tax revenues registered growth in four out of five Brazilian regions, particularly North and Northeast. We can highlight the North region, which obtained high state collections. Even reaching high rates, from 1995 to 2015, states and municipalities maintained their tax collection at similar levels, in total revenue.

The observed scenario is different concerning revenue from government transfers. In 1985, current transfers from Brazilian municipalities accounted for about 60% of the total municipal revenue, while in the last period of the analysis (2010-2015), they accounted for 64%. However, when the regions are observed, much more significant variations are visible. In the less developed ones (North and Northeast), there is a greater weight of current transfers, over the 80% level, while in the Southeast region, the participation is about 50%. As for the states, the portion related to the tax revenue remained constant from 1985 to 2015, despite slight variations.

When comparing such results with the tax revenues, it is possible to notice an inverse relationship. A system of tax transfers favors the less developed regions and counterbalances the concentration of tax revenues in the most developed ones. Besides, most municipalities do not possess a revenue that can sustain their demands. However, they have experienced significant growth during the period in question, as shown in figure 1.

Figure 1

Evolution of the average of municipal own tax collection, 1995-2015, R$ million



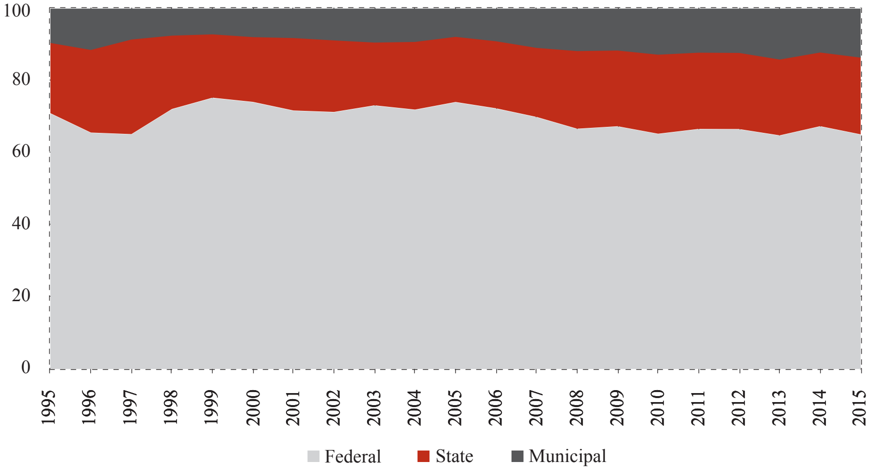
Note: Data from the National Treasury Secretariat.

In the period between 1995 and 2015, the growth of municipal revenue was rapid. However, from the second half of the last decade, the average level of municipal tax collection showed less significant variations, suggesting a stabilization. In figure 2, we show the expenditure per level of the Brazilian government, from 1995 to 2015.

We demonstrate that, in the initial period (1995), the federal government controlled approximately 71% of the aggregate expenditure, while state and municipal governments controlled about 20% and 9%, respectively. In 2015, the share of federal government expenditure was 66%, while the share of state governments was approximately 23% and, finally, the expenses of the municipal governments accounted for about 11% of the total. The sizes of state and local governments, therefore, changed, reflecting the trend of decentralization in the sense that the sub-national expenditure increased its share of the total. However, it is important to be careful in making such claims, as the concept of decentralization is quite complex and includes other dimensions.

Figure 2

Expenditure per level of the Brazilian government, 1995-2015, by percentage



Note: Data from the National Treasury Secretariat.

3 OVERVIEW OF FISCAL DECENTRALIZATION

Empirical studies on the effects of fiscal decentralization on economic growth present mixed results. Part of the literature believes that the relationship between them is direct (Davoodi and Zou, 1998), while others argue that such a relationship is indirect, depending on the quality of the institutions (Libman, 2010; Huynh and Tran, 2021).

This ambiguity may be related to the different methodologies and measures of decentralization used since there is no consensus in the literature on which method best measures the fiscal independence of sub-governments (Martinez-Vazquez and McNab, 2003). To better understand these controversial results, we have separated this review into two parts. The first will present works addressing the connections between fiscal decentralization and economic growth, and the second will show those that used different measures of decentralization.

3.1 FISCAL DECENTRALIZATION AND THE EFFECTS ON ECONOMIC PERFORMANCE

Before discussing the most recent works addressing the research question approached here, let us start with the theoretical and empirical literature that underlies them. A small portion of the literature has considered cross-country evidence on the impact of fiscal decentralization on economic growth. Once again, it is worth noting that these studies have reached mixed conclusions on the subject. From the literature that relates fiscal decentralization to the economic growth of specific countries, we highlight the works by Akai and Sakata (2002), Jin, Qian and Weingast (2005) and Young (2000). Akai and Sakata (2002), through decentralization measures, such as indicators of public revenue, expenses, and autonomy, evaluated this theme using data from 50 American states, from 1992-1996. They provided evidence that decentralization contributes to economic growth, suggesting that recent movements by developed countries towards may stimulate this effect.

Jin, Qian and Weingast (2005) addressed how decentralization in other parts of the country directly impacts the increase in provincial protectionism. They used data from 1982 to 1992, from 30 provinces in China, to show that fiscal decentralization was positively related to the regional growth of GDP per capita, non-farm employment, and non-state industrial production, controlling for provincial tax rates and forcing the growth of regional work. In this way, the authors found that administrative decentralization had a significant positive correlation with local fixed investment, the proportion of local government as compared with central government investment, growth of non-farm employment, and non-state industrial production. To justify the results, the authors revealed that due to amended tax contracts between local and central governments, the first were allowed to withhold a larger fraction of the tax revenues collected from 1982 to 1992, than in the previous decade. Similarly, Young (2000), also studying China, argued that fiscal decentralization has contributed to economic growth in a general form, due to its success in dealing with control and incentive problems. These studies seem to agree that the overall effects of fiscal decentralization were positive in China.

Turning now to cross-country empirical literature, Davoodi and Zou (1998), using panel data for 46 countries, from 1970 to 1989, measured the sub-national fiscal decentralization as a sub-local part of total government expenditure. The authors found a negative relationship between fiscal decentralization and economic growth for developed countries and no relation for developing ones. In the same way, Martinez-Vazquez and McNab (2003) and Iimi (2005) have found positive impacts of decentralization on the growth of a combined set of countries. However, these studies did not focus on developing countries. Using an instrumental variable technique and data on 51 countries, from 1997 to 2001, to describe the effect of decentralization on economic growth, Iimi (2005) discovered that the sub-national share of total government expenditure is significantly and positively correlated with per capita growth.

On the decentralization of revenues and expenditures, Rodríguez-Pose and Ezcurra (2011) observed a negative and meaningful relationship between revenue, expenditure decentralization, and economic growth, in a set of 21 OECD countries, between 1990 and 2005. Baskaran and Feld (2013), analyzing the effect of income and expenditure autonomy on economic growth, also among OECD countries, between 1975 and 2008, reported a negative effect. In a different study, Gemmell, Kneller and Sanz (2013), from a panel dataset of OECD countries, confirmed that expenditure decentralization tends to be associated with lower economic growth, while revenue decentralization is associated with higher growth.

More recently, Bojanic and Collins (2021), in a work on OECD countries, found a significant positive relationship between expenditure and revenue, but not with economic performance. Thus, there is still no consensus on the impact of federalism on economic growth. Göcen, Bayhanay and Göktaş (2017) revealed that, for OECD countries, the impact on growth depends on the econometric strategies used and the measurement criteria.

3.2 DECENTRALIZATION, ECONOMIC GROWTH, AND DIFFERENT VARIABLES

The different decentralization measures used, and other variables that make up the model may explain the controversial results found in the empirical literature. Therefore, it is also important to highlight how the forms of decentralization affect economic efficiency by distorting the efficient allocation of resources.

Filippetti and Sacchi (2016) used the following variables for decentralization: Tax Decentralization (TD), representing tax revenues of local governments due to the total fiscal revenue of the general government; Income Tax Decentralization (TDI), which is the local government income tax due to total general government tax revenue; Property Tax Decentralization (TDP), which is the local government property taxes due to total general government tax revenue; TD_1, which is the proportion of total local government tax revenue regarding the general government tax revenue; and the regional authority index.

Recently, Ligthart and van Oudheusden (2017) observed the relationship between fiscal decentralization and economic growth for 56 countries from various continents, from 1990 to 2007. The authors used instruments based on characteristics such as the original characteristics of the countries, their descentralization system, size, and geographic position. They have concluded that countries with similar characteristics experience a similar process of fiscal decentralization. The results indicated that this relationship remains valid after controlling for endogeneity problems using instrumental variables based on the origin of the common legal system and country size. They do not seem to be able to reject fiscal decentralization for being exogenous, but there is no concrete evidence of causality arising from it for economic growth.

Taking data samples from 2001 to 2011, Ma and Mao (2018) studied the effects of decentralization on Chinese economic growth after the province-managing-county (PMC) reform. Using dummy variables for the moments before and after the reform, they found that it increased the average annual growth rate of GDP by 1.4% over the period studied. The reform abolished the subordinate fiscal relationship between counties and municipalities, transferring much of the fiscal and spending authority from the municipality to the county.

All the articles in this review have the same identification problem: economic growth is impacted by decentralization, but the latter can also be affected by the former. Furthermore, most works use the GMM or Instrumental Variables model to deal with this. Its implementation aims to address endogeneity since decentralization is contemporaneous with other processes, such as infrastructure development. Among the works that more deeply analyze the problem of endogeneity present in the estimates, we highlight those by Aritenang and Chandramidi (2022) and Canavire-Bacarreza, Martinez-Vazquez and Yedgenov (2020).

Unlike this work, that of Aritenang and Chandramidi (2022) applied regional distribution indices, spatial cluster analysis, convergence analysis (GMM and spatial models), and spatial econometrics to Indonesian district-level panel data to address the endogeneity present in the model, since studies indicate that geographic proximity, spatial links, and repercussions have major impacts on economic growth. As in the paper previously mentioned, Canavire-Bacarreza, Martinez-Vazquez and Yedgenov (2020) used the Geographic Fragmentation Index (GFI) and country size as instruments for fiscal decentralization.

4 METHODOLOGY

4.1 MEASURING DECENTRALIZATION

Although fiscal decentralization is a political issue common to many countries, the term is not clear enough, even in the fields of political science and public administration. It is a comprehensive system that includes a framework for decentralizing expenditure, revenue, and corresponding responsibilities to a lower level of government (Dunn and Wetzel, 1998). In this way, it encompasses the decentralization of fiscal expenditures and fiscal revenues. To study its relationship with economic growth, we will follow the metrics adopted by Akai and Sakata (2002), Fu (2010), and Zhang, (2016). First, one should bear in mind that the authority associated with decision-making was allocated according to the legal relationships between the upper and lower levels of government.

Given this, the standard approach to measuring the allocation of authority is to make use of accounting measures such as income and expenses. However, sub-national government expenditures may be financed by transfers from higher levels of government and, therefore, their expenditures do not necessarily reflect the level of authority of local governments.

Furthermore, even if the revenue or expenditure shares are small, the sub-national government can be considered fiscally decentralized, whether it was originally allocated sufficient resources for its own expenses. Therefore, the level of autonomy must be used as a proxy for fiscal decentralization. However, as mentioned earlier, studies have used subnational revenue and expenditure shares as indicators of fiscal decentralization. As it is difficult to develop a unique and completely satisfactory measure, we have considered five different ones based on studies by Akai and Sakata (2002), Fu (2010), and Zhang (2016). They are detailed below:

- Autonomy Indicator 1 (A1) was defined, for each sub-national government, as the fraction of their total revenue generated by tax collection or received by intergovernmental transfer. Other revenues, such as credit operations, are not accounted for.
- Autonomy Indicator 2 (A2) was defined, for each sub-national government, as the fraction of their tax revenue. This indicator is close to the true

fiscal independence of sub-national governments, as it only accounts for their capacity for tax autonomy.

- The Revenue Indicator (RI) reflects the share of the total revenue of each state regarding the consolidated revenue (total collected from all government entities in the country).
- The Production Indicator (PI) reflects the share of the total expenditure of each state regarding consolidated expenditure (the sum of all local expenditures).
- The Production Revenue Indicator (PRI) was defined by the average of the RI and PI indicators.

There are two problems with using accounting information to obtain accurate measures of decentralization, as the authors point out.

First, expenditure by lower levels of the government may be financed by intergovernmental grants from higher levels. Hence, the share of expenditure in the total budget does not necessarily reflect the level of authority allocated to a lower-level government because, to some extent, its grant relates to expenditure authorized by a higher-level government. Therefore, it is inappropriate to regard expenditure shares as necessarily an accurate measure of shares of authority. Given the allocation of lump-sum grants, neither do revenue shares necessarily reflect shares of authority. This is because the authority associated with the spending of the lump-sum grant is attributed to the lower-level government.

Second, even if expenditure shares or revenue shares are small, authority is fiscally decentralized provided that sufficient resources for public spending are originally allocated to the lower-level government; that is if autonomy is achieved. Therefore, autonomy should be considered one of the indicators of fiscal decentralization (Akai and Sakata, 2002).

Because of this, the authors have concluded that to obtain a convincing general result and respond to discussions outside the economic field, it is necessary to build indicators of fiscal decentralization that reflect various points of view. According to them, when the grantor directs the purposes for which the funds are to be used in detail, the grants must be allocated to the level of government that collects the revenues, allowing that the revenue share (A1) in the total budget measures the degree of authority. As for A2, as lump-sum or unconditional donations must be attributed to the level of government that conducts the expenditures, the share of expenditures in the total budget is also an approximation of the degree of the fiscal authority. As A1 and A2 are extreme cases of decentralization, the PRI is an option that combines both. The RI and PI reflect the fiscal autonomy of the local government, considering how public spending at a lower level of government is financed on its revenue or expenditure.

4.2 EMPIRICAL STRATEGY

4.2.1 GENERALIZED METHOD OF MOMENTS (GMM)

In this paper, we explore how variations in decentralization across Brazilian states affect their growth. Variations in fiscal policies, as measured by the proposed indices of decentralization, are likely to be correlated with local capacity, institutions, and other confounders. The rules governing the degree of fiscal decentralization in Brazil according to the Constitution from 1988 allow states and municipalities the freedom to collect taxes locally, but also to boost their revenue from transfers from the federal government. For this reason, it is possible to visualize that several factors can affect growth and are related to decentralization, making it impossible to reject the hypothesis of the presence of endogeneity in the models.

To address this endogeneity, Arellano and Bond (1991) proposed an estimate using instrumental variables from the difference between the current period and the lag of the endogenous variable. That is, this estimator applies the first difference to remove panel-level effects and uses instruments to provide momentum conditions. In this way, it is possible to accommodate large autoregressive parameters and variance ratios in the panel-level effect for the idiosyncratic error variance.

However, as shown by Blundell and Bond (2000), the Arellano-Bond estimator presents weaknesses concerning lagged-level instruments. Due to persistent autoregressive processes or variance ratio of panel effects and idiosyncratic error, it becomes very large. Thus, the model we used fits the panel of estimators of dynamic data of the estimator used by Arellano and Bover (1995) and Blundell and Bond (2000), which was designed for data with many panels and few periods, assuming that there is no autocorrelation in idiosyncratic errors and that it does not require, in the initial condition, that the independent variables do not correlate with the first difference of the first observation of the dependent variable. In this way, the Arellano-Bover and Blundell-Bond estimator presents the estimated model as follows

$$y_{i,t} = \sum_{j=1}^{p} \alpha_j y_{it-j} + X_{it}\beta_1 + W_{it}\beta_2 + v_i + \varepsilon_{it}\, i.i.d. \sim N\left(\mu, \sigma^2\right) \qquad (1)$$

where α_j are the p parameters to be estimated, x_{it} is a 1 x k_1 vector of strictly exogenous covariates, β_1 is a k_1 x 1 vector of parameters to be estimated, w_{it} is a 1 x k_2 vector of predetermined or endogenous covariates, β_2 is a k_2 x 1 vector of parameters to be estimated, v_i are the effects on the level of the panel (which may be correlated with the covariates) and it is i.i.d. in the whole sample, with mean μ and variance σ^2. Adapted from Akai and Sakata (2002), we can express the regression model as

$$\Delta GDP_{i,t} = LnGDP_{it} - LnGDP_{i(t-1)}, i = 1, \ldots, 27; t = 1995, \ldots, 2015 \qquad (2)$$

$$\Delta GDP_{i,t} = \alpha_j + \Delta GDP_{i(t-j)} + LnX_{it}\beta_1 + LnW_{it}\beta_2 + v_i + \varepsilon_{it} \rightarrow i.i.d. \sim N\left(\mu, \sigma^2\right) \qquad (3)$$

In the equation above, *i* refers to the state changing in each year *t*; $LnGDP_{it}$ represents the natural logarithm of GDP, so that our variable of interest is represented in terms of the GDP growth rate of each state. Our model will explain economic growth from the endogenous indicators of fiscal decentralization, and degree of trade openness; Gini index and a X_{it} vector contain the exogenous controls. Finally, vi represents the panel-level effects, while uit is the error term.

Here, we have emphasized that the use of lags may not be the most appropriate for the problem of endogeneity since the series presents a strong temporal persistence. However, given the limitations in the definition of more appropriate instruments, we understand that the GMM-System meets the research needs.

4.3 DATA SOURCES

Our data compiles a set of social, economic, and public fiscal data, composing a panel of the Brazilian states from 1995 to 2015. Firstly, our four indicators of fiscal decentralization were obtained through the Secretariat of the National Treasury. By law, all spheres of government must disclose their income and expense accounting information.

The degree of trade openness, Gini index, population, homicide rate, and employed population were inserted in the model as control variables. To control adverse effects within the public budget, we have also included an electoral dummy. The data used for collecting the Gini Index, population and the employed population were obtained from the Brazilian Institute of Geography and Statistics (IBGE), while the degree of trade openness was taken from the data of the statistics of foreign trade (AliceWeb – MDIC). Variables that could improve our estimates, such as literacy rate, human capital, and investment could not be included because of the limited public data available.

Table 2

Definition of variables and reason for inclusion

Explanatory variable	Variable	Reason for inclusion	Source
Main variables			
Revenue indicator	RI	Ratio between the state revenue *i* and the consolidated revenue	STN
Production indicator	PI	Ratio between the expenditure of state *i* and the consolidated expenditure	STN
Production and revenue indicator	PRI	Weighted average between Revenue and production indicator	STN
Autonomy indicator 1	A1	The ratio between the own revenue of the states and their total revenue, excluding transfers	STN
Autonomy indicator 2	A2	The ratio between the own revenue of the states and the total of their revenue	STN

Explanatory variable	Variable	Reason for inclusion	Source
Control variables			
Degree of commercial opening	OPNESS	Ratio between the trade balance result and GDP	Comex Stat
Gini Index	Gini	Gini Index for income concentration	IPEADATA/ IBGE
Population	POP	Population value	IBGE
Employed population	POP OCUP	Number of people who are employed	IBGE
Dummy election	ELECTION	Dumymy variable indicating state election years	Superior electoral court (TSE)
Homicide rate	HOM	Homicide rate per one hundred thousand inhabitants	IBGE
School effectiveness	PSE	Efficiency indicator created from school attendance, years of schooling, and illiteracy rate variables	IPEADATA/ IBGE

The characteristics of the variables are summarized below, in table 2. The data[2] concern annual frequency from 1996 to 2015, where the rate of growth of GDP is the dependent variable of the model.

To obtain a measure of the performance and efficiency of the public sector, it is necessary to add several indicators that compose their obligations. Drawing on Afonso, Schuknecht and Tanzi (2005), it is possible to use the public performance indicator (PSP), the public expenditure indicator, and the public sector efficiency indicator (PSE). We have measured them by their weighted average, where the indicators in the year *i* and state *j* are divided by the national average. The average concerning their respective sectors divides the expenditures, the year *i* and state *j*. Finally, the efficiency meter can be described as follows

$$PSE_{nj} = PSP_{nj} \Big/ \frac{\sum_{1}^{n} PSP_j}{n} \tag{4}$$

Thus, the ratio between each state and the sum of *n* government areas (the areas formed by an arithmetic mean of the syndicators) comprise the efficiency indicator. Values greater than one represent efficiency, while those inferior to one represent inefficiency. We have composed the efficiency indicator using the variables of average years of schooling, school attendance in primary and secondary education, and illiteracy rate for each state.

[2] When necessary, the collected data was deflated, as the literature recommends.

4.4 DESCRIPTIVE STATISTICS

In table 3, we show the results of the descriptive statistics of the data used in the paper for the years between 1996 and 2015[3]. The results show that only the growth rate of the services sector, homicide rate, population, and employed population present dispersion between the minimum and maximum in relation to the mean. For the control variables, we have applied the natural logarithm to stabilize the series.

Investigating the economic growth of the states, we could see that the service sector drives this growth the most, which is the sector with the largest share of GDP in all Brazilian states. Over the years, the agricultural and industrial sectors have reduced their participation regarding GDP. This was mainly due to national and state policies that allowed the scenario to become more favorable for the growth of the service sector. Another factor that we could observe is that, on average, this value is large when we analyze Brazil in a general overview. This happens because states considered economically very small grew until the last year of our sample.

Table 3

Averages, standard deviations, and definitions of the variables used

Variables	Mean	Std. dev.	Min	Max
Dependent variables				
Δ GDP	0.124593	0.064578	-0.07	0.33
Δ Agriculture	0.123953	0.270231	-0.62	2.29
Δ Industry	0.131147	0.243828	-0.43	2.40
Δ Services	0.151147	0.205321	-0.50	4.03
Mainly variables				
A1	0.809024	0.092698	0.14	1.00
A2	0.498713	0.173038	0.09	0.87
PI	0.007712	0.012192	0.0006	0.13
RI	0.007681	0.012195	0.000533	0.14
PRI	0.007697	0.012186	0.000546	0.13
Control variables				
Education	1.00	0.094177	0.77	1.21
Openness	0.145475	0.127781	0.01	0.59
POP	6,762,641	8,029,421	254,499	44,000,000
Gini	0.55241	0.049231	0.42	0.69
Homicide rate	28.13537	13.15587	4.50	71.40
Occupied population	2,952,986	3,778,686	70,996	22,000,000

Note: All data collected are at the state level of Brazil and aggregate of the period of our analysis.

For Autonomy Indicators 1 and 2, the scenarios are different. As presented in table 3, on average, states have revenue gains of 30% when we include federal transfers, with considerable weight in the state budget. When dealing with the RI and PI indicators, it is possible to observe that, on average, the variables are remarkably similar.

[3] The defined period is due to the large distortions of the economic growth rate before the system of the Real Plan, and the last results of variables only extended to 2015.

However, this analysis needs to be performed carefully, since both indicators may have different meanings; for example, as an attempt to boost the local market, a federal state may increase its level of spending in relation to other states and federative entities, even if its revenue does not grow by the same amount. If this same state has not reached its goal of economic growth, it is possible that this expenditure has caused only a crowding out effect and, therefore, it is being inefficient.

To evaluate the real importance of fiscal decentralization on economic growth in Brazilian states, we have estimated the model in equation 5 and presented our main results in the following section.

5 RESULTS AND DISCUSSION

We will now discuss the main results obtained through the estimates based on the previous discussion.

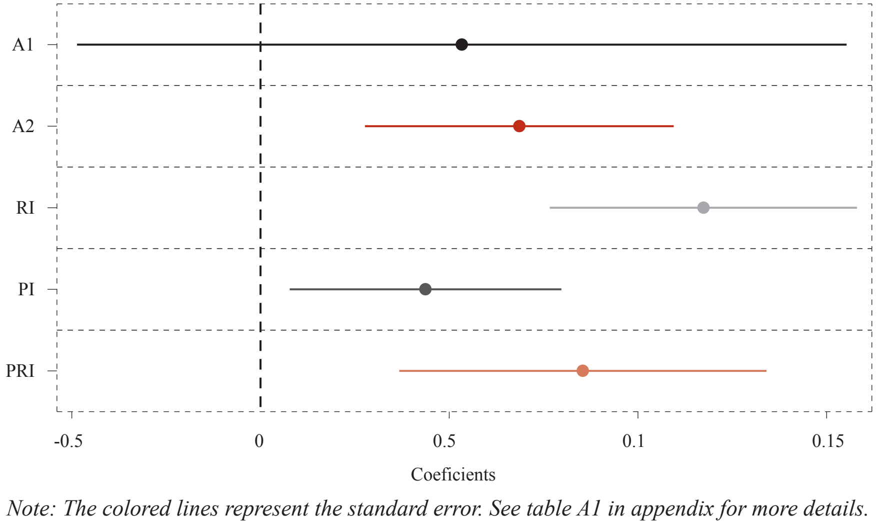
5.1 MAIN RESULTS

Recently, the effect of taxation enforcement on economic growth has been the subject of empirical studies, becoming the focus of debates on government reforms. To contribute to the discussion, this research uses four types of indicators besides control variables that measure fiscal decentralization to understand its relationship with economic growth in Brazilian states.

In figure 3, we present the relationship between the variables studied in this article. The results are based on data from the Stata software package for the estimation of the data model in a dynamic panel (GMM).

Figure 3

Results for coefficients of decentralization



Note: The colored lines represent the standard error. See table A1 in appendix for more details.

Within the estimates, the variables for population, employed population, and life expectancy were considered exogenous for the model, while the variable of fiscal decentralization, Gini index, and degree of trade openness were endogenous. The dynamic GMM specification uses its lags as instruments for correcting endogeneity, therefore we have included the variables that were most indicated to be endogenous and checked if they are valid together, using the Sargan and Arellano-Bond tests.

According to the results, our main finding is that the indicators of fiscal decentralization A2, PI, RI, and PRI are positive and statistically significant to the economic growth of the states in the analyzed period, especially RI. As in the studies by Akai and Sakata (2002) and Gemmell, Kneller and Sanz (2013), we have found evidence of a positive effect of these measures of fiscal decentralization on economic growth. However, unlike Qiao, Martinez-Vazquez and Xu (2008), we have found no evidence of any effect of fiscal decentralization on economic growth when measured in terms of autonomy, which is the A1 variable. We can thus see that the variable used to measure decentralization influenced our result.

All coefficients of the estimations are interpreted as elasticity. Therefore, we can interpret that a 1% increase in the autonomy of states is capable of increasing GDP growth by 8%. Regarding the control variables, we have found a positive and significant effect of trade openness on economic growth. This result diverges from the literature since it shows that the increase in the coefficient contributes to the increase in economic growth. Comparable results were found by Rodríguez-Pose and Ezcurra (2011), Filippetti and Sacchi (2016), and Ligthart and van Oudheusden (2017).

In addition, we must highlight the positive and significant effect of the Gini index on economic growth. This result was not in line with our expectations, since it shows that the increase in inequality contributes positively to economic growth. According to Mirrlees (1971), the possibility of earning a higher income makes the individual strive harder. In this way, it contributes to higher levels of productivity. In this sense, the result corroborates those by Forbes (2000), in which an increase in the level of income inequality has a positive relation to the economic growth of a country.

Subsequently, the Sargan test was conducted to identify overidentification constraints. Using instrumental models with a lag in the dependent variable, we have as a result that the Sargan and Arellano-Bond tests indicated that there was no residual correlation of the second order and that the instruments are valid for all the estimated models (see appendix for more details). The results have demonstrated that the model evaluated does not reject the hypothesis that the restrictions are valid, leading to the conclusion that the instruments used are valid, that is, not correlated with the error term and are, therefore, correctly excluded from the estimated equation, allowing the existence of the model. The Arellano-Bond test seeks to show the autocorrelation for p differences in the error term. The results show that, for the first difference in the error term, the probability of not rejecting the null hypothesis of no autocorrelation is approximately zero.

The results in table A1 are presented in a more simplified form, in which it is possible to see that, among the indicators used in the estimation, RI is the variable that best fits the explanations of economic growth (that is, it has obtained the highest statistical significance and degree of reliability). We have thus chosen this indicator to represent decentralization in the next steps.

5.2 ANALYSIS BY SECTOR

Our next step was to evaluate which sectors of the economy are responsible for the observed positive effect of decentralization on economic growth. Tables A2, A3 and A4, in the appendix section, present these results.

Historically, the Brazilian economy has undergone a major structural change. Since the 1950s, the service sector has become the one with the greatest share of gross value added. The data presented shows that this growth led to reduced participation of agriculture and stable participation of industry in GDP.

According to de Andrade Jacinto and Ribeiro (2015), the productivity of services (except for commerce) is high and showed growth between the mid-1990s and the end of the 2000s. In this sense, the expansion of the participation of services in employment had the effect of increasing the aggregate productivity of the economy.

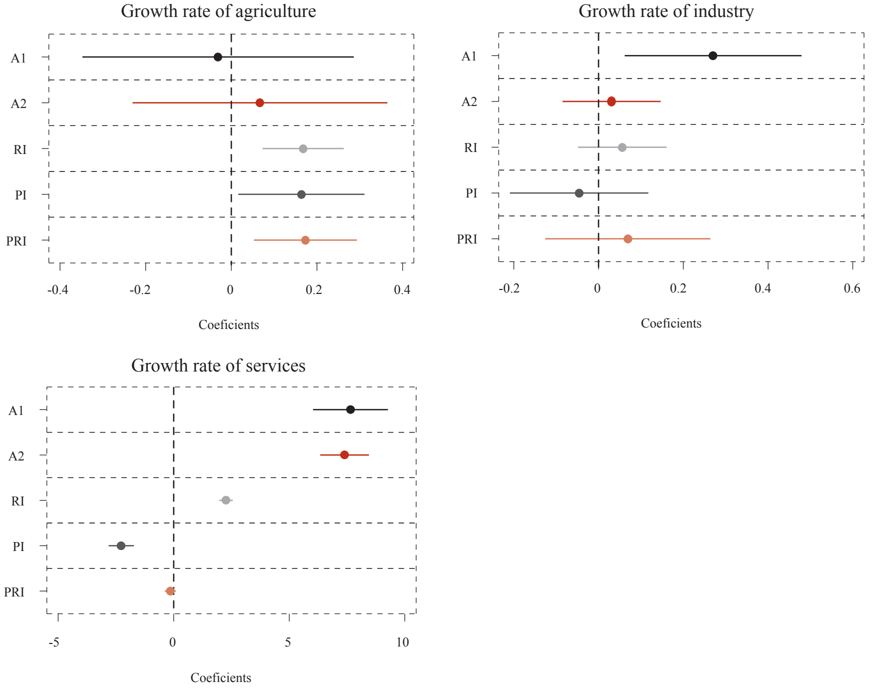
The results indicate the robust performance of the three sectors, indicating that it may be in the service sector that fiscal decentralization generates the greatest positive effects on the Brazilian economy. In fact, Christl, Köppl-Turyna and Kucsera indicated that fiscal decentralization increases efficiency and that the interaction between research, technology, and productivity has been relevant to explain economic growth.

Thus, for the Brazilian states during the period of analysis, decentralization boosts the economic growth of all sectors. This result corroborates those of Ma and Mao (2018), who evidenced the contribution of fiscal decentralization to industrial economic activity. This is known to have positive effects on economic growth. This result is valid for the Arellano-Bond tests for all sectors, except for some estimates that had the AR (2) coefficient significant at 10%.

The results show that greater autonomy in state revenue is positively associated with higher agricultural growth rates. At the same time, states with greater autonomy linked to transfers from the general government are associated with higher rates of growth in the industry. And, finally, the autonomy and revenue variables indicate positive effects on the growth rate in the service sector, but the expenditure variable is negatively associated with this growth. These results are briefly presented in figure 4.

Figure 4

The impact of decentralization on the economic growth of various sectors



Note: The colored lines represent the standard error. See tables A2, A3 and A4 for more details.

6 FINAL CONSIDERATIONS

Fiscal decentralization is a relevant issue in the economic literature. In Brazil, this topic became more relevant after implementation of the Constitution in 1988, in which states and municipalities gained more freedom to provide public goods and services. Thus, this work aimed to identify the relationship between fiscal decentralization and economic growth in Brazilian states.

Through five measures of decentralization proposed by Akai and Sakata (2002), the estimation performed found positive and significant effects for the variable of decentralization A2, which measures decentralization as the ratio between state revenue, derived from transfers and own revenue, and total revenue. The result also showed a significant result for the revenue, expense and PRI indicator. The positive result agrees with the expected theoretical support.

This result shows that fiscal decentralization is an important instrument to achieve higher growth rates. In addition, the positive relationship between the rate of growth, human capital, and trade openness shows which policies can achieve better results in the long term.

These results are important, as they contribute to the debate on public policies concerning higher rates of economic growth. Policymakers should improve the mechanisms for decentralization to identify means of strengthening the tax structure and solving the problems of expenditure and revenue redistribution of the government.

Thus, future efforts that aim to contribute to a greater decentralization of the federative entities of the country can also contribute to its economic growth. It is important to note, however, that such evidence should be treated with caution since the causes behind this positive effect of decentralization on growth are not known.

Therefore, this study can lead to further endeavours to identify the causes of the positive effects of fiscal decentralization on the GDP of Brazilian states. Confirming such causes with greater accuracy would enable more efficient public policies. Finally, some issues deserve to be further investigated to improve the understanding of the relationship between the growth rate and fiscal decentralization in Brazilian states. The first is the incorporation of newer and more accurate indicators in relation to the growth rate. The second one involves simulations of the impacts of the growth rate through the expansion of transfers or the tax base itself. Some variables, such as literacy rate, human capital, and investment should be incorporated, but because of data limitations, this was not possible. More findings are needed to explore their effect on economic growth. Lastly, the case of fiscal decentralization should be analyzed at the municipal level in Brazil.

Disclosure statement

The authors declare that there is no conflict of interest.

REFERENCES

1. Afonso, A., Schuknecht, L. and Tanzi, V., 2005. Public sector efficiency: An international comparison. *Public Choice,* 123(3-4), pp. 321-347. https://doi.org/10.1007/s11127-005-7165-2
2. Akai, N. and Sakata, M., 2002. Fiscal decentralization contributes to economic growth: Evidence from state-level cross-section data for the United States. *Journal of Urban Economics*, 52(1), pp. 93-108. https://doi.org/10.1016/S0094-1190(02)00018-9
3. Araujo, J. M. and Siqueira, R. B., 2016. Demanda por gastos públicos locais: Evidências dos efeitos de ilusão fiscal no Brasil. *Estudos Economicos*, 46(1), pp. 189-219. https://doi.org/10.1590/0101-416146116jar
4. Arellano, M. and Bond, S., 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(2), pp. 277-297. https://doi.org/10.2307/2297968_
5. Arellano, M. and Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), pp. 29-51. https://doi.org/10.1016/0304-4076(94)01642-D
6. Aritenang, A. F. and Chandramidi, A. N., 2022. The spatial effects of fiscal decentralization on regional convergence: the case of regions in indonesia. *GeoJournal*. https://doi.org/10.1007/s10708-022-10724-2
7. Baskaran, T. and Feld, L. P., 2013. Fiscal Decentralization and Economic Growth in OECD Countries: Is There a Relationship? *Public Finance Review*, 41(4), pp. 421-445. https://doi.org/10.1177/1091142112463726
8. Blundell, R. and Bond, S., 2000. GMM Estimation with persistent panel data: An application to production functions. *Econometric Reviews*, 19(3), pp. 321-340. https://doi.org/10.1080/07474930008800475
9. Bojanic, A. N. and Collins, L. P. A., 2021. Differential effects of decentralization on income inequality: evidence from developed and developing countries. *Empirical Economics*, 60(4), str. 1969-2004. https://doi.org/10.1007/s00181-019-01813-2
10. Canavire-Bacarreza, G., Martinez-Vazquez, J. and Yedgenov, B., 2020. Identifying and disentangling the impact of fiscal decentralization on economic growth. *World Development*, 127. https://doi.org/10.1016/j.worlddev.2019.104742
11. Christl, M., Köppl-Turyna, M. and Kucsera, D., 2020. Determinants of Public-Sector Efficiency: Decentralization and Fiscal Rules. *Kyklos*, 73(2), pp. 253-290. https://doi.org/10.1111/kykl.12224
12. Colombo, D. G. and Martinez-Vazquez, J., 2020. Fiscal decentralization and public investment in Innovation: A country panel analysis. *Publius: The Journal of Federalism*, 50(4), pp. 672-697. https://doi.org/10.1093/publius/pjz033
13. Davoodi, H. and Zou, H. F., 1998. Fiscal Decentralization and Economic Growth: A Cross-Country Study. *Journal of Urban Economics*, 43(2), pp. 244-257. https://doi.org/10.1006/juec.1997.2042

14. de Andrade Jacinto, P. and Ribeiro, E. P., 2015. Crescimento da produtividade no setor de serviços e da indústria no Brasil: Dinâmica e heterogeneidade. *Economia Aplicada*, 19(3), pp. 401-427. https://doi.org/10.1590/1413-8050/ea119450
15. Dunn, J. and Wetzel, D., 1998. Fiscal Decentralization in Former Socialist Economies: Progress and Prospects. *World Development*, *1994*.
16. Filippetti, A. and Sacchi, A., 2016. Decentralization and economic growth reconsidered: The role of regional authority. *Environment and Planning C: Government and Policy*, 34(8), pp. 1793-1824. https://doi.org/10.1177/0263774X16642230
17. Forbes, K. J., 2000. A reassessment of the relationship between inequality and growth. *American Economic Review*, 90(4), pp. 869-887. https://doi.org/10.1257/aer.90.4.869
18. Fu, Y., 2010. Fiscal Decentralization, Governance and Non-Economic Public Goods Provision. *Economic Research*, 3(September).
19. Gadenne, L. and Singhal, M., 2014. Decentralization in developing economies. *Annual Review of Economics*, (6), pp. 581-604. https://doi.org/10.1146/annurev-economics-080213-040833
20. Gemmell, N., Kneller, R. and Sanz, I., 2013. Fiscal decentralization and economic growth: Spending versus revenue decentralization. *Economic Inquiry*, 51(4), pp. 1915-1931. https://doi.org/10.1111/j.1465-7295.2012.00508.x
21. Göcen, S., Bayhanay, A. and Göktaş, N., 2017. Fiscal Decentralization and Economic Growth: Theory and Application. *MPRA Paper*, No. 84523.
22. Huynh, C. M. and Tran, H. N., 2021. Moderating effects of corruption and informality on the fiscal decentralization – economic growth nexus: Insights from OECD countries. *Annals of Public and Cooperative Economics*, 92(2), pp. 355-373. https://doi.org/10.1111/apce.12298
23. Iimi, A., 2005. Decentralization and economic growth revisited: An empirical note. *Journal of Urban Economics*, 57(3), pp. 449-461. https://doi.org/10.1016/j.jue.2004.12.007
24. Jametti, M. and Joanis, M., 2016. Electoral Competition as a Determinant of Fiscal Decentralisation. *Fiscal Studies*, 37(2), pp. 285-300. https://doi.org/10.1111/j.1475-5890.2015.12061
25. Jin, H., Qian, Y. and Weingast, B. R., 2005. Regional decentralization and fiscal incentives: Federalism, Chinese style. *Journal of Public Economics*, 89(9-10), pp. 1719-1742. https://doi.org/10.1016/j.jpubeco.2004.11.008
26. Libman, A., 2010. Constitutions, regulations, and taxes: Contradictions of different aspects of decentralization. *Journal of Comparative Economics*, 38(4), pp. 395-418. https://doi.org/10.1016/j.jce.2010.05.004
27. Ligthart, J. E. and van Oudheusden, P., 2017. The Fiscal Decentralisation and Economic Growth Nexus Revisited. *Fiscal Studies*, 38(1), pp. 141-171. https://doi.org/10.1111/1475-5890.12099

28. Ma, G. and Mao, J., 2018. Fiscal Decentralisation and Local Economic Growth: Evidence from a Fiscal Reform in China. *Fiscal Studies*, 39(1), pp. 159-187. https://doi.org/10.1111/j.1475-5890.2017.12148
29. Martinez-Vazquez, J. and McNab, R. M., 2003. Fiscal Decentralization and Economic Growth. *World Development*, 31(9), pp. 1597-1616. https://doi.org/10.1016/S0305-750X(03)00109-8
30. Mirrlees, J. A., 1971. An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies*, 38(2), pp. 175-208. https://doi.org/10.2307/2296779
31. Qiao, B., Martinez-Vazquez, J. and Xu, Y., 2008. The tradeoff between growth and equity in decentralization policy: China's experience. *Journal of Development Economics*, 86(1), pp. 112-128. https://doi.org/10.1016/j.jdeveco.2007.05.002
32. Qiao, M., Ding, S. and Liu, Y., 2019. Fiscal decentralization and government size: The role of democracy. *European Journal of Political Economy*, 59, pp. 316-330. https://doi.org/10.1016/j.ejpoleco.2019.04.002
33. Rodríguez-Pose, A. and Ezcurra, R., 2011. Is fiscal decentralization harmful for economic growth? Evidence from the OECD countries. *Journal of Economic Geography*, 11(4), pp. 619-643. https://doi.org/10.1093/jeg/lbq025
34. Shah, A., 2006. Fiscal decentralization and macroeconomic management. *International Tax and Public Finance*, 13(4), pp. 437-462. https://doi.org/10.1007/s10797-006-8948-1
35. Thanh, S. D. and Canh, N. P., 2020. Fiscal Decentralization and Economic growth of Vietnamese provinces: the role of local public governance. *Annals of Public and Cooperative Economics*, 91(1), pp. 119-149. https://doi.org/10.1111/apce.12255
36. Young, A., 2000. The razor's edge: Distortions and incremental reform in the People's Republic of China. *Quarterly Journal of Economics*, 115(4), pp. 1091-1135. https://doi.org/10.1162/003355300555024
37. Zhang, S., 2016. Fiscal Decentralization, Budgetary Transparency, and Local Government Size in China. *Emerging Markets Finance and Trade*, 52(7), pp. 1679-1697. https://doi.org/10.1080/1540496X.2016.1142213

Table A1

Main results of the estimation, 1996-2015

Estimator: GMM	Equations				
Variables	**(1.1)**	**(1.2)**	**(1.3)**	**(1.4)**	**(1.5)**
Δ GDP L1.	0.01	-0.00	0.04	0.02	0.03
	(0.38)	(-0.16)	(1.24)	(0.74)	(1.09)
A1	0.05	–	–	–	–
	(1.03)	–	–	–	–
A2	–	0.07***	–	–	–
	–	(3.29)	–	–	–
RI	–	–	0.12***	–	–
	–	–	(5.65)	–	–
PI	–	–	–	0.04**	–
	–	–	–	(2.38)	–
PRI	–	–	–	–	0.09***
	–	–	–	–	(3.44)
Education	-0.14	-0.05	0.29**	0.08	0.15
	(-1.13)	(-0.61)	(2.56)	(0.96)	(1.03)
Openness	0.03**	0.02	0.05***	0.03**	0.04**
	(2.37)	(1.53)	(2.73)	(2.04)	(2.23)
Gini	0.19***	0.14***	0.34***	0.24***	0.27***
	(4.95)	(4.41)	(8.79)	(7.43)	(7.81)
Pop	-0.03	-0.05*	-0.14***	-0.05	-0.10**
	(-1.50)	(-1.65)	(-3.90)	(-1.60)	(-2.39)
Homicide rate	0.00*	0.00	-0.00	0.00	0.00
	(1.72)	(0.85)	(-0.17)	(0.88)	(0.33)
Occupied population	0.00	0.01*	0.00	0.00	0.00
	(0.98)	(1.67)	(0.78)	(0.47)	(0.72)
Dummy election	-0.00	-0.00*	-0.01***	-0.01***	-0.00**
	(-1.37)	(-1.90)	(-4.40)	(-3.36)	(-2.20)
Constant	0.72**	0.93**	3.14***	1.29**	2.35***
	(2.08)	(2.19)	(4.39)	(2.35)	(2.83)
Observations	513	513	513	513	513
Wald Test	178.48	260.44	397.17	354.57	270.53
No. of instruments	286	439	286	286	286
Sargan Test Chi2	25.28	24.97	24.62	25.81	25.37
Prob > Chi2	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Arellano-Bond Test					
Order 1	-4.23***	-4.14***	-4.40***	-4.18***	-4.34***
Prob > z	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Order 2	-1.80*	-1.88*	-1.35	-1.79*	-1.57
Prob > z	(0.08)	(0.07)	(0.12)	(0.05)	(0.11)

*Note: * Significant at 10% ** Significant at 5% and *** Significant at 1%.*

PUBLIC SECTOR ECONOMICS
47 (2) 255-280 (2023)

Table A2

Main results of the estimation for Agriculture Value Added, 1996-2015

Estimator: GMM	Equations				
Variables	**(1.1)**	**(1.2)**	**(1.3)**	**(1.4)**	**(1.5)**
Δ GDP L1.	-0.04	-0.06	-0.05	-0.04	-0.05
	(-0.79)	(-1.23)	(-1.64)	(-1.40)	(-1.53)
A1	-0.03	–	–	–	–
	(-0.19)	–	–	–	–
A2	–	0.07	–	–	–
	–	(0.44)	–	–	–
RI	–	–	0.17***	–	–
	–	–	(3.47)	–	–
PI	–	–	–	0.16**	–
	–	–	–	(2.18)	–
PRI	–	–	–	–	0.17***
	–	–	–	–	(2.82)
Education	0.30	0.44	1.22***	1.12***	1.22***
	(0.48)	(0.75)	(2.99)	(2.70)	(2.84)
Openness	0.04	0.04	0.06	0.05	0.05
	(0.64)	(0.59)	(1.45)	(1.06)	(1.21)
Gini	0.17	0.24*	0.40***	0.29***	0.37***
	(1.46)	(1.67)	(3.27)	(2.69)	(3.17)
Pop	-0.09	-0.07	-0.16	-0.13	-0.15
	(-0.48)	(-0.35)	(-1.24)	(-0.97)	(-1.13)
Homicide rate	0.01***	0.00	0.00	0.00	0.00
	(2.90)	(1.34)	(1.21)	(1.03)	(1.19)
Occupied population	-0.00	-0.01	-0.02	-0.03	-0.03
	(-0.00)	(-0.16)	(-0.23)	(-0.38)	(-0.31)
Dummy election	-0.02	-0.03***	-0.02**	-0.02	-0.02*
	(-1.60)	(-3.42)	(-1.97)	(-1.20)	(-1.70)
Constant	1.61	1.56	4.18***	3.69**	4.02**
	(0.65)	(0.56)	(2.68)	(2.13)	(2.38)
Observations	513	513	513	513	513
Wald Test	32.27	58.71	57.19	63.72	55.09
No. of instruments	286	286	286	286	286
Sargan Test Chi2	22.77	18.23	20.79	20.85	24.50
Prob > Chi2	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Arellano-Bond Test					
Order 1	-3.68***	-3.51	-3.67***	-3.77***	-3.73***
Prob > z	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Order 2	-1.41	-1.97**	-1.65*	-1.36	-1.52
Prob > z	(0.15)	(0.04)	(0.09)	(0.17)	(0.12)

*Note: * Significant at 10% ** Significant at 5% and *** Significant at 1%.*

Table A3

Main results of the estimation Industry Value Added, 1996-2015

Estimator: GMM			**Equations**		
Variables	**(1.1)**	**(1.2)**	**(1.3)**	**(1.4)**	**(1.5)**
Δ GDP L1.	-0.18	-0.08	0.56	-0.06	0.51
	(-0.68)	(-0.20)	(0.96)	(-0.15)	(1.03)
A1	0.27**	–	–	–	–
	(2.54)	–	–	–	–
A2 -	–	0.03	–	–	–
	–	(0.52)	–	–	–
RI -	–	–	0.06	–	–
	–	–	(1.05)	–	–
PI -	–	–	–	-0.05	–
	–	–	–	(-0.55)	–
PRI -	–	–	–	–	0.07
	–	–	–	–	(0.69)
Education	-0.18	-0.08	0.56	-0.06	0.51
	(-0.68)	(-0.20)	(0.96)	(-0.15)	(1.03)
Openness	0.04**	0.00	0.08**	0.04*	0.09**
	(1.98)	(0.20)	(2.40)	(1.91)	(2.36)
Gini	0.72***	0.66***	0.79***	0.81***	0.86***
	(10.21)	(6.80)	(5.04)	(7.50)	(4.87)
Pop	-0.01	-0.00	-0.15*	-0.01	-0.17
	(-0.16)	(-0.01)	(-1.68)	(-0.12)	(-1.63)
Homicide rate	0.00	0.00	0.00	0.00*	0.00
	(1.03)	(0.47)	(0.45)	(1.89)	(0.81)
Occupied population	0.03***	0.04***	0.03***	0.02***	0.05
	(6.34)	(6.56)	(4.32)	(3.25)	(0.89)
Dummy election	0.00	-0.00	-0.00	-0.01*	-0.00
	(0.15)	(-0.06)	(-0.09)	(-1.95)	(-0.78)
Constant	0.38	0.03	2.90*	0.20	3.14
	(0.54)	(0.04)	(1.67)	(0.12)	(1.45)
Observations	513	513	513	513	513
Wald Test	651.69	795.83	482.41	404.99	306.72
No. of instruments	286	286	286	286	286
Sargan Test Chi2	21.58	24.61	21.11	22.20	20.44
Prob > Chi2	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Arellano-Bond Test					
Order 1	-3.04***	-2.86***	-3.00***	-2.94***	-2.96***
Prob > z	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Order 2	-1.87*	-1.84*	-2.04**	-1.93*	-2.32**
Prob > z	(0.06)	(0.05)	(0.04)	(0.05)	(0.02)

*Note: * Significant at 10% ** Significant at 5% and *** Significant at 1%.*

PUBLIC SECTOR ECONOMICS
47 (2) 255-280 (2023)

Table A4

Main results of the estimation Services Value Added, 1996-2015

Estimator: GMM			**Equations**		
Variables	**(1.1)**	**(1.2)**	**(1.3)**	**(1.4)**	**(1.5)**
Δ GDP L1.	-0.07***	-0.07***	-0.07***	-0.08***	-0.07***
	(-31.90)	(-27.23)	(-37.12)	(-17.04)	(-44.84)
A1	7.65***	–	–	–	–
	(9.25)	–	–	–	–
A2 -	–	7.40***	–	–	–
	–	(13.73)	–	–	–
RI	–	–	2.27***	–	–
	–	–	(15.23)	–	–
PI	–	–	–	-2.27***	–
	–	–	–	(-8.15)	–
PRI	–	–	–	–	-0.15
	–	–	–	–	(-1.19)
Education	-7.14**	-0.82	2.30	-9.64***	-3.41**
	(-2.05)	(-0.50)	(1.02)	(-4.98)	(-2.10)
Openness	-1.00***	-0.77***	-0.71***	-1.34***	-0.99***
	(-4.59)	(-3.80)	(-4.44)	(-9.90)	(-8.02)
Gini	-6.92***	-8.43***	-1.53***	-5.01***	-3.43***
	(-10.96)	(-9.55)	(-4.36)	(-11.40)	(-7.04)
Pop	0.60	-3.28***	-1.55***	2.56***	0.78**
	(0.46)	(-6.45)	(-3.41)	(3.92)	(2.25)
Homicide rate	0.00	0.00	0.00	0.03*	0.01
	(0.13)	(0.32)	(0.43)	(1.96)	(1.40)
Occupied population	0.65***	0.95***	0.43***	0.51***	0.40***
	(3.39)	(6.22)	(4.21)	(5.22)	(4.53)
Dummy election	-0.55***	-0.47***	-0.69***	-0.66***	-0.63***
	(-10.28)	(-7.00)	(-13.49)	(-10.98)	(-18.43)
Constant	-22.27	36.00***	28.08***	-64.55***	-22.18***
	(-1.26)	(4.29)	(3.85)	(-6.02)	(-3.74)
Observations	513	513	513	513	513
Wald Test	17,021.67	14,106.26	19,232.38	26,595.50	87,624.82
No. of instruments	286	286	286	286	286
Sargan Test Chi2	26.30	26.53	26.6256	26.82	26.77
Prob > Chi2	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Arellano-Bond Test					
Order 1	-1.17	-1.18	-1.14	-1.16	-1.15
Prob > z	(0.24)	(0.23)	(0.25)	(0.24)	(0.24)
Order 2	-1.45	-1.45	-1.54	-1.56	-1.55
Prob > z	(0.14)	(0.14)	(0.12)	(0.11)	(0.12)

*Note: * Significant at 10% ** Significant at 5% and *** Significant at 1%.*

Participatory budgeting (contexts, models and practical experience)

DANIEL KLIMOVSKÝ ET AL.
Participatívne rozpočtovanie (kontexty, modely a praktické skúsenosti)
Wolters Kluwer, Bratislava, 2021, pp. 212

Book review by WOJCIECH SOŃTA*

https://doi.org/10.3326/pse.47.2.6

** Received: September 18, 2022
Accepted: October 11, 2022

Wojciech SOŃTA, dr hab., Assoc. Prof.
Kazimierz Pulaski University of Humanities and Technology in Radom, Bolesława Chrobrego 27, 26-600 Radom, Poland
e-mail: wojciechsonta@o2.pl
ORCiD: 0000-0003-3137-3646

PUBLIC SECTOR
ECONOMICS
47 (2) 281-284 (2023)

The budgetary procedure for local government units in Slovakia and Czechia is strictly defined in the legislation. Nevertheless, some local governments allocate earmarked funds as the basis for participatory budgeting. A participatory budget is the result of public consultations about how to spend part of a local government budget on proposals for local projects that have been put forward by the citizens and that are within the area of public services provided by the local government. Within the framework of the participatory budget, residents can personally contribute to the selection of investments that they think should be implemented within the portion of the local budget available to them. The textbook "Participatory budgeting (contexts, models and practical experience), 2nd extended edition" written by Klimovský et al., discusses the current issue of participatory budgeting from both theoretical and practical points of view. It builds upon the first edition of the same title published in 2021; both editions are the results of work on the research project "Innovations in Local Government Budgeting in Slovakia" (INLOGOB) supported by the Slovak Research and Development Agency [APVV-19-0108].

Although the textbook is in Slovak and focuses on Czechia and Slovakia, the text is comprehensible for all Slavic language speakers (I am from Poland and had no problems reading the book). Given the importance of the topic and the relevance of participatory budgeting in other Central and Eastern Europe (CEE) countries, the book could be of interest to a wider audience than comprised by Slovaks and Czechs.

The introduction of the textbook shows that public sector economics is an inherent feature of democracy and pluralism in every sense. One of the most important features of postmodern culture is manifested in the individualisation of the product, participatory budgeting. To conclude the above arguments, it must be stated that participatory budgeting is a developing financial trend because it is an instrument of civic engagement. The introduction of this edition also describes the atmosphere in which the textbook was created (COVID-19 and the war in Ukraine). These crisis situations have affected local budgets, in which resources have been reallocated to purposes for which they were not originally planned. In this situation, participatory budgets are in many cases limited, but still work.

Particular attention should be paid to the three main pillars of public choice theory: (1) democracy, (2) political pluralism, and (3) economy of the public sector.

All three pillars are supported by participatory budgeting. If any one of the pillars is threatened, it will be difficult to design and implement a participatory budget. I would also add a fourth pillar, the rule of law, because it is often the case that democracy and political pluralism are generally respected, while respect for the law is deficitary. However, there is no doubt that both the first and the second pillar have a decisive impact on the public sector economy.

The structure of the textbook consists of seven chapters. The first chapter, "Civil society and the third sector as drivers of public participation" written by Klimovský and Murray Svidroňová, draws the reader's attention to the fact that, in general, everyone performs several roles in parallel throughout their lives, without always realising it. Understanding the phenomena is important in the context of the strengths and weaknesses of innovations promoting the quality of democracy as one of the pillars of participatory budgeting.

The second chapter is entitled "Public participation and selected contexts", written by Bakoš, Klimovský and Murray Svidroňová. The authors refer to the community in every sense of the word and also to the family, the commune, the state, religious organisations and civil associations. In the context of this approach, the aim of the chapter is to explain the importance of public participation in public finance and to define the concepts of co-creation and co-production.

The third chapter, "Participatory budgeting as a tool for public participation" written by Bakoš, Klimovský and Kukučková, presents the organisational, substantive and procedural challenges of participatory budgeting and explains the basic mechanisms of its implementation and operation.

In the fourth chapter entitled "Perceptions of participatory budgeting and its barriers" (by Balážová, Bakoš, Gašparík, Klimovský and Kukučková) the basic intention is not only to introduce the reader to the processes related to participatory budgeting, but also to reveal the barriers and limitations associated with its implementation.

The next chapter "Practical experience with participatory budgeting" written by Balážová, Bakoš, Murray Svidroňová and Klimovský, presents the development and diversity related to the implementation and further realisation of participatory budgeting in the world, with a focus on the countries of Central and Eastern Europe, including Croatia, Germany, France and Poland. Through the presentation of different models and approaches, the reader has the opportunity to confront the practical experience with a more theoretical approach.

In the sixth chapter "Participatory Budgeting in Czechia and Slovakia" by Balážová, Bardovič, and Gašparík, the national (Slovak and Czech) experiences with participatory budgeting are evaluated. This chapter allows the reader not only to perceive how the theoretical definition of participatory budgeting is applied in national policies, but also to compare it with its development in other countries.

The last chapter, "Selected examples of participatory budget initiatives in Czechia and Slovakia", written by Balážová, Bardovič, and Gašparík, presents a closer observation of the implementation and use of participatory budgeting in selected regional and local governments.

Each chapter ends with questions and tasks that allow the reader (student) to review the material covered and, in case of any doubts about the correctness of the

PUBLIC SECTOR
ECONOMICS
47 (2) 281-284 (2023)

answer, to re-familiarise themselves with it and dispel any doubts. This is a typical procedure for such editions.

Although the textbook is intended for students of various economics and administration courses and for those who deal with participatory budgeting in practice, it may be an asset for Czech and Slovak policy makers, worthy of further in-depth study because the two countries face similar socio-economic realities, are both in the aftermath of political transformation and were admitted to the EU at the same time and both must make effective use of the structural funds, while Czechia faces the particular challenge of joining the eurozone.

It is also worth noting that both countries can be classified as countries with moderate involvement in innovation projects, so the manifestation of any novelty deserves attention. A street survey showed that cities that had introduced participatory budgeting have changed their appearance for the better, resulting in a better quality of life and higher citizen satisfaction after the two countries joined the EU. The change has been driven by projects implemented with resources from both the EU funds and participatory budgets. However, significant parts of the projects implemented were related to education, which indirectly influences the image of the city. Spending on infrastructure, revitalisation, urban transport and water management has played a significant role in this process. This is important information that affects the functioning of any municipality.

Numerous tables, figures, graphs and examples in the textbook complement and illustrate the content of each chapter. The list of references has 489 entries. In most cases, these are articles published within a not very long time frame, given that the institution of participatory budgeting (and logically the literature covering the issue) is relatively young. It is noteworthy that most of the cited sources are listed in world databases (WOS, Scopus).

Finally, the book was prepared by a group of authors from academic centres in Czechia and Slovakia. They represent different scientific fields and differ in experience and length of professional experience, as well as in the rank attained in the scientific hierarchy, but they all have a wealth of experience in participatory budgeting research.

All of the above have resulted in a textbook that comprehensively and exhaustively highlights the development and current issues of participatory budgeting.

Institute of
Public Finance